



HAL
open science

The spatial regulation of mRNA metabolism : mechanisms and consequences of targeting to nuclear pore complexes

Arianna Penzo

► **To cite this version:**

Arianna Penzo. The spatial regulation of mRNA metabolism : mechanisms and consequences of targeting to nuclear pore complexes. Genetics. Université Paris Cité, 2022. English. NNT : 2022UNIP5248 . tel-04747981

HAL Id: tel-04747981

<https://theses.hal.science/tel-04747981v1>

Submitted on 22 Oct 2024

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

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

Université Paris Cité
École doctorale Bio Sorbonne Paris Cité (562)

Laboratoire **Institut Jacques Monod**

The spatial regulation of mRNA metabolism: mechanisms and consequences of targeting to nuclear pore complexes

Par **Arianna PENZO**

Thèse de Doctorat en **Génétique**
Dirigée par **Benoît PALANCADE**

Présentée et soutenue publiquement le 26 Septembre 2022

Devant un jury composé de :

Dr. **Emmanuelle FABRE**, Directrice de recherche, Université Paris Cité, Présidente

Dr. **Manuel MENDOZA**, Directeur de Recherche, IGBMC Strasbourg, Rapporteur

Dr. **Vincent VANOOSTHUYSE**, Directeur de Recherche, ENS Lyon, Rapporteur

Dr. **Sophie POLO**, Directrice de recherche, Université Paris Cité, Examinatrice

Dr. **Mathieu ROUGEMAILLE**, Chargé de recherche HDR, Université Paris-Saclay, Examinateur

Dr. **Benoît PALANCADE**, Directeur de Recherche, Université Paris Cité, Directeur de thèse

*Rem tene
verba sequentur¹*

¹ *Get hold of the matter, the words will come of themselves.* Cato, the Elder

ABSTRACT

The spatial regulation of mRNA metabolism: mechanisms and consequences of targeting to nuclear pore complexes

The nuclear pore complex (NPC) is responsible for the exchanges between the nucleus and the cytoplasm. NPCs are remarkable for the sophistication and the conservation of their structure, but also for the wide variety of biological processes in which they are involved, notably gene expression regulation and the maintenance of genome integrity.

During my thesis, I was interested in deciphering the mechanisms responsible for the spatial restriction of mRNA-related processes to the NPC, and the consequences of this localization for cell homeostasis. On the cytoplasmic side, I contributed to describe how mRNAs coding for certain nuclear proteins are associated with the pore itself as soon as they are engaged in the translation process, and how localized translation prevents nascent protein aggregation. On the nuclear side, I have shown that gene relocation to the pore can be triggered by the formation of R-loops, which are co-transcriptional, potentially genotoxic DNA:RNA hybrids. In this context, I characterised the mechanism of relocation, which requires ssDNA coating by RPA and SUMOylation events, and gathered evidence supporting a protective effect of the nuclear pore environment against R-loop dependent genetic instability.

In both cases, the characterisation of the mechanisms and consequences of localized transcription or translation revealed that the proper recognition of potentially toxic molecules and their repositioning at the NPC likely allows their appropriate processing, avoiding the accumulation of protein aggregates in the cytoplasm or genotoxic structures in the nucleus. These results highlight the fundamental role of the NPC in the constitution of the nuclear proteome and the maintenance of genome homeostasis.

Keywords: mRNA localisation, localised translation, nuclear pore complex, R-loops, RPA, SUMOylation, transcription dependent genetic instability, *Saccharomyces cerevisiae*

RÉSUMÉ

Étude de la régulation spatiale du métabolisme des ARN messagers : mécanismes et conséquences du ciblage aux pores nucléaires.

Le complexe du pore nucléaire (NPC) est responsable des échanges entre le noyau et le cytoplasme. Ces structures sont remarquables par leur sophistication et leur conservation, mais aussi par la grande variété de processus biologiques dans lesquels elles sont impliquées, notamment la régulation de l'expression des gènes et le maintien de l'intégrité du génome.

Au cours de ma thèse, je me suis intéressée à déterminer les mécanismes par lesquels certaines étapes du métabolisme des ARN messagers se déroulent aux pores nucléaires, et les conséquences de cette localisation sur l'homéostasie cellulaire. Du côté cytoplasmique, j'ai contribué à décrire comment les ARNm codant pour certaines protéines nucléaires sont associés au pore lui-même dès qu'ils sont engagés dans le processus de traduction, et comment la traduction localisée empêche l'agrégation des protéines naissantes. Du côté nucléaire, j'ai montré que la relocalisation des gènes vers le pore peut être déclenchée par la formation de R-loops, des hybrides ADN:ARN co-transcriptionnels potentiellement génotoxiques. Dans ce contexte, j'ai caractérisé le mécanisme de relocalisation, qui nécessite la liaison de l'ADN simple brin par RPA et des événements de SUMOylation, et obtenu des données en faveur d'un effet protecteur de l'environnement du pore nucléaire contre l'instabilité génétique dépendante des R-loops.

Dans les deux cas, la caractérisation des mécanismes et des conséquences de la transcription ou de la traduction localisée a révélé que la reconnaissance de molécules potentiellement toxiques et leur repositionnement au niveau du NPC permet vraisemblablement leur prise en charge, évitant l'accumulation d'agrégats protéiques dans le cytoplasme ou de structures génotoxiques dans le noyau. Ces résultats soulignent le rôle fondamental des pores nucléaires dans la constitution du protéome nucléaire et le maintien de l'homéostasie du génome.

Mots-clés: Localisation des ARN messagers, traduction localisée, pore nucléaire, R-loop, RPA, SUMOylation, instabilité génétique dépendant de la transcription, *Saccharomyces cerevisiae*

RÉSUMÉ DÉTAILLÉ

Le complexe du pore nucléaire (NPC) est une structure multiprotéique hautement conservée au cours de l'évolution. Les pores nucléaires sont composés de plusieurs copies d'environ 30 sous-unités, appelées nucléoporines, organisées en plusieurs sous-complexes assemblés pour former un canal central, flanqué de structures filamenteuses des deux côtés de l'enveloppe nucléaire. La coordination de l'assemblage d'un complexe de cette taille et de cette stœchiométrie nécessite un réglage précis de la production de ses sous-unités, notamment en préparation pour la division cellulaire, ou suite à des dommages affectant son intégrité.

Le rôle principal des pores nucléaires est de réguler les échanges de macromolécules entre le noyau et le cytoplasme, notamment de par leur association dynamique avec des récepteurs de transport tels que les karyophérines. Toutefois, plusieurs nucléoporines sont connues pour être également impliquées dans le métabolisme nucléaire, la régulation transcriptionnelle et la maintenance du génome. Ce répertoire varié de fonctions est rendu possible par le grand nombre d'interactions que les sous-unités de cette structure remarquable sont capables d'établir. Notamment, parmi les facteurs dont la localisation est limitée à l'environnement du pore nucléaire figurent plusieurs enzymes impliquées dans les processus de SUMOylation, une modification post-traductionnelle réversible impliquant le couplage covalent du polypeptide SUMO (*small ubiquitin-like modifier*) à ses protéines cibles, et ayant une grande variété de conséquences fonctionnelles, en particulier la modulation des interactions protéine-protéine. Le pore nucléaire constitue ainsi une plateforme dynamique et polyvalente vers laquelle convergent une grande diversité de molécules, permettant de ce fait la régulation et la coordination temporelle et spatiale de nombreux processus biologiques fondamentaux. Si certaines nucléoporines exercent de telles fonctions dans le nucléoplasme, ces régulations se produisent généralement au niveau de l'enveloppe nucléaire, comme montré chez la levure *S. cerevisiae*, où de nombreuses études ont constaté que les NPCs interagissaient à la fois avec des gènes fortement exprimés ou des loci endommagés (présentant des cassures double brin ou des fourches de réplication bloquées), influençant localement la transcription ou la résolution de structures génotoxiques.

Parmi les facteurs susceptibles de menacer à la fois l'expression et la stabilité du génome figure l'accumulation de « R-loops », des structures comprenant un hybride ADN:ARN formé entre l'ARN naissant et son brin d'ADN matrice, et un simple brin d'ADN déplacé. Malgré l'existence de nombreux processus biologiques nécessitant la formation de R-loops, leur accumulation excessive peut interférer avec la transcription et la réplication, provoquant des défauts dans l'expression des gènes, une augmentation des taux de mutagenèse et de recombinaison, et une perturbation de l'épigénome. Cependant, les mécanismes par lesquels ces structures sont détectées et résolues dans l'environnement nucléaire demeurent peu connus.

Au cours de ma thèse, je me suis consacrée à l'étude des processus biologiques spatialement restreints au pore nucléaire, avec un intérêt particulier pour le ciblage des structures contenant des ARN messagers (ARNm), tant du côté nucléoplasmique que cytoplasmique.

Dans le cadre d'une étude sur la régulation co-traductionnelle de l'assemblage du NPC, mon laboratoire d'accueil avait identifié deux ARNm, codant pour des nucléoporines de la face nucléaire, montrant une interaction stable avec le pore une fois engagés dans le processus de traduction. Cette interaction dépend de l'association des ARNm avec les ribosomes, et de l'interaction des karyophérines avec l'extrémité N-

terminale de la protéine naissante. Ma contribution à cette étude a été de mettre en place des approches génétiques pour déterminer l'importance physiologique de la traduction localisée de ces ARNm au pore nucléaire. Dans ce but, j'ai employé deux stratégies complémentaires pour induire des événements de traduction de la protéine Nup1 à distance de son site de traduction préférentiel, afin de pouvoir évaluer les conséquences d'une telle perturbation sur l'homéostasie cellulaire.

D'une part, nous avons fusionné la région codante de *NUP1* avec la séquence terminatrice de l'ARNm *ASH1*, ce qui est suffisant pour induire une distribution asymétrique des ARNm chimères à l'extrémité du bourgeon des cellules de levure en division. D'autre part, nous avons inactivé le répresseur traductionnel Hek2, précédemment identifié par la laboratoire comme contrôlant la traduction de l'ARNm *NUP1*, augmentant ainsi le nombre d'ARNm engagés dans la traduction sans affecter leur association au NPC. De manière frappante, ces deux situations aboutissent à la formation d'agrégats cytoplasmiques de protéines Nup1. Ces résultats soutiennent l'hypothèse selon laquelle la traduction localisée au NPC aurait un rôle important dans la prévention de l'agrégation cytoplasmique des protéines, en assurant le transfert immédiat de protéines potentiellement nocives vers leur localisation finale. Ce processus serait particulièrement critique pour les Nups à répétitions phenylalanine-glycine, telles que Nup1, en raison de leurs nombreux domaines hydrophobes. À l'avenir, il sera intéressant d'étudier comment la traduction localisée peut être régulée dans différentes circonstances, notamment au cours du cycle cellulaire, en réponse à un stress ou en cas de dommages au NPC. De plus, nous avons identifié d'autres protéines nucléaires dont les ARNm s'associent au pore. Étendre notre étude à ces ARNm pourrait permettre de mieux comprendre les mécanismes moléculaires régissant le phénomène de traduction localisée au pore et leur importance pour d'autres aspects de l'homéostasie nucléaire.

Du côté nucléaire, je me suis concentrée sur les mécanismes d'interaction du pore nucléaire avec la chromatine transcriptionnellement active. Compte tenu de la dépendance de la formation des R-loops à l'activité transcriptionnelle, et de leur génotoxicité, nous nous sommes demandé si leur accumulation pouvait aussi déclencher le repositionnement des gènes au NPC, à l'instar des processus déjà décrits pour d'autres situations de dommages à l'ADN.

Pour tester cette hypothèse, nous avons dans un premier temps comparé la localisation des sites de contact entre la chromatine et les NPCs aux cartes disponibles de distribution des hybrides ADN:ARN. Cette analyse a révélé une corrélation entre la capacité des gènes à former des R-loops et leur association avec les NPCs. Pour évaluer directement si la formation d'hybrides déclenche la localisation des gènes aux NPCs, nous avons ensuite induit l'accumulation de R-loops sur deux loci modèles: le rapporteur *YAT1*, inductible par le galactose et inséré ectopiquement dans le génome, et le gène endogène *HSP104*, répondant au choc thermique. Dans ce cadre, la formation de R-loops a été induite par l'inactivation du complexe THO, responsable du recrutement co-transcriptionnel de différents facteurs associés à l'ARN, et réprimée en *cis* par l'insertion d'un intron, comme précédemment établi par le laboratoire d'accueil, ou en *trans* par la surexpression de la RNase H, une enzyme ciblant spécifiquement les hybrides ADN:ARN. En combinant différentes approches biochimiques et de microscopie, nous avons montré que ces gènes modèles migrent près des NPCs lorsque leur transcription est induite, et ce de manière dépendante de la formation des R-loops.

Ensuite, par le biais d'une approche candidat, j'ai caractérisé les facteurs requis pour la reconnaissance, le ciblage et l'interaction physique des R-loops avec le NPC. J'ai identifié RPA (*Replication protein A*), un complexe protéique reconnaissant l'ADN simple brin, comme le facteur de détection clé déclenchant la

relocalisation des R-loops. Le processus de relocalisation nécessite également des événements de mono-SUMOylation, spécifiquement médiés par la SUMO-ligase Mms21, sous-unité du complexe Smc5/6, et ciblant notamment RPA. Nos données suggèrent que la présence de RPA sur l'ADN et sa SUMOylation constituent ainsi une "étiquette" de reconnaissance qui rend la R-loop compétente pour l'établissement d'interactions physiques avec les NPCs. Dans ce cadre, nous avons pu montrer que l'ancrage au pore des R-loops liées par des protéines SUMOylées implique des facteurs contenant des motifs de reconnaissance du polypeptide SUMO, tels que Slx5/8, une ubiquitine ligase associée au NPC. Il reste à déterminer si d'autres protéines également SUMOylées sont aussi reconnus par Slx5/8 et contribuent également à l'ancrage des R-loops aux pores.

Pour déterminer les conséquences physiologiques de l'association des gènes formant des R-loops aux NPCs, nous avons utilisé plusieurs approches génétiques. D'une part, nous avons empêché le repositionnement des R-loops aux NPCs en interférant avec la voie RPA/Mms21, et montré que ce processus est essentiel dans des mutants formant des hybrides. D'autre part, nous avons utilisé un système d'ancrage constitutif pour forcer l'association d'un gène formant des R-loops au NPC, et observé une diminution de l'instabilité génétique associée aux hybrides, suggérant une action protectrice du pore nucléaire contre la toxicité de ces structures. Les mécanismes par lesquels l'environnement du NPC influence ainsi le métabolisme des R-loops restent en revanche à déterminer. La proximité du pore pourrait permettre à l'ARNm de s'engager plus rapidement dans la voie d'export, facilitant son éviction du site de transcription et empêchant ainsi la formation de R-loops. Alternativement, l'association avec les NPCs pourrait donner accès à des enzymes dédiées à la résolution des R-loops, ou encore faciliter l'élimination de protéines liées aux R-loops et pouvant stabiliser les hybrides.

À ce stade, il serait essentiel d'évaluer si le phénotype de relocalisation peut être généralisé à d'autres mutants accumulant des R-loops, et de mettre en place les outils appropriés pour pouvoir détecter et quantifier directement la formation d'hybrides aux loci d'intérêt. En ce qui concerne le mécanisme sous-jacent à la relocalisation, il sera aussi important de réconcilier nos résultats avec des études précédentes suggérant l'implication de facteurs de maturation de l'extrémité 3' des ARNm. De manière intéressante, nos observations préliminaires montrent que la relocalisation de loci modèles peut être affectée par l'accumulation d'un excès d'ARN dans la cellule, un phénotype majeur des mutants des facteurs de maturation de l'ARN.

Un point important de notre étude a été la comparaison du mécanisme de relocalisation des gènes formant des R-loops, que nous avons identifié dans le cadre de ce travail, avec les autres voies de relocalisation déjà connues dans la littérature. Dans ce cadre, nous avons pu montrer que l'association des R-loops au NPC implique un mécanisme distinct de ceux requis pour le repositionnement des gènes fortement transcrits ou des loci endommagés. En effet, nous avons pu notamment détecter l'interaction *HSP104*-NPC dans des conditions où ces autres voies sont abolies, i.e. en cas d'inactivation de facteurs de transcription et d'export, en empêchant la conversion de la R-loop en cassure double-brin, ou en l'absence de réplication. Cependant, le mécanisme de repositionnement que nous avons caractérisé met en jeu des acteurs communs aux voies de relocalisation dépendantes de la transcription ou des dommages à l'ADN, notamment de par l'implication des modifications par SUMO dans ces différentes situations. Les SUMO-ligases mises en jeu et la nature des protéines SUMOylées pourraient alors permettre de distinguer les différents types de loci et structures ainsi relocalisées aux NPCs.

En conclusion, la caractérisation des mécanismes de transcription ou traduction localisée aux pores a révélé que la reconnaissance de molécules potentiellement toxiques et leur repositionnement au niveau du NPC permet vraisemblablement leur prise en charge, évitant ainsi l'accumulation d'agrégats protéiques dans le cytoplasme ou de structures génotoxiques dans le noyau. Ces résultats soulignent le rôle fondamental des pores nucléaires dans la constitution du protéome nucléaire et le maintien de l'homéostasie du génome.

TABLE OF CONTENT

ABSTRACT	3
RÉSUMÉ	4
RÉSUMÉ DÉTAILLÉ	7
TABLE OF CONTENT	9
LIST OF FIGURES AND TABLES	13
LIST OF ABBREVIATIONS	15
INTRODUCTION	16
1 Structure and functions of the nuclear pore complex	17
1.1 Anatomy of the nuclear pore complex	18
1.1.1 Structural components of the NPC	18
Variability in the NPC architecture	19
Post-translational modifications of nucleoporins	20
NPC-associated factors	21
<i>NPC as a hub for SUMOylation events</i>	23
1.1.2 Biogenesis and turnover of nuclear pore complexes	24
Nuclear pore assembly	24
Nuclear pore maintenance and quality control	25
Translational regulation of NPC production	26
1.1.3 Transport through the nuclear pore complex	27
Protein transit through the NPC	27
mRNA export	28
<i>mRNA quality control</i>	29
1.1.4 Transport-unrelated roles of nucleoporins	30
Moonlighting nucleoporins	30
1.2 The nuclear pore and transcription regulation	31
1.2.1 Strategies for the study of gene relocation to the periphery	32
<i>In vivo</i> single-locus tracking by microscopy	32
Biochemical approaches	34
<i>ChIP</i>	34
<i>ChEC</i>	34
<i>Chromatin Fractionation</i>	34
1.2.2 Dynamics of loci diffusion in the nucleus	35
1.2.3 The nuclear pore complex as a boundary between silenced and active chromatin	36
1.2.4 Mechanisms mediating the interaction of active genes to the nuclear pore	37
Promoter activation driven model	39
Promoter de-repression driven model	41
A role for mRNA associated proteins in gene-NPC association	42
1.2.5 Consequences of NPC association on gene expression	43
1.2.6 Nucleoporins contribution to gene regulation	44

1.2.7 Nuclear pore association and transcriptional memory	46
1.3 The nuclear pore and maintenance of genome stability	48
1.3.1 Strategies to investigate damage compartmentalisation	49
DSBs generation	49
Replication stress induction	49
Telomere erosion	50
1.3.2 Chromatin dynamics upon DSB formation	51
1.3.3 Mechanisms driving damaged loci relocation to the nuclear periphery	52
Nucleoporins involved in lesion relocation	53
Persistent DSBs have multiple NE docking sites	54
1.3.4 Role of SUMOylation in lesion relocation and repair activation	54
Relevance of STUbL in peripheral anchoring and repair efficiency	55
1.3.5 Physiological consequences of lesion compartmentalisation at the nuclear periphery	56
2 Biogenesis, regulation and functional relevance of R-loops	59
2.1 Biogenesis of R-loops	61
2.1.1 R-loop prone sequences	61
R-loops and torsional stress	61
ssDNA lesions	61
2.1.2 R-loop 3D structures	62
2.1.3 Post-transcriptional R-loops	62
TERRA R-loops	62
lncRNA form R-loops in trans	62
CRISPR-Cas system	63
2.2 Strategies for R-loop detection	64
2.2.1 S9.6 antibody-based methods	64
DNA:RNA hybrid immunoprecipitation	64
Bisulfite treatment	65
2.2.2 RNase H based methods	66
mapR	66
R-loop CUT&TAG	66
RNase H CRAC	67
2.2.3 Caveats of R-loop detecting strategies	68
Differences between S9.6 and RNase H based strategies	68
R-loops fate during sample preparation	69
Aspecificity of the S9.6 antibody	69
Best practices to ensure detection of bona fide R-loops	70
2.2.4 R-loop levels manipulation strategies	70
2.3 Physiological roles of R-loops	72
2.3.1 R-loop in immunoglobulin class switch recombination	72
2.3.2 R-loop and Replication initiation	73
2.3.3 R-loops and transcription regulation	74
Transcription initiation	74

Transcription termination	75
2.3.4 R-loops in cell differentiation and development	75
2.4 Factors involved in R-loop control	76
2.4.1 <i>Sed quis custodiet ipsos custodes?</i>	76
R-loop associated chromatin marks	76
2.4.2 R-loop prevention	77
THO/TREX	77
Spliceosome	78
Topoisomerases	79
2.4.3 R-loop Removal	79
RNases H	80
<i>RNase H1</i>	80
<i>RNase H2</i>	80
Senataxin	81
Fanconi Anaemia Pathway	81
2.4.4 R-loop sensing	82
RPA	82
2.4.5 The expanding R-loop interactome	83
R-loop proteome studies	83
Indirect effects of R-loop resolving proteins	84
2.5 R-loop as a threat for genome stability	85
2.5.1 ssDNA modifying enzymes and nucleases	85
2.5.2 Structure-specific nucleases	86
2.5.3 Transcription replication conflicts	86
2.5.4 R-loops in DNA repair	87
Possible models for DNA:RNA hybrid or R-loop formation at DSBs	87
Physiological relevance of DSB DNA:RNA hybrids	88
2.5.5. Pathological R-loops	88
R-loop and cancer	88
R-loops in genetic diseases	89
<i>R-loops in repeat expansion disorders</i>	89
R-loops as therapeutic targets	89
OBJECTIVES OF THE STUDY AND BIOLOGICAL QUESTIONS	91
RESULTS	93
ARTICLE 1 Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis	94
Supplemental material	121
ARTICLE 2 A R-loop sensing pathway mediates the relocation of transcribed genes to nuclear pore complexes	139
Supplemental material	170

DISCUSSION	180
1 Mechanisms underlying compartmentalisation of mRNA-related processes at NPCs	181
1.1 Karyopherin-dependent localised translation at nuclear pores	181
1.2 R-loops as the signal for active genes relocation to the nuclear pore	182
1.2.1 R-loop gating in R-loop accumulating mutants	182
<i>HSP104</i> relocalisation in absence of RNase H	182
Role of Sen1 in R-loop gating	183
1.2.2 R-loop detection upon stress-induced transcription	183
1.3 R-loop sensing mechanisms: role of RPA and SUMOylation	185
1.3.1 RNA metabolism and R-loop sensing by RPA	185
Assessing the effect of excess RNA on R-loop gating	185
Tools to investigate the titration of gating factors by RNA	185
1.3.2 Is the Smc5/6 complex being specifically recruited at R-loops?	186
1.4 Factors mediating the interactions between SUMO-RPA-coated-R-loops and the nuclear pore	187
1.4.1 SIM-containing NPC-partners anchor R-loops to the nuclear pore	187
1.4.2 Additional intermediate factors reinforce R-loop-NPC association?	188
1.5 Comparison of NPC relocation pathways	189
1.5.1 How does R-loop gating relates to transcription-dependent relocation?	189
1.5.2 Does R-loop gating requires damage formation and checkpoint activation?	190
1.5.3 Common features of NPC relocation pathways	191
2 Functional consequences of gene / mRNA relocalisation to the pore	193
2.1 Physiological relevance of localised mRNA translation at the nuclear pore complex	193
2.2 Consequences of loci relocation to the nuclear pore for R-loop metabolism	194
2.2.1 Optimising the detection of R-loop dependent genetic instability	194
2.2.2 How does the nuclear pore act on R-loop metabolism?	194
Resolution of the pioneer R-loop?	194
Prevention of R-loop genotoxicity?	195
Prevention of further R-loop accumulation?	195
3 Conservation of relocation mechanisms across organisms	196
3.1 Conservation of R-loop gating	196
3.2 Conservation of localised translation at NPCs	197
BIBLIOGRAPHY	198

LIST OF FIGURES AND TABLES

INTRODUCTION

Figure 1	Electron micrographs of nuclear pore complexes	17
Figure 2	Anatomy of the nuclear pore complex	19
Figure 3	Schematic representation of the SUMOylation pathway	22
Figure 4	Schematic representation of protein transit through the NPC	27
Figure 5	Schematic representation of export of an mRNAs through the NPC	29
Figure 6	Overview of the methodologies used to determine loci localisation relatively to the nuclear pore in <i>S. cerevisiae</i>	33
Figure 7	Schematic representation of the proposed models for active loci relocation to the nuclear periphery.	40
Figure 8	Strategies to induce genetic instability	50
Figure 9	Schematic representation of challenged loci relocation to the nuclear periphery described in budding yeast	53
Figure 10	schematic representation of an R-loop	59
Figure 11	Schematic representation of S9.6-based R-loop detection approaches:	65
Figure 12	Schematic representation of RNAse H-based R-loop detection meeting	67
Figure 13	Schematic representation of the “collapsed R-loop” model explaining the mechanism of transcription dependent AID-mediated CSR	72
Figure 14	Factors implicated in the prevention of R-loop formation	77
Figure 15	Schematic representation of the main factors involved in R-loop removal	79
Figure 16	Schematic representation of the factors involved in R-loop processing into DSBs	85
Table 1	SUMO pathway components	22
Table 2	Summary of the major findings investigating the sequence and protein determinants required for inducible loci relocation to the nuclear pore	38

RESULTS

ARTICLE 1

Figure 1	Distinct modes of co-translational interactions between NPC subunits	117
Figure 2	Translation-dependent targeting of mRNAs encoding NPC-bound proteins	118
Figure 3	Karyopherin-mediated recognition of nascent proteins mediates their co-translational association with NPCs	119
Figure 4	Uncoupling co-translational interactions affects nucleoporin homeostasis	120

Figure S1	Control experiments related to the co-translational interaction screen	122
Figure S2	Identification and characterization of NPC-associated mRNAs	123
Figure S3	Mechanisms of targeting of NPC-bound mRNAs	125
Figure S4	Co-translational interactions and nucleoporin homeostasis	127

ARTICLE 2

Figure 1	R-loops can be a signal for repositioning to NPCs	144
Figure 2	Heat shock-induced transcriptional activation leads to R-loop dependent relocalization to the NPC	148
Figure 3	R-loop-dependent relocalization to NPCs requires ssDNA coating by RPA	150
Figure 4	The SUMOylation pathway is involved in R-loop-dependent repositioning to NPCs	154
Figure 5	NPC association alleviates R-loop toxicity	156
Figure S1	Validation of the gene datasets and reporter systems used to analyze the relationships between NPC association and R-loop levels	171
Figure S2	RNase H1 over-expression suppresses HSP104 co-fractionation with NPCs	172
Figure S3	RPA ChIP-seq analysis	173
Figure S4	NPC association alleviates R-loop toxicity	174

DISCUSSION

Figure 1	<i>HSP104</i> co-fractionation with the nuclear pore in absence of RNase H1 and 2	182
Figure 2	<i>HSP104</i> transcript levels in wt and <i>tho</i> mutant	184
Figure 3	Tools to investigate titration of gating factors by RNA	186
Figure 4	<i>HSP104</i> co-fractionation with the nuclear pore in mutants affecting Smc5/6 complex recruitment to chromatin	187
Figure 5	<i>HSP104</i> co-fractionation with the nuclear pore in a checkpoint mutant	190
Figure 6	General principle of nucleic acid relocation to the nuclear pore complex	192
Table 1	Comparison between the three proposed model for transcription-dependent relocation and the R-loop gating mechanism identified in this study	189
Table 2	Comparison between damage-dependent and R-loop-dependent relocation mechanisms identified in budding yeast	191

LIST OF ABBREVIATIONS

AID	Activation-induced cytidine Deaminase
ALT	Alternative Lengthening of Telomeres
BIR	Break Induced Replication
ChEC	Chromatin Endogenous Cleavage
CHIP	Chromatin ImmunoPrecipitation
CRAC	Crosslinking Analysis of cDNAs
CSR	Class Switch Recombination
CUT&RUN	Cleavage Under Targets and Release Using Nuclease
CUT&TAG	Cleavage Under Targets and TAGmentation
DRIP	DNA:RNA hybrid ImmunoPrecipitation
DSB	Double Strand Break
GCRs	Gross Chromosomal Rearrangements
GRS	Gene Recruitment Sequences
HR	Homologous Recombination
IR	Ionizing Radiation
MMEJ	Microhomology Mediated End Joining
mRNP	messenger RiboNucleoParticle
NHEJ	Non-Homologous End Joining
NPC	Nuclear Pore Complex
NLS	Nuclear Localisation Signal
ORF	Open Reading Frame
RFB	Replication Fork Barrier
SIM	SUMO-Interacting Motif
STUbL	SUMO-Targeted Ubiquitin Ligase
SUMO	Small-Ubiquitin-like Modifier
TAR	Transcription Associated Recombination
TBP	TATA Binding Protein
TNR	Triplet Nucleotide Repeats
UAS	Upstream Activation Sequence

INTRODUCTION

Chapter1

Structure and functions of the nuclear pore complex

A milestone in the evolution of life is the internal compartmentalisation of the cell into distinct organelles surrounded by membranes, in particular the nucleus, the defining feature of eukaryotic cells. The formation of the nuclear envelope led to the isolation of the genome from the cytoplasm and the subsequent spatial separation of the transcription and translation processes. Consequently, adequate communication between spatially isolated and functionally specialised compartments is essential to guarantee the proper coordination and synchronisation of cellular functions. To this aim, molecular trafficking between the nucleus and the cytoplasm strictly takes place via the nuclear pore complex, which finely regulates molecular flows in and out of the nucleus.

In the 1950s, pioneering electron microscopy observations of *Xenopus* oocyte nuclei described the porous nature of the nuclear envelope (Callan and Tomlin, 1950 - Fig. 1A), sparking the interest to decipher the morphology and function of the remarkable architecture that is the nuclear pore complex.

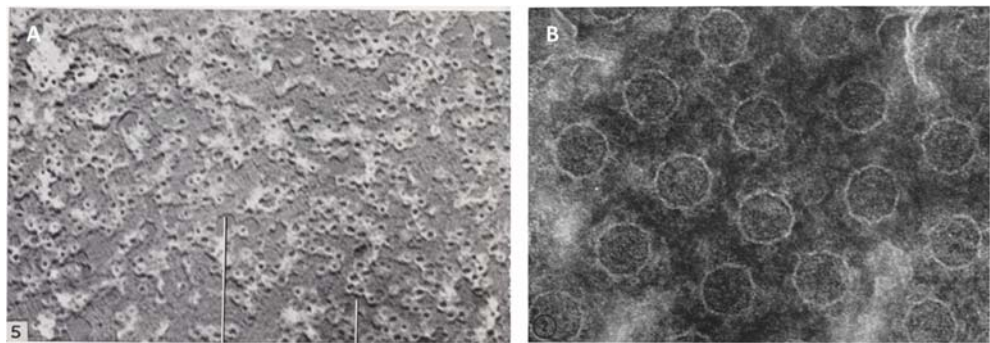


Figure 1. **Electron micrographs of nuclear pore complexes.** A. Nuclear envelope from oocyte of *Xenopus laevis* from Callan and Tomlin, 1950 (Magnified x 26,000). B. Nuclear envelope from oocyte of the newt *Triturus* from Gall, 1967 (Magnified x 200,000).

Over the years, technological advancement has refined our knowledge about the pore, enabling the description of its shape (Fig. 1B), distribution, composition, and conservation across organisms. To date, the nuclear pore structure has been deciphered at outstanding resolution, and numerous functions have been identified beyond its pivotal role in mediating nucleocytoplasmic molecular trafficking. The nuclear pore emerges as a major regulator of nuclear homeostasis, with important implications in gene expression, chromatin organisation, and maintenance of genome stability.

The following paragraphs summarise the current knowledge on the structure of the nuclear pore and its assembly and turnover dynamics. In addition to a synthetic description of transport processes through the pore, a major focus will be on the functions of the nuclear pore in nuclear organisation, in particular its interaction with active genes, its role in the regulation of transcription, and its contribution to the maintenance of genetic stability through the spatial regulation of repair pathways.

1.1 | Anatomy of the nuclear pore complex

Since the first isolation of yeast nuclear pores (Rout and Blobel, 1993) and the first proteomic analysis of nuclear pore structure in yeast (Rout et al., 2000) vertebrates (Cronshaw et al., 2002) and plants (Tamura et al., 2010), a great number of studies have joined efforts to decipher the subtleties of the nuclear pore structure and its functional implications. A wide variety of techniques have been employed, combining bottom-up and top-down approaches to depict a more and more clear portrait of this gigantic architecture. Recently, the combination of native NPC isolation with mass spectrometry, *in vivo* imaging, and cryo-electron tomography have allowed the dissection of the yeast nuclear pore anatomy at sub-nanometre resolution (Kim et al., 2018; Allegretti et al., 2020; Akey et al., 2022).

1.1.1 | Structural components of the NPC

The nuclear pore architecture is highly conserved throughout the eukaryotic tree. It is composed of around 30 subunits (the nucleoporins), each present in 8, 16 or 32 copies, for a total of over 500 molecules making up the complete structure. This megadalton sized assembly consists of several subcomplexes hinged together to form the eight spokes constituting the backbone of the channel, arranged in an eight-fold symmetry, with filamentous components protruding on both sides of the nuclear envelope.

Four transmembrane nucleoporins form the membrane ring (Fig. 2, *in orange*), which anchor the structure to the nuclear envelope and interconnect adjacent spokes, which compose the three latitudinal rings (Fig. 2, *red and blue modules*) and contribute to shape the membrane curvature. Flexible connectors (Fig. 2, *in pink*) hinge together the inner with the outer ring, welding all the major building blocks together.

The nucleoporins forming the central channel (Fig. 2, *in dark grey*) branch off towards the central axis, forming the hydrogel-like molecular sieve responsible for the formation of the permeability barrier. These nucleoporins are characterised by highly disordered domains containing Phenylalanine-Glycine (FG) – rich repeats, which can interact with several transport receptors to orchestrate the active transport of cargoes and maintain the permeability barrier through intra- and inter-molecule FG domain interactions. FG nucleoporins can be classified in 2 groups based on the type of FG repeats interspersed in their sequence: ‘Phe-X-Phe-Gly/Phe-Gly’ (FXFG/FG) containing nucleoporins are located at the cytoplasmic and nucleoplasmic entrance of the pore, regulating the access of macromolecules to the channel, while ‘Gly-Leu-Phe-Gly’ (GLFG) repeats are more concentrated within the central scaffold. GLFG nucleoporins have been proposed to be responsible for the passive permeability barrier (Strawn et al., 2004) and have been shown to participate in the formation and stability of the nuclear pore structure (Onischenko et al., 2017).

From the outer cytoplasmic ring, filamentous nucleoporins branch out towards the cytoplasm (Fig. 2, *in green*), forming a platform hosting the helicase Dbp5 (hsDDX19), important for the directionality of mRNA export. From the outer

nucleoplasmic ring, Mlp1 and Mlp2 (hsTPR) extends towards the chromatin, forming the nuclear basket (Fig. 2, *in purple*). Their anchoring to the NPC involves mainly Nup1 and Nup60 (hsNup153). The nuclear basket, beyond its role as a docking site for export-competent messenger RNA ribonucleoproteins (mRNPs), has been shown to harbour a dynamic network of interactions with a great number of factors, broadening the list of functions in which the nuclear pore is involved.

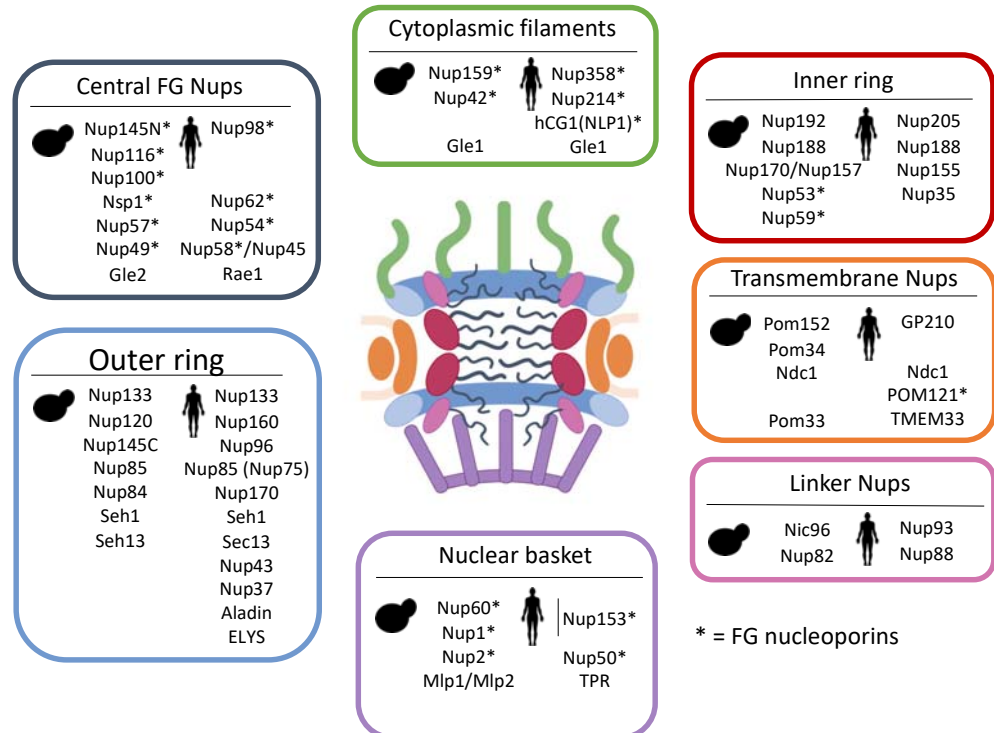


Figure 2. **Anatomy of the nuclear pore complex.** Schematic representation of the nuclear pore structure, indicating the position of nucleoporins in both yeast and mammalian nomenclatures. Adapted from Tingey et al., 2022.

Variability in the NPC architecture

Although the essence of the NPC architecture is extremely well conserved, numerous organism-specific and cell-type-specific variations can be observed, notably in the number, stoichiometry and organisation of the nucleoporins. The outer rings, for example, are constituted by head-to-tail interacting Y-shaped modules, which are heptameric in yeast (Nup84 complex) and decameric in mammals (Nup107/160 complex). Moreover, they show differential composition among different organisms: notably, each outer ring is composed of two layers of Y-complex rings in mammalian and *Xenopus* cells, only one layer is observed in budding yeast, while fission yeast and the algae *Chlamidomonas reinhardtii* have an asymmetric distribution, with a two-layered nucleoplasmic outer ring, and a single layered cytoplasmic ring (reviewed in Dultz et al., 2022).

The nuclear pore conformation can also change based on the life cycle and developmental stage of the cell: in exponentially growing Tobacco cells, NPCs appear to have a larger pore diameter compared to stationary phase or senescent cells (Fiserova et al., 2009). Furthermore, a study on NPC composition during development in *Caenorabditis elegans* showed how the expression of some nucleoporins is restricted at early developmental stages and appear absent in post-

mitotic cells. Moreover, while the scaffold nucleoporins display prolonged stability, more distal components can undergo deterioration over time and appear absent in aging non-dividing cells, leading to age-related leakiness of the pores (D'Angelo et al., 2009). Tissue-specific nucleoporin expression can also be observed in several organisms (reviewed in Capelson et al., 2010).

Finally, nuclear pore composition can vary even between complexes embedded in the same nuclear envelope; notably, in budding yeast, the nuclear pores facing the nucleolus are deprived of the nuclear basket nucleoporins Mlp1 and Mlp2 (Galy et al., 2004) and their associated factors. Moreover, recent structural studies identified a subset of NPC harbouring double-layered outer nucleoplasmic rings coexisting with canonical NPCs in *S. cerevisiae* (Akey et al., 2022).

Post-translational modifications of nucleoporins

Nucleoporins have been shown to undergo several kinds of post-translational modifications, which add additional layers of structural and functional variability.

A peculiar case of post-translational modification is the post-translational maturation of the yeast nucleoporin Nup145, which is cleaved into two distinct moieties: a carboxy-terminal domain (Nup145C) which assembles into the Nup84 complex, and an GLFG-rich amino-terminal domain (Nup145N) that has homology to Nup100 and Nup116 (Teixeira et al., 1997). The cleavage reaction is self-catalysed (Teixeira et al., 1999) and conserved: mammalian nucleoporins Nup98 and Nup96 are also produced from a unique 186kDa precursor molecule which self-cleaves in two distinct polypeptide moieties (Fontoura et al., 1999).

In metazoan, nucleoporins undergo phosphorylation by PLK1 and CDK1 at the onset of mitosis triggering nuclear envelope breakdown, and are later dephosphorylated to allow NPC reassembly during nucleus reformation (reviewed in Kutay et al., 2021). In yeast, phosphorylation of nuclear basket nucleoproteins has been shown to cause the detachment of active chromatin from the NPC during S-phase (Brickner and Brickner, 2010; Bermejo et al., 2011).

Around half of the nucleoporins can be mono/poly/multi-ubiquitylated, without major effects on protein stability (Hayakawa et al., 2012). In particular, Nup60 ubiquitination in response to genotoxic stress increases the strength of its interactions with its partner nucleoporins and participate to the DNA damage response (Niño et al., 2016). Moreover, SUMOylation of components of the nuclear basket has been shown to be stress- and cell cycle-dependent (Folz et al., 2019).

Protein O-glycosylation consists in the addition of a sugar moiety to the hydroxyl group of Serine or Threonine side chains and, at the NPC, has been shown to be important for stress tolerance and cell cycle progression (Li and Kohler, 2014). Specifically, addition of N-acetylglucosamine to FG-Nucleoporins has been shown to prevent nucleoporin degradation via the ubiquitin–proteasome pathway and contribute to preserve the permeability barrier (Zhu et al., 2016). Strikingly, an innovative optogenetic strategy allowing high-throughput quantification of the nuclear import and export kinetics in live human cells revealed the role of NPC glycosylation in enhancing active and passive transport in both directions, possibly by hindering the hydrophobic cohesions of FG domains (Yoo and Mitchison, 2021).

Finally, Nup60 acetylation has been shown to promote mRNA export by facilitating the enrichment of the TREX-2 subunit Sac3 at the nuclear basket (Gomar-Alba et al., 2022). Interestingly, the deacetylase Hos3 has been shown to transfer from the septins ring to the nuclear pore in dividing *S. cerevisiae* cells, mediating deacetylation of nuclear basket and central ring nucleoporins embedded in pores destined to be transferred to the daughter cell (Kumar et al., 2018). This event seems to be crucial to delay G1/S transition of the daughter cell, which typically experience longer G1 phases compared to the larger mother cell. S-phase Start delay is therefore achieved by modulating gene expression of cell cycle related genes and mRNA export. Notably, nucleoporin acetylation and association with acetylases and deacetylases has also been observed in mammalian cells, although their relevance in nuclear and cellular processes has yet to be elucidated (Gomar-Alba and Mendoza, 2019).

NPC-associated factors

The nuclear pore microenvironment is an extremely crowded space, with ≈ 1000 molecules estimated to transit through each pore every second (Ribbeck and Görlich, 2001). The nuclear pore is therefore constantly interacting with a wide variety of proteins, notably nuclear transport receptors and other factors involved in protein import and mRNA export.

Among the direct interactors of the yeast nuclear basket nucleoporins Mlp1 and Mlp2 it is important to mention the peripheral nuclear protein Esc1, involved in telomere silencing, NE structural organization and DNA repair (Niepel et al., 2013 and references therein). The nuclear basket also interacts with the proteasome, responsible for ubiquitin-mediated selective turnover of short-lived and misfolded proteins, and localised predominantly in the Nuclear Envelope-Endoplasmic Reticulum network in proliferating yeasts (Enekel et al., 1998).

In mammalian cells, Mad1 and Mad2 association to the nuclear basket is important for the correct function of the mitotic checkpoint, which delays sister chromatid separation until all the chromosomes are correctly aligned at the metaphase plate (Scott et al., 2005; Lee et al., 2008). Notably Mad1-TPR interaction is also important for the correct localization of Mad1 at kinetochores during mitosis (Lee et al., 2008).

In metazoans, structural support to the nucleus is provided by a tight fibrillary network composed of intermediate filaments (lamins) and membrane associated proteins that coat the inner face of the nuclear envelope. Strikingly, Nup153 has been shown to co-immunoprecipitate with LaminB and directly interact with it through its C-terminal domain. Moreover, disrupting lamina assembly leads to loss of Nup153 recruitment to the nuclear pore complex (Smythe et al., 2000). Conversely, depletion of Nup153 by RNAi resulted in alteration of the nuclear lamina organisation, and consequent rearrangements of the cytoskeleton (Zhou and Panté, 2010). Nuclear pores are therefore directly involved in the support network of the nuclear structure, which in turn may have consequences in their mobility and distribution along the membrane.

Box1 | SUMOylation pathway: an overview

SUMO (Small Ubiquitin-related MOdifier) is a post-translational protein modifier which can be covalently attached to lysine residues of an acceptor protein. This 11kDa SUMO polypeptide is ubiquitously expressed throughout the eukaryotic kingdom and appears structurally similar to Ubiquitin. All SUMO proteins are expressed as a precursor polypeptide and require the cleavage of a C-terminal portion by specific SUMO-proteases to complete maturation. SUMO particles can be added as a monomer in one (mono-SUMOylation) or multiple lysines (multi-SUMOylation) or can form chains (poly-SUMOylation) in which the SUMO particle is itself SUMOylated. Similarly to ubiquitylation, SUMOylation require an enzymatic cascade involving three enzymes: an E1 activating enzyme, an E2 conjugating enzyme and an E3 ligase (Fig. 3).

While only one E1 and E2 enzymes are known, several ligases have been identified (see Table 1 below). SUMOylation is a reversible modification that can alter protein localisation, activity or stability by reshaping its intra- or inter-molecular interactions.

The influence of SUMOylation on the target interactions results from the generation or the masking of interaction domains; consequence on protein stability may depend on the recognition of the SUMOylated substrate by SUMO-targeted Ubiquitin ligases (STUbLs), which target SUMOylated substrate to proteasome degradation.

SUMOylation is involved in a plethora of biological processes, and is essential in most eukaryotes, with the sole exception of fission yeast. SUMO ligases do not seem to recognise specific substrates, but mainly act by triggering waves of SUMOylation that can simultaneously target a group of conjugation-competent proteins located at their vicinity, favouring their interactions, as is the case for the assembly of the DNA repair machinery (Psakhye and Jentsch, 2012). This mode of action may explain the wide number of identified targets and assigned functions, which makes this dynamic modification a major regulator of cellular homeostasis (reviewed in Geiss-Friedlander and Melchior, 2007; Vertegaal, 2022).




			
SUMO particle	SUMO1–5	SUMO	Smt3
E1 – activating enzyme	SAE1–UBA2 dimer	Aos1–Uba2 dimer	Aos1–Uba2 dimer
E2 – conjugating enzyme	UBC9 (also known as UBE2I)	UBC9 (also known as Iwr)	Ubc9
E3 -SUMO ligases	NSE2, PIAS1–4, ZMIZ1, ZMIZ2, RANBP2	Qjt, Cerv, dPIAS/Su(var) 2-10	Mms21, Siz1, Siz2, Zip3 (meiosis-specific)
SUMO Proteases	SEN1–3, SEN5–7, DES11, DES12, USPL1	Ulp1, velo, pira	Ulp1, Ulp2
STUbLs	RNF4, RNF111	Dgrn	Slx5/8, Uls1, Rad18

Table 1 | SUMO pathway components

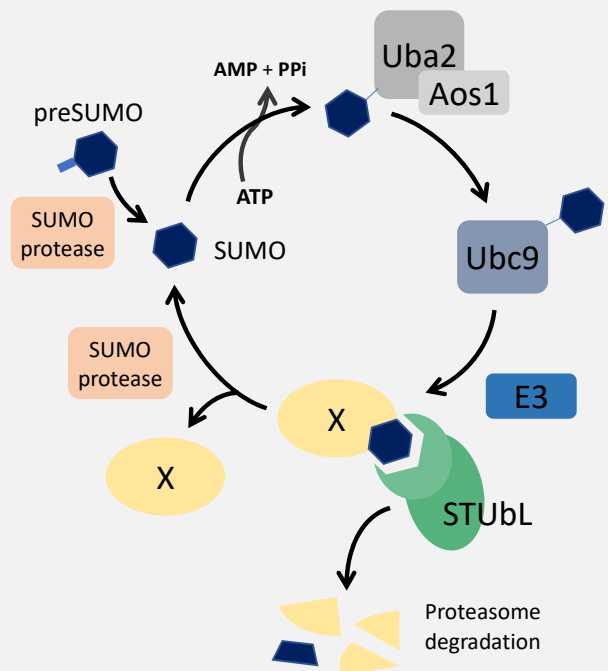


Figure 3. Schematic representation of the SUMOylation pathway

NPC as a hub for SUMOylation events

In addition to the aforementioned interactors, the nuclear pore has been shown to be a hub for factors involved in SUMOylation, a reversible post translational modification with a wide variety of functions, in particular the modulation of protein-protein interactions (see Box 1).

In vertebrate cells a sub-population of the SUMO-conjugating enzyme Ubc9 has been shown to stably localize at the nuclear pore, associated with the cytoplasmic filaments (Saitoh et al., 2002; Zhang et al., 2002) by interacting with Nup358/RanBP2, which displays SUMO-E3 ligase activity (Pichler et al., 2002). Together, they form a stable complex with SUMOylated RanGAP1 (Zhang et al., 2002).

The SUMO protease Ulp1 has also been shown to localize at the nuclear pore complex in budding yeast. Its interaction involves the nuclear basket nucleoporins Mlp1-2 (Zhao et al., 2004), Nup60, the outer ring Nup84 complex (Palancade et al., 2007) and the karyopherins Kap60 and Kap95 (Panse et al., 2003). The restriction of Ulp1 at the nuclear pore microenvironment is important to prevent damage formation (Palancade et al., 2007) and for transcription regulation (Texari et al., 2013 – see paragraphs 1.2.4 and 1.3.4). Consistent with the importance of such functions, interaction with the nuclear pore has been detected also for the *S. pombe*, *Drosophila*, *Arabidopsis* and mammalian orthologues (reviewed in Palancade and Doye, 2008).

The SUMO-targeted Ubiquitin ligase (STUbL) Slx5/8 has also been shown to associate with the nuclear pore by interacting with the Nup84 complex (Nagai et al., 2008). In budding yeast, Slx5 and Slx8 exist as a stable dimer harbouring several SUMO interacting motifs (SIM) for substrate recognition, with a preference for poly-SUMOylated factors, and RING domains that confer the ubiquitin ligase activity (li et al., 2007; Mullen and Brill, 2008; Xie et al., 2007). Mutants of this complex are synthetic lethal with the SUMO machinery and show accumulation of SUMO-conjugates, increased recombination rates and Gross Chromosomal Rearrangements (GCR), and loss of telomeric silencing (Wang et al., 2006; Zhang et al., 2006; li et al., 2007; Darst et al., 2008). The role of Slx5/8 in the maintenance of genetic stability is likely exerted through the targeting to proteasome degradation of poly-SUMOylated damage-related factors (further discussed in paragraph 1.3.4).

1.1.2 | Biogenesis and turnover of nuclear pore complexes

Nuclear pores are extremely long-lived structures, which need constant surveillance and maintenance to ensure their integrity and functionality throughout the life of the cell. Problematic or irreparably damaged structures must be quickly repaired or removed to ensure the maintenance of the permeability barrier and safeguard the nuclear environment. Furthermore, at each cell generation, new nuclear pores must be produced to provide the necessary amount for daughter cells and to meet the demands of nucleocytoplasmic exchange as the nucleus increases in volume during interphase. The number of nuclear pores doubles during interphase as the size of the nucleus increases, so that the overall pore density remains unchanged (D'Angelo et al., 2006).

Nuclear pore assembly

Two main mechanisms have been characterised for nuclear pore assembly: in organism with closed mitosis, e.g. budding yeast, the only assembly pathway possible consists in *de novo* production of NPCs and insertion in the NE. In cells with open mitosis, i.e. undergoing nuclear envelope breakdown, both NPC insertion in interphase and post-mitotic assembly are possible. Notably, the former is one order of magnitude slower than the latter, which require only around 10 minutes, and the order of assembly of the subunits is different between the two processes (Dultz and Ellenberg, 2010).

During interphase, pores form *de novo*, independently from the existent pores, and assembly happen through an inside-out mechanism, which require the Y-complex and its nuclear import (D'Angelo et al., 2006). Nuclear envelope associated proteins have been proposed to contribute to the bending of the membranes prior to fusion (Vitale et al., 2022). The mechanics of the membrane fusion process and how the permeability barrier is maintained during interphase NPC assembly, however, has not yet been fully elucidated.

The recent development of a bottom-up high throughput strategy based on metabolic labelling (KARMA - Kinetic Analysis of incorporation Rates in Macromolecular Assemblies; Onischenko et al., 2020) to assess the dynamics of NPC assembly in yeast showed how the biogenesis of the complex starts from the symmetrical core nucleoporins, followed by the asymmetrical ones, and finish with the assembly of the distal nuclear basket nucleoporins Mlp1/2.

During mitosis, nuclear envelope breakdown causes the disassembly of the NPC in smaller, soluble building blocks, triggered by PLK1 and CDK1 dependent nucleoporins phosphorylation. Once chromosome segregation is successfully achieved, nuclear reformation starts, and the now de-phosphorylated nucleoporins can reassemble, with a much rapid kinetic than the one scored for interphase insertion. It is still a subject of study whether the seeding structure for the reassembly are the core proteins which could remain associated to residual membrane structures, or the formation of Y complex rings recruited to chromatin by the DNA-binding nucleoporin ELYS, or a combination of both (reviewed in Kutay et al., 2021).

A third assembly mechanism has been characterised in embryonic cells, in which maternally provided NPCs, produced during oogenesis, are stocked in a subdomain of the ER consisting in stacked cytoplasmic membranes called *Annulate Lamellae* (AL). In *Drosophila* embryos, at the syncytial blastoderm stadium where transcription activity is still absent, AL-NPC, composed only of the scaffold, have been observed to progressively transfer to the nuclear envelope to maintain the right pore density along with nuclear growth and acquire transport competence only once correctly inserted at the NE (Hampoelz et al., 2019).

Nuclear pore maintenance and quality control

Given the long life of the NPC, the intense crowding around and through them, and the mechanical solicitation given by nuclear envelope movements, these structures are subjected to considerable strains that can lead to deterioration and structural damage in the absence of proper maintenance.

The modularity of the pore architecture provides at the same time robustness and rigidity to sustain the structure, while allowing a certain degree of flexibility to buffer mechanical solicitation. Indeed, recent studies in human and *S. Pombe* cells showed how the nuclear pore can dilate and constrict depending on the membrane tension of the NE (Schuller et al., 2021; Zimmerli et al., 2021), highlighting the flexibility of the pore structure in response to environmental cues. Moreover, nucleoporins show a certain degree of redundancy, which allows the structure to remain solid and continue to perform its functions even in the absence of some of its components. Indeed, 2/3 of the nucleoporins are non-essential in budding yeast and their removal has limited, if any, impact on the structure, assembly and turnover dynamics of the complex (Hakhverdyan et al., 2021), with the obvious exception of the anchoring of the most peripheral subunits. However, such flexibility is guaranteed only in the case of the absence of a single subunit, while double mutants, especially of functionally or structurally related nucleoporins often result in lethality.

Measurement of the dissociation dynamics of single nucleoporins showed how pore subunits are slowly but constantly replaced, with peripheral components being exchanged to a higher rate than the core nucleoporins (Rabut et al., 2004). Notably, the rate of exchange was shown to not be a mere reflection of the subunit position relatively to the structure, but to also correlate with the strength of the interactions with the neighbour nucleoporins (Hakhverdyan et al., 2021). Such steady and continuous turnover, although slow, may ensure resilience to structural damage and adaptation to environmental solicitations in both dividing and quiescent cells.

Interestingly, in budding yeast, a quality control mechanism ensure that only functional NPCs are transferred to the daughter cells: NPC lacking the nucleoporin Nsp1, a signal of precedent stress-induced NPC changes, are actively excluded from crossing the barrier at the bud neck (Colombi et al., 2013; Makio et al., 2013). The proportion of NPCs that are transferred to the daughter cells is anyway limited (around 15%), and the majority of but NPC are *de novo* inserted (Shcheprova et al., 2008). This results in a rejuvenated daughter cell, while the mother cell takes on the burden of the pre-existing and flawed pores, which will contribute to its ageing.

In mammalian cells, instead, a checkpoint mechanism prevents cell abscission in presence of abnormal NPCs, notably due to improper nuclear basket assembly (Mackay et al., 2010).

Finally, a surveillance mechanism involving the Endosomal Sorting Required for Transport Complex III (ESCRT-III), recruited by Inner Nuclear membrane proteins, has been shown to clear out defective NPC assembly intermediates (Webster et al., 2014). Moreover, a mechanism of autophagy-mediated NPC degradation has also been identified in stress conditions in budding yeast (Lee et al., 2020; Tomioka et al., 2020).

Translational regulation of NPC production

To coordinate the assembly of a complex of this size and stoichiometry, the production of the subunits must be finely regulated to cope with the need for *de novo* assemblies and to sustain the pace of regular maintenance.

Interestingly, little is known about the transcriptional regulation of nucleoporin-encoding genes. Although a downregulation of scaffold nucleoporin RNAs have been observed by RT-qPCR in non-dividing cells in adult *C. elegans* worms (D'Angelo et al., 2009), there are no information on the mechanism causing this decrease, nor are there any known general mechanisms for coordinating transcription of Nup mRNAs, such as dedicated transcription factors/repressors or self-sustaining feedback mechanisms.

Regarding translation, instead, recent advances has been made in elucidating nucleoporin translation dynamics and co-translational regulation of NPC assembly. Previous work from the lab showed how in budding yeast, a subset of mRNAs coding for nucleoporins displays a characteristic bimodal distribution in polysome profiling, indicating the existence of two distinct pools: a population of actively translated mRNAs, associated with heavy polysomes, and a larger fraction of untranslated molecules. The latter is bound by the translational repressor Hek2, whose activity is regulated by its SUMOylation. Strikingly, a feedback loop has been identified for the regulation of NPC mRNAs which is triggered by the mislocalisation and degradation of Ulp1, possibly following NPC structural damage. This results in accumulation of SUMOylated inactive Hek2, allowing the rapid production of new subunits from the available mRNAs already residing in the cytoplasm (Rouvière et al., 2018).

In addition, recent studies, included our own (Lautier, Penzo et al., 2021; Seidel et al., 2022), identified the existence of co-translational assembly events targeting a subset of nucleoporins, which appear to interact with the nascent polypeptide of their direct interactor while the latter is still engaged in the translation process. Such binary interactions are consistent with known reported interactions in the main pore structure, which could act as seeds for the formation of the subcomplexes.

Together, these two modes of regulation allow a fine coordination of the assembly process and ensure the proper, rapid supply of essential pore components for the quick production of new nuclear pores in case of need.

1.1.3 | Transport through the nuclear pore complex

The nuclear pore main role is to finely regulate the exchanges between the nucleus and the cytoplasm, acting as a gate that only selected molecules can penetrate, with a designated directionality.

Only molecule smaller than 30kDa/5nm can freely diffuse through the pore, while bigger molecules (proteins and RNA-containing particles) need to be escorted by dedicated transport receptors establishing interactions with the FG nucleoporins to be able to cross the passage. Strikingly, combined deletion of all the FG domains in asymmetric nucleoporins do not have major effect on transport across the NPC, while symmetric FG domains proved to be essential (Strawn et al., 2004). Moreover, deletion of specific combination of FG domains impacted transport of molecules escorted by different transport receptors, indicating the co-existence of distinct pathways (Strawn et al., 2004; Terry and Wentz, 2007).

DNA do not travel through the nuclear pore except for the import of plasmid or retrotransposon DNA, the entrance or exit of viral DNA in infected cells, and the export of by-products of DNA repair pathways, which have also been observed to leak to the cytoplasm in conditions of genotoxic stress (Ashkenazy-Titelman et al., 2020; Wolf et al., 2016).

Protein transit through the NPC

Protein can cross the channel in both directions, in association with transport receptors of the Importin- β family, which includes both importins and exportins, also known as karyopherins. Binding of the soluble receptors to their cargo is determined by their interaction with the small GTPase Ran, which can be bound to either GTP (Guanosine triphosphate) or GDP (guanosine diphosphate). The directionality of the transport is given by the establishment of a RanGTP gradient across the nuclear envelope, with RanGDP predominantly present in the cytoplasm while RanGTP is mainly nuclear.

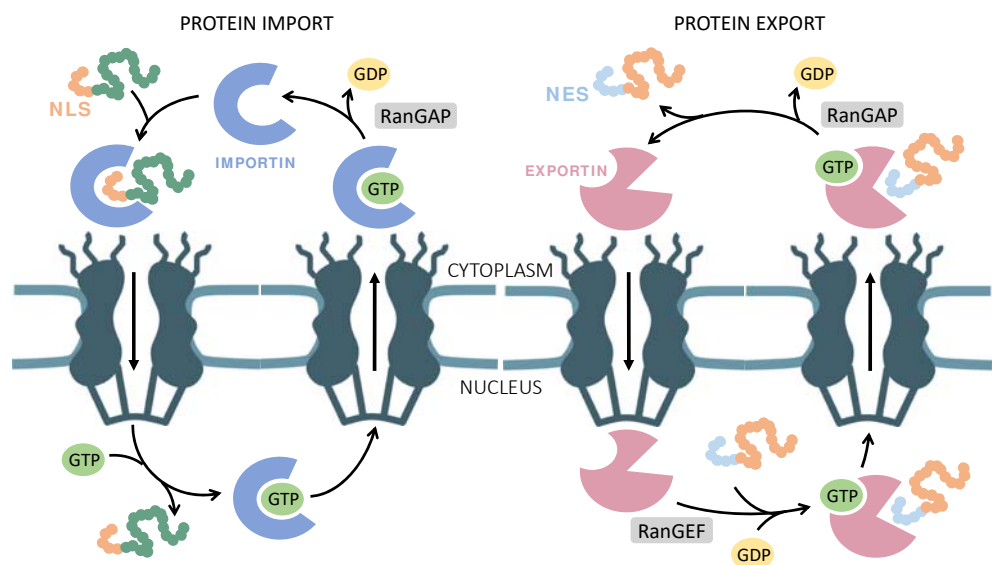


Figure 4. Schematic representation of protein transit through the NPC. Adapted from (Stewart, 2010).

Importins can associate with their cargo only in the absence of RanGTP, thus in the cytoplasm, while the formation of a trimeric complex with RanGTP cause loss of affinity for the substrate in the nucleus (Fig. 4, *left panel*). Viceversa, exportins retain affinity for the cargo only in association with RanGTP, while GTP hydrolysis in GDP by the Ran GTPase-activating proteins (RanGAP) causes its release in the cytoplasm (Fig. 4, *right panel*). The nuclear transport factor NTF2 take care of relocating RanGDP to the nucleoplasm where Ran Guanine nucleotide Exchanging factors (RanGEF) contribute to maintain the gradient by converting Ran-GDP to Ran-GTP (reviewed in Hampoelz et al., 2019).

Karyopherins recognize specific Nuclear Localisation Signal (NLS) or Nuclear Export Signal (NES) present in the cargo protein. Remarkably, unlike other types of import signals, such as those targeting to the mitochondria or the ER, NLSs can be found anywhere in the protein sequence, and are not removed once transport is accomplished, leaving the possibility of multiple passages. Although several consensus sequences have been identified as functional NLS, many other non-canonical NLS have been identified (reviewed in Lu et al., 2021); such degree of variability makes the prediction of NLS from the protein sequence inaccurate, and identification through gradual deletion approaches are sometimes required in order to identify its position.

mRNA export

RNA transit is mostly unidirectional toward the cytoplasm, apart from the nuclear import of hnRNP forming the spliceosome machinery, the RNA component of the telomerase, and the retrograde import of tRNAs in certain conditions (Ashkenazy-Titelman et al., 2020). Depending on the class of RNAs, different transport receptors are involved for the transcript export to the cytoplasm. The export of ribosomal RNA can be mediated by Cmr1 or exportin 5 (Xpo5), while tRNA export require exportin-t (Xpot) in vertebrates and its ortholog Los1 in *S. cerevisiae*. Both of these export pathways rely on exportins and therefore are regulated by the small GTPase Ran similarly to the protein export pathway described above (reviewed in Okamura et al., 2015).

Competency for mRNA export, instead, is acquired through the binding to the nascent transcript of several factors that associate to the RNA molecule co-transcriptionally, forming a mRNP (messenger ribonucleoparticle, Fig. 5). Among the RNA-binding proteins that compose the mRNP, the THO/TREX complex (Hpr1, Mft1, Tho2, Thp2, Tex1, Sub2, Yra1 in budding yeast), which travels with the transcription machinery, associates with the transcript during transcription elongation and recruits the export receptor Mex67-Mtr2 (Tap-p15/NXF1-NXT1 in metazoans). Other RNA-binding proteins (e.g. Nab2, Npl3 in yeast) have been proposed to similarly function as adaptors to recruit Mex67-Mtr2 onto mRNA (reviewed in Rougemaille et al., 2008). Mex67-Mtr2 has also been shown to interact with NPC associated proteins, i.e. Sus1, subunit of the SAGA complex, the TREX-2 complex, and nuclear basket nucleoporins (Fischer et al., 2002; Rodríguez-Navarro et al., 2004 and references therein), bridging transcription elongation with nuclear pore docking in preparation to the export. The interaction of the export receptor with the FG nucleoporins then allows the passage of the mRNP through the pore.

Interestingly, Mex67-Mtr2 appears to interact with a separate set of FG domains than the one typically contacted by karyopherins and do not rely on the RanGTP gradient (Terry and Wentz, 2007); the directionality of the transport is given by the requirement for cytoplasmic-restricted factors to complete the export process. Once the mRNP reaches the cytoplasm, the DEAD-box helicase Dbp5 (hsDDX19), docked at the NPC by Nup159 (hsNup214), is responsible for the release of the RNA by removing the associated proteins, including the export receptor and adaptors, which are imported in the nucleus to be recycled (Tran et al., 2007; reviewed in Carmody and Wentz, 2009).

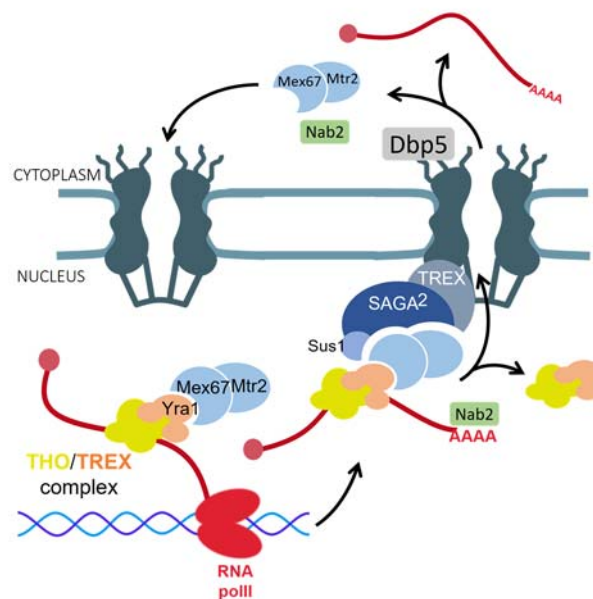


Figure 5. Schematic representation of export of mRNAs through the NPC.

Despite the high conservation of the aforementioned factors among several organisms, species specific discrepancies have been observed regarding their absolute requirement for the export process: Yra1 and Mex67 for example are essential in *S. cerevisiae*, while Mex67 homolog in *S. Pombe*, and Yra1 homologs in *D. melanogaster* or *C. elegans* are not (Carmody and Wentz, 2009).

mRNA quality control

It is important to note that the full export competency of mRNA is reached not only by the recruitment of the export receptors, but also requires the successful completion of the three major steps in mRNA maturation: 5' capping, splicing, 3' end cleavage and polyadenylation. Several mRNA surveillance mechanisms are in place to ensure that only properly processed and assembled mRNPs are allowed to exit the nucleus, some of which involve nucleoporins. Unspliced mRNAs, for example, have been shown to be retained in the nucleus by a process requiring the nuclear basket Mlp1-2/TPR and their associated protein Pml39 (Galy et al., 2004; Palancade et al., 2005; Coyle et al., 2011). Interestingly, the necessity of quality control mechanism can be overcome in stress conditions: while bulk mRNA experience Mex67 dissociation upon temperature shift at 37°C, heatshock-induced mRNA maintain the ability to bind the transport receptor, and are allowed to be exported bypassing the quality control of the nuclear basket to ensure survival in extreme situation (Zander et al., 2016).

1.1.4 | Transport-unrelated roles of nucleoporins

The nuclear pore complex is remarkable not only for the conservation and sophistication of its architecture, but also for the wide variety of biological processes in which it is implicated. Beyond its roles in nucleo-cytoplasmic transport, several nucleoporins have been implicated in the regulation of transcription and DNA damage repair, two major nuclear pore functions which will be discussed in detail in the following dedicated paragraphs.

Moreover, nucleoporins have been shown to be involved also in the translation process. The cytoplasmic nucleoporin Nup358/RanBP2, for example, is important for efficient translation of secreted, ER-bound proteins (Mahadevan et al., 2013), but is also involved in miRNA-mediated translation repression through its SUMO-ligase activity (Shen et al., 2021). Moreover, the yeast cytoplasmic nucleoporin Gle1 interacts with translation factors and is involved in both translation initiation and termination, the latter via activation of the NPC-associated helicase Dbp5 (Bolger et al., 2008).

Several nucleoporins have been associated with developmental and differentiation processes. For example, nucleoporins of the Y-complex have been shown to be required for embryonic development and mouse embryonic stem cell differentiation (Lupu et al., 2008; Gonzalez-Estevez et al., 2021). Similarly, mutants of the plant homolog of TPR, Tpr/NUA, affect flowering time, seed production and leaf morphology (Xu et al., 2007). Furthermore, several nucleoporins mutants lead to alteration in plant stress responses and defence signalling (reviewed in Yang et al., 2017; Wu et al., 2022). Notably, the mechanisms underlying the involvement of nuclear pores in these processes is not yet fully elucidated, but it most likely relies on the NPC-dependent control of either nucleocytoplasmic transport or gene expression.

Moonlighting nucleoporins

Interestingly, nucleoporins also display secondary separate functions and localisation from their NPC-bound life, significantly broadening the array of processes in which the nuclear pore and its subunits are implicated.

In flies a distinct sub-pool of Nup98, Nup62 and Nup50 has been shown to reside in the nucleoplasmic space, interacting with active chromatin (Capelson et al., 2010; Kalverda et al., 2010, further discussed in paragraph 1.2.6).

In the cytoplasm of mammalian cells, nucleoporins have been found associated with the cytoskeleton and the spindle pole body during mitosis. For example, Nup358/RanBP2 and the Y-complex are involved in the stabilization of microtubule-kinetochore interactions (Salina et al., 2003; Joseph et al., 2004; Zuccolo et al., 2007; Berto et al., 2018). Moreover, nucleoporins have also been shown to interact with the microtubule network in S-phase (Joseph and Dasso, 2008).

Finally, inner ring nucleoporins have been shown to localize at the basal bodies of cilia, where they assemble with a distinct organization compared to their canonical conformation in the NPC (Kee et al., 2012; Del Viso et al., 2016).

1.2 | The nuclear pore and transcription regulation

From the earliest electron microscopy observations of the nucleus of differentiated eukaryotic cells, it became evident that chromosomes are organised into distinct territories, which assume a non-random distribution. In particular, the nuclear periphery is characterised by inactive heterochromatic zones interspersed with less dense euchromatic domains, indicative of transcriptional activity, which are found in correspondence of the nuclear pores (Akhtar and Gasser, 2007). As Gunter Blobel points out in his “gene gating” manifesto, although the information encoding this three-dimensional configuration is contained in the DNA sequence, the DNA molecule *per se* would not be able to assume this conformation without the assistance of protein factors designed to coordinate its three-dimensional organisation (Blobel, 1985). He therefore proposed the nuclear pore complex and lamina proteins as key factors in the coordination of chromosome conformation and general nuclear organisation. This hypothesis was supported by the observation of the non-random distribution of nuclear pores along the nuclear envelope (Maul, 1971), suggesting that their position may be related to the nature of the underlying chromatin. Numerous subsequent observations have confirmed these hypotheses; even in organisms without a lamina, such as the yeast *S. cerevisiae*, the nuclear periphery appears to be divided into two types of non-overlapping functional compartments, with areas of silenced chromatin, particularly at the telomere anchorage points, where there is the highest concentration of SIR-silencing regulatory proteins (Andrulis et al., 1998), while the pore microenvironment appears more open and transcriptionally active. In his pioneering vision, Blobel proposes the nuclear pore not only as a key factor in chromatin organisation and territory maintenance, but also as a direct regulator of transcription. In fact, he hypothesises the spatial restriction of transcription to the nuclear pore for a subset of mRNAs, favouring their expression in a certain area or moment in the life of the cell. Blobel argues for the importance of such spatial organisation in the establishment of cellular asymmetry and polarity, and even goes so far as to theorise a function of the pore in the maintenance of epigenetic memory (Blobel, 1985).

Over the years, thanks also to technological advancement, an ever-growing body of evidence has accumulated and helped paint an increasingly accurate picture of the nature and extent of interactions between the nuclear pore and chromatin. In the next paragraph, I will discuss the evidence supporting the association of the nuclear pore with active genes, the possible consequences for transcription regulation, and the relevance of this role for cell homeostasis.

1.2.1 | Strategies for the study of gene relocation to the periphery

Studies on the role of the nuclear pore in regulating transcription in *S. cerevisiae* have focused on genes whose transcription is induced by stress or nutrient shifts, which allow the timing and extent of induction to be finely controlled, and whose mechanisms of activation and repression are well known. The locus most used for such studies is the *GAL* locus, a region of chromosome II consisting of three neighbouring genes, *GAL1*, *GAL7*, *GAL10*, whose transcription is induced by the presence of galactose, and repressed in the presence of glucose. *GAL* promoters can also exist in an intermediate state, where they are neither bound by activators nor repressors, when cells are grown in media containing alternative carbon sources, such as raffinose or glycerol.

Other loci that have been extensively characterised for this purpose are *INO1*, an unfolded protein response gene, induced by inositol starvation, the heat-shock gene *HSP104*, and other galactose-induced genes such as *GAL2*, in chromosome XII, and the subtelomeric locus *HXX1*. It is also important to note that although most of the studies were performed on endogenous, tagged loci, studies performed on high-copy number reporter systems recapitulated the dynamics observed in a genomic context (Abruzzi et al. 2006; Vodala et al. 2008).

The characterisation of the association of transcriptionally active genes with the nuclear pore has been made possible by the development and optimisation of techniques with increasing resolution and reliability. The combination of different approaches has made it possible to delineate the dynamics and factors involved in the relocation of genes to the pore under a plethora of different conditions. The techniques used to assess the association of loci to the nuclear pore fall into two main categories: live imaging and biochemical approaches.

***In vivo* single-locus tracking by microscopy**

Early microscopy observations of loci localisation in the nucleus were performed using DNA Fluorescence In Situ Hybridization on fixed cells (Casolari et al., 2004). The development of new labelling strategies allowed the observation of nuclear organisation in live cells at high resolution, both for population studies and single particle tracking over time. Such technique consists in the labelling of the gene of interest with an array of tandem operator repeats, which when bound by fluorophore-tagged repressors appear as a bright nuclear focus, whose position can be determined relatively to the stained nuclear envelope (Fig. 6A; Robinett et al., 1996). The main advantage of this methodology is the possibility to obtain *in vivo* direct evidence about the localization of a given locus compared to the periphery, the dynamicity of the interaction and its behaviour over time. However, the insertion of the operator array may have an impact on the position of the locus and induces its silencing or replication stress (Dubarry et al., 2011). Moreover, such observations do not allow to conclude on the actual presence and nature of the interaction with nuclear envelope proteins. The inability, at present, to resolve nuclear pores from other proteins or regions of the nuclear envelope in fluorescence microscopy can be bypassed, in yeast, using nuclear pore clustering mutants (e.g. *nup133-ΔN*), which cause the aggregation of the nuclear pores on one

side of the nuclear envelope (Doye et al., 1994). Only in this context, colocalization of the locus with the NPC staining allows to conclude on loci-pore association.

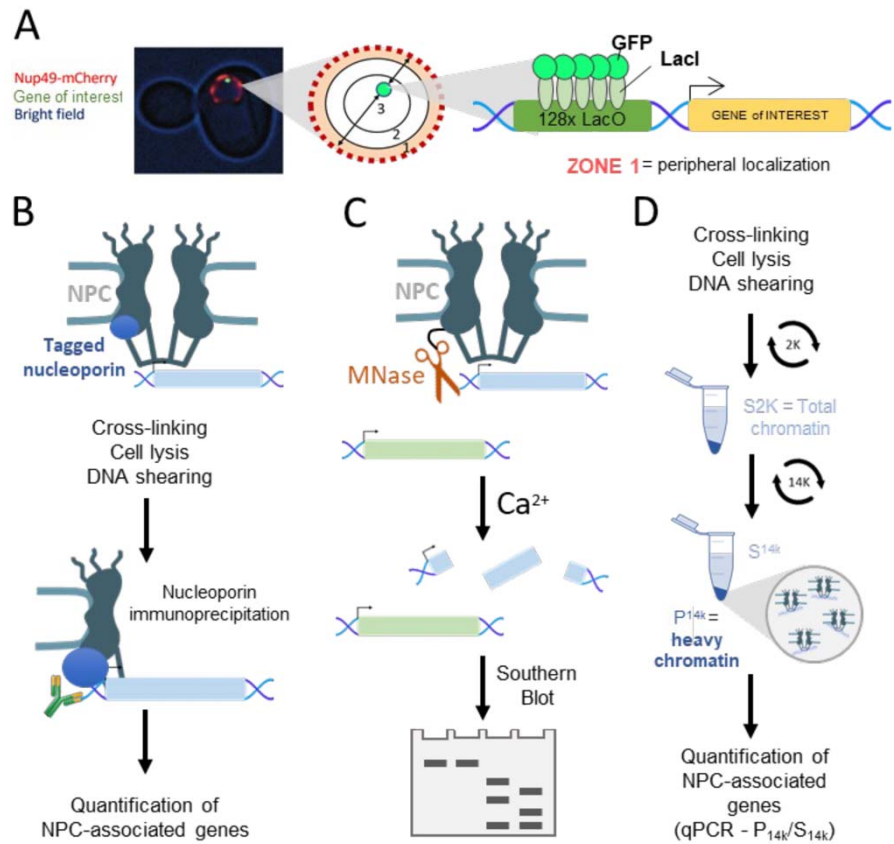


Figure 6. **Overview of the methodologies used to determine loci localisation relative to the nuclear pore in *S. cerevisiae*.** A. Principle of single locus tracking: the gene of interest is tagged with multiple tandem repeats of the bacterial Lac operator, which can be bound by the LacI repressor, fused with GFP. The locus appears as a bright green dot, whose distance can be measured relatively to the nuclear envelope, stained thanks to a fluorophore tagged nucleoporin. The nucleus is then divided in 3 equivolumetric concentric zones, and only the dots localised in zone 1 are considered peripheral. B. Principle of Chromatin Immunoprecipitation (ChIP): tagged nucleoporins are used as bait to immunoprecipitate the nuclear pore and nuclear pore associated genes. Upon elution, the DNA is retrieved and can be quantified by qPCR (ChIP-qPCR), microarray (ChIP-on-chip) or next generation sequencing (ChIP-seq). C. Principle of Chromatin Endogenous Cleavage (ChEC): the micrococcal nuclease is fused to a component of the nuclear pore. Upon its induction by Calcium addition, MNase cleaves the accessible chromatin localised at the vicinity of the nuclear pore. The cleavage profile is then analysed by Southern blot with locus-specific probes. D. Principle of Chromatin fractionation: fixed chromatin is lysed, sheared by sonication and then subjected to several rounds of centrifugation at increasing speed in order to isolate the “heavy chromatin” fraction, which comprises the nuclear pores and the nuclear pore associated genes. After decrosslinking, the presence of the genes of interest in such fractionation is assessed by qPCR and the enrichment is expressed as ratio between the insoluble and soluble part.

Biochemical approaches

Information about the physical interaction of loci with the nuclear pore can be achieved with the use of biochemical assays, which however, contrary to live imaging approaches, give little to no information about the dynamicity and stability of the interaction over time.

ChIP The first genome wide observations of gene-NPC association were carried out in the Silver lab by performing Chromatin ImmunoPrecipitation combined with tiling microarray (ChIP-on-chip; Fig. 6B) using several different nucleoporins or other nuclear pore associated factors as bait (Casolari et al., 2005, 2004). Moreover, ChIP has also been used to assess the binding and distribution of factors to the loci of interest (Dieppois et al., 2006). However, such pioneering experiments needs to be interpreted carefully due to lack of controls to ensure that the tagging of the protein used as bait is neutral to its localization, and to account for the “hyper-ChIPability” of highly transcribed genes (Teytelman et al., 2013).

ChEC The Chromatin Endogenous Cleavage (ChEC; Fig. 6C) method exploits the inducible cleavage activity of the micrococcal nuclease (MNase) fused to a protein of interest to obtain the footprinting of the binding of such protein to the DNA molecule *in vivo* and at high base pair resolution (Schmid et al., 2006). Such technique allows the mapping of interactions in native conditions, without need for fixation. However, caution must be kept in the setup of the experiment to account for the bias of the MNase for AT rich regions (Dingwall et al., 1981) and for hypersensitive sites in the tested conditions (Dieppois and Stutz 2010).

Chromatin Fractionation

Another method to assess loci association to the pore is chromatin fractionation (Fig. 6D), which, through several rounds of centrifugation of fixed chromatin, allows the isolation of the “Heavy chromatin” fraction, which comprises components of the NPC and its associated genes (Rougemaille et al., 2008a; Mouaikel et al., 2013). It must be noted that such assay is less specific compared to the other described above since it assesses the co-fractionation of 2 elements in a determined fraction, without using baits for identification. Despite the confirmed presence of nucleoporins in the heavy chromatin fraction, they do not constitute the sole content of such fraction, therefore the possible presence of different “heavy” complexes should be carefully considered.

All these biochemical strategies can be applied to locus-specific approaches, by assessing the enrichment of a model gene by qPCR, or to genome-wide analyses, by microarray or next generation sequencing. In light of the limitations highlighted in these methods, it appears evident how investigations of loci-NPC association should use a complement of both biochemical and live imaging approaches, as neither of them by itself can give a complete picture of the studied phenomenon, while combined they constitute a powerful tool to investigate such dynamic and complex interactions.

1.2.2 | Dynamics of loci diffusion in the nucleus

Before diving into the characterisation of the mechanisms underlying the transcription-dependent relocation of genes to the periphery, it is important to have an idea of the mobility and position of these genes before they engage in interactions with the pore. The development of microscopy techniques capable of tracking single loci at high resolution *in vivo* has made it possible to refine our knowledge of the dynamics and characteristics of the peripheral relocation of inducible genes upon activation.

Early observations of chromatin motility in *Drosophila* and budding yeast nuclei describe a Brownian motion, which varies depending on the phase of the cell cycle and depends on ATP availability but not on the microtubule network, suggesting the active involvement of ATP-dependent protein factors in coordinating this mobility, such as chromatin remodellers or the transcription and replication machineries (Heun et al. 2001; Vazquez et al. 2001).

Chromosomes behave as semi-polymers, whose mobility is highest in the nucleoplasmic zone and limited at the extremity of the molecule or at the vicinity of tethered regions, such as the centromere and the telomeres in the yeast nucleus (Albert et al., 2012). Single-particle tracking experiments show that induction of transcription causes restriction of the movement of the locus of interest to a limited area of the nucleus, close to the nuclear periphery. Interestingly, such confinement, while limiting the volume explored by the locus, does not reduce its motility: the locus continues to move, sliding laterally along the nuclear envelope (Cabal et al., 2006; Taddei et al., 2006; Sumner et al., 2021). The persistence of such two-dimensional movement could be at least partially accounted by the diffusion within the NE of the NPCs to which the locus would be tethered. However, the observation that the residence of *INO1* at the periphery upon transcription induction, although more persistent, remains discontinuous over time (Sumner et al., 2021), rather suggests that the interaction with the nuclear pore is transient and dynamic.

Another important question for understanding this phenomenon is whether the induced loci are driven towards the pore by a directed cytoskeleton-dependent movement. Although some studies have shown an involvement of actin-related proteins in gene localisation (Yoshida et al., 2010; Maruyama et al., 2012; Wang et al., 2020), this effect seems to be accounted by the requirement of Arp6 for histone H2A.Z deposition, rather than an active involvement in transport. Consistently, the measurement of the repositioning timing of an induced gene compared to the motility of a free diffusing locus seems to exclude the presence of active movement (Sumner et al., 2021).

It is important to point out that, especially in a dense genome like the one of *S. cerevisiae*, localization of a specific locus is tightly linked to the position of neighbouring genes, adding another layer of complexity to the interpretation of the observations. Both ChIP and ChEC experiments showed how transcription induced gene gating interested way more than just the known loci targeted by the transcription induction, indicating the occurrence of large-scale rearrangements of

the genome organisation following local changes in the transcriptional program (Casolari et al., 2005; Schmid et al., 2006). Gene relocation to the nuclear pore upon transcription activation has also been shown to be accompanied by transcription-factor mediated inter-chromosomal clustering (Brickner, 2017). Furthermore, artificial tethering of a locus to the pore has been shown to provoke rearrangement of loci position up to 100kb upstream and downstream the tether region (Green et al., 2012). Finally, localisation and dynamicity of the reorganisation of a given chromosome region can be influenced by the eventual constraint of neighbouring regions, especially in nuclei adopting a Rab1 configuration, i.e. the clustering of centromeres and telomeres at opposite sides of the nuclear envelope. Indeed, the low rate of actively transcribed *HXX1* colocalization with the pore in *nup133-ΔN* mutants compared to its extent of peripheral localisation in wt (17% and 80% respectively) may be explained by the competition between its pore association and the binding to the nuclear envelope of the neighbouring subtelomeric loci (Taddei et al., 2006).

All together, these observations imply that the association of active loci with the nuclear pore should not be considered as a stable tethering mechanism, but rather a continuous dynamic formation and dissociation of interactions between chromatin and the nuclear pore, highly influenced by the surrounding chromatin context.

1.2.3 | The nuclear pore complex as a boundary between silenced and active chromatin

The pioneering gene gating hypothesis remained unexplored for years, owing to the abundance of observations defining the nuclear periphery as a dedicated chromatin silencing environment (Cockell and Gasser, 1999). This view was supported by observations showing that tethering of genes to the nuclear periphery can lead to gene silencing (Andrulis et al., 1998), and for the role in the establishment of perinuclear silent chromatin domains assigned to the nuclear basket nucleoporins Mlp1 and Mlp2 (Galy et al., 2000). This was thought until a genetic screen aiming at identifying factors acting as insulators established a role for transportins in exerting boundary activities. Strikingly, the ability of transportins in preventing the spreading of heterochromatin towards more active loci requires direct or indirect tethering to the nucleoporin Nup2 (Ishii et al., 2002). This suggested the existence of two distinct peripheral compartments: a silenced, rich in Sir proteins, heterochromatic domain, in which telomeres reside, and more active domains in correspondence to the NPCs, molecularly defined by insulators.

This first evidence of active transcription being possible at the nuclear periphery prompted the start of a vast series of studies investigating the correlation between transcription activation and the gene position with respect to the nuclear pore, the dynamics of such interactions and the mechanisms involved.

1.2.4 | Mechanisms mediating the interaction of active genes to the nuclear pore

The nature of the mechanisms and the protein factors involved in the interaction between active loci and the NPC is still object of debate, given the plethora of studies on the subject, which point to different and sometimes contradictory conclusions.

As evident from the overview depicted in Table 2, the mechanisms identified show a high degree of inconsistencies between them. However, a few cardinal statements appear to find a consensus:

1. **Active transcription is not sufficient (and, in some instances, not necessary) for loci relocation to the pore.** Transcription inhibition using the thermosensitive RNA polII subunit mutant *rpb1-1* did not affect the position of the locus of interest, indicating that ongoing transcription is dispensable for relocation (Schmid et al., 2006; Brickner et al., 2007, 2016). Moreover, both *INO1* and *GAL1* have been shown to remain to the nuclear periphery way after transcriptional shut off (Brickner et al., 2007).
2. Viceversa, **loci interaction with the pore is not necessary nor sufficient for the transcription activation of the relocated loci.** Active transcription is detectable even in absence of relocation (Cabal et al. 2006; Dieppois et al. 2006; Taddei et al. 2006).
3. **The interaction between the DNA locus and one or more nucleoporins is bridged by a nucleic acid binding protein or complex.** Only a couple of studies found a requirement for the RNA molecule (Casolari et al., 2005; Abruzzi et al., 2006), indicating that the interaction is more likely dependent on DNA. However, none of the yeast nucleoporins have an identified DNA binding domain, while in metazoan nuclear pores only ELYS has been shown to directly bind DNA (Gillespie et al., 2007), but this domain do not seem involved for its gene gating function (Scholz et al., 2019). This imply that the NPC interaction must be mediated by DNA binding factors bridging the association between DNA and nucleoporins.

To some extent, the dissimilarities between the studies could be attributed to differences in the conditions and approaches and to the variety of loci analysed. However, it seems to be a shared belief in the field that the inconsistencies and sometime contradictory pieces of evidence collected in this subject very likely mirror the presence of multiple, partially overlapping mechanisms that coexist and concur to the fine tuning of gene localization and expression, particularly in situations of stress and challenge.

On top of the differences in sequence and protein determinants identified, a major criterion in which the studies appear to be divided is the moment at which the relocation to the pore is triggered: pre, during or post- transcription activation. In the following paragraph I propose a categorisation of the findings in three groups, based on similar dynamics of relocation, discussing differences, similarities, and limitations of such possible models.

Reference	Technique	Locus analysed	DNA bound factors	Factors Required	Factors not Required	Sequence Determinant
Casolari et al., 2004	ChIP-on-chip	genome wide	Rap1, Xpo1, Cse1, Kap95, Nic96, Nup116, Nup2, Nup60, Mlp1, Mlp2			Promoter
Casolari et al., 2005	ChIP-on-chip	alpha-factor induced	Mlp1	RNA		Bias for 3' in Mlp1 interaction
Brickner and Walter 2004, Brickner et al., 2007, Ahmed et al., 2010, Light et al., 2010, Randise-Hinchliff et al., 2016	Live microscopy	<i>INO1</i>		Hac1, Put3, Cbf1, Nup1, Nup2, Mlp2, Nup60, Nup157, Gle2, Nup42, SAGA, TREX-2	Transcription, Mlp1, Nup100, Nup116	GRSI, GRSII
Taddei et al., 2006	Live microscopy	<i>HXK1</i>				Mode of activation, 3' end
Abruzzi et al., 2006	Microscopy (fixed cells) + FISH	pGAL- <i>GFP</i> reporter		post-transcriptional mRNP		3' end
Schmid et al., 2006	Nup2 ChEC	<i>GAL</i> , <i>HXK1</i> , chr6		GAL4	SAGA, TBP, transcription	Promoter (UAS, not TATA box)
Cabal et al., 2006	Live microscopy	<i>GAL</i> locus		SAGA (Sus1, Ada2), TREX-2 (Sac3), NUP1	Nup2, Nup53, SAGA activity (Gcn5), mlp1, Nup60	Promoter
Green et al., 2012	Live microscopy	<i>GAL1</i>		SAGA (ada2), Nup1	Transcription	UAS
Brickner et al., 2007, 2016	Live microscopy	<i>GAL1</i>		Nup2, Nup1, Mlp2, Nup60	Transcription Mlp1	GRS4, GRS5
Dieppo et al., 2006	Live microscopy + Mex67 ChIP	<i>GAL</i> locus, <i>HSP104</i> (EtOH induced)	TBP, Mex67	Mlp1 and Mex67	Transcription	UAS and TATA box, (3' end)
Brickner et al., 2012	Microscopy	<i>HSP104</i>				GRSIII
Rougemaille et al., 2008; Mouaikel et al., 2013	Chromatin fractionation	<i>HSP104</i>	Nup60, Nup116, Gle1	Hsf1, Rrp6, Rna14/15		Promoter and terminator region

Table 2. Summary of the major findings investigating the sequence and protein determinants required for inducible loci relocation to the nuclear pore. GRS = Gene Recruitment Sequences; TBP= TATA Binding Protein; UAS = Upstream Activation Sequence.

Promoter activation driven model

A great number of studies suggested that gene – NPC interaction is an early event involving gene activation, pointing to transcription activators binding the promoter sequence as the determinants required for triggering loci relocation to the pore. The SAGA complex has been appointed by several studies as a key player in mediating the interaction to the NPC, given its role in transcription activation of a high proportion of loci, among which galactose inducible genes, and its interaction with the NPC through the TREX-2 export complex (Rodríguez-Navarro et al., 2004). Indeed, transcription dependent peripheral localisation of the *GAL* locus was lost in SAGA subunits mutants *ada2* and *sus1* (Cabal et al., 2006), but not in mutants for the Gcn5 subunit, responsible for the histone acetyltransferase activity, indicating that the role of the complex in mediating the interaction is purely structural, rather than functional (Cabal et al., 2006; Luthra et al., 2007). However, the requirement for the SAGA complex has not been observed consistently: Schmid et al. (2006) found the SAGA complex and the TATA Binding protein to be dispensable for galactose dependent interaction with Nup2, despite defining the promoter region as determinant for relocation. Randise-Hinchliff et al. (2016) found instead a requirement for SAGA only for a subset of the analysed genes, further pointing to the possible existence of multiple distinct mechanisms of active genes recruitment to the periphery. In addition, TREX-2 influence on gene-NPC association has been shown to be exerted also thanks to its interaction with the Mediator complex; impairment of NPC targeting of the inducible genes *GAL1* and *HXK1* have been observed in mutants of mediator subunits. Once more, different subunits were shown to be specifically required for the gating of the tested genes, due to the ability of the mediator complex to “interpret” a gene specific promoter complex by interacting with the associated transcription factors (Schneider et al., 2015). Mediator has also been shown together with the MRX complex to be enriched at highly transcribed genes contacting the NPC through Nup60. Disruption of MRX complex reduced gene gating for galactose-induced and mating pheromone-induced loci, in a mechanism independent on its role in damage checkpoint activation (Forey et al., 2021).

Furthermore, the ability of mediating the interaction between promoter sequences and the nuclear pore has recently been recognised to a large group of yeast transcription factors, which have been shown to be sufficient to reposition an ectopic site to the nuclear periphery, with distinct but possibly partially overlapping mechanisms (Brickner et al., 2019). However, transcription activation does not by default result in the relocation of a gene to the periphery: not all transcription factors have this property (Brickner et al., 2019), and alternative activation of the *HXK1* gene using the transcription activator VP16, instead of galactose, causes the complete loss of peripheral positioning rendering the locus prevalently nucleoplasmic and even causing the mislocalization of the telomere of the corresponding chromosome arm (Taddei et al., 2006).

Concerning the sequence determinant required for relocation, ChEC experiments mapped the Nup2 galactose-dependent interaction of the *GAL* promoter to a 50bp region positioned 40bp upstream of the TSS (Schmid et al., 2006). Furthermore, a

requirement for the UAS only (Schmid et al., 2006) or the UAS and the TATA box containing region of the promoter (Dieppois et al. 2006) was found to be necessary and sufficient to trigger galactose-dependent anchoring to the NPC. Moreover, the persistence of the anchoring even in absence of the coding region (Dieppois et al. 2006; Vodala et al. 2008) confirms that the promoter region is necessary and sufficient for efficient repositioning to the nuclear pore.

Taken together, these data compose a model suggesting that the acquisition of DNA-associated factors at the time of promoter activation renders the gene more competent to interact with the nuclear pore in the event of the locus shift to the nuclear periphery, increasing its time of residency in that area. Moreover, the involvement of the SAGA complex, direct interactor of the TREX-2 complex, suggest a link between activation induced relocation with transcription elongation and mRNA export.

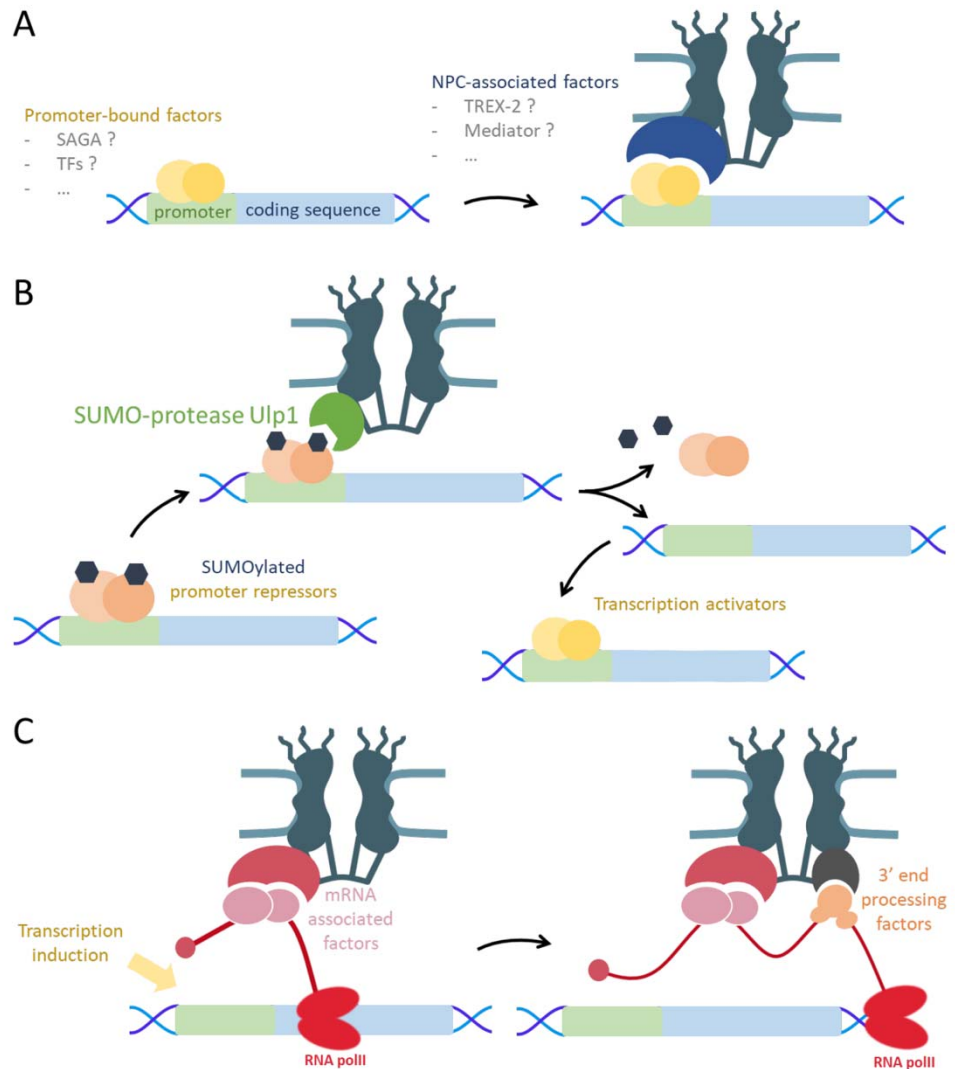


Figure 7. Schematic representation of the proposed models for active loci relocation to the nuclear periphery. A. promoter-activation driven model; B. Promoter de-repression driven model; C. mRNA dependent model.

**Promoter
de-repression
driven model**

In contrast with the previous model, a few studies highlighted how peripheral localisation, while still involving the promoter region, may depend on events occurring prior to or distinct from transcription activation.

The Gene Recruitment Sequences (GRS), or zip-codes, identified as necessary and sufficient to target *INO1* and *GAL1* the nuclear periphery (Ahmed et al., 2010; Brickner et al., 2016), despite being located at the promoter region are distinct from the promoter sequences required to activate transcription. *INO1* relocation to the nuclear pore has been shown to depend on the binding to the GRSs of Tup3 and Cpf1, which is possible only upon eviction of UAS-associated repressors that usually mask the targeting sequences (Randise-Hinchliff et al., 2016). Consistently, the ectopic insertion of *INO1* or *GAL1* zip-codes in a differently regulated promoter region leads to constitutive peripheral localization even in uninduced conditions (Ahmed et al., 2010; Brickner et al., 2016), while the artificial tethering of transcriptional repressors to the endogenous GRSs blocked targeting to the nuclear periphery (Randise-Hinchliff et al., 2016). Furthermore, the relocation of *INO1* to the nuclear periphery has been shown to require Siz2-dependent SUMOylation of factors associated to the GRSI region (Saik et al., 2020; Ptak et al., 2021).

In addition, the *GAL* locus has been observed to be predominantly peripheral not only upon galactose induction but also in cells grown in raffinose or glycerol, when the promoter is in a derepressed state, but transcriptionally inactive (Green et al., 2012). Interestingly, the eviction of repressors from glucose-regulated promoters requires their post-translational modification modulated at the nuclear pore proximity. This is the case for Ssn6 and Tup1, which are deSUMOylated by Ulp1, a SUMO-specific protease whose localisation is spatially restricted to the nuclear pore microenvironment through interactions with nuclear basket nucleoporins, notably Mlp1 and Mlp2 (Texari et al., 2013). Thus, in cells where Ulp1 is correctly restricted to the pore microenvironment, migration of the to-be-activated *GAL* locus to the pore allows accessibility of the gene to such enzyme favouring de-repression. Consistently, mislocalization of Ulp1 might promote transcription derepression also in presence of glucose, given the increased chance of encounters between the enzyme and the locus (Texari et al., 2013), which may explain how aberrant *GAL1* expression can be rescued by its tethering to the periphery in a nuclear basket mutant (Green et al., 2012).

Overall, this points to a role of the nuclear pore in the kinetic of derepression rather than in actively enhancing activation. Moreover, the nucleoporins function in mediating gene repositioning and activation may be an indirect effect of their role in spatially restricting the Ulp1 sumo protease, which in turns regulates TFs binding or activity at the promoter region.

A role for mRNA associated proteins in gene-NPC association

Despite a great body of evidence supporting a major role for promoters in triggering loci relocation to the periphery, several studies also highlighted the importance of the nascent transcript and its associated proteins in mediating the interaction to the NPC. The interaction with the NPC of *HXX1* or alpha-factor inducible loci appear to be dependent on RNA and show a bias for the 3' of the observed genes (Casolari et al., 2005; Taddei et al., 2006). Moreover, experiences carried out in high copy number plasmid reporter systems show how relocation of *GFP* under the control of the *GAL* promoter has a dependency on the nature of the 3'UTR sequence, notably the poly(A) site (Abruzzi et al., 2006). Furthermore, *GAL1* NPC association has been shown to be mediated by the formation of a persistent RNA dot, adjacent but spatially distinct from the transcription site, indicating the presence of a physical link between the mRNPs and the locus (Abruzzi et al., 2006). All together, these observations lead to speculate a role of post-transcriptional mRNPs in mediating the interaction with the NPC. Consistently, a requirement of the mRNA export factor Mex67 for anchoring at the pore has been observed for the *GAL* locus, with an increase occupancy at the centre of the gene body, as scored by ChIP (Dieppo et al. 2006). Surprisingly, Mex67 role in bridging the interaction with the NPC appear to be RNA independent, suggesting its prior recruitment to the locus of interest at early stages of transcription activation, when Mex67 is travelling with the transcription machinery (Gwizdek et al., 2006). Mex67 could be then only subsequently transferred to the RNA, stabilizing the Gene-NPC interaction, and favouring mRNA export through the pore.

Another example of 3' end dependent association to the NPC has been identified for *HSP104* and several other heat-shock inducible genes, found associated to the pore upon heat shock induction in cells lacking the THO complex, which is important for mRNP formation and export competency (Rougemaille et al., 2008a). *HSP104* shows a clear bias for its NPC association at the 3' end of the gene, and increased residency of Mex67 and 3' processing factors, both required for the association, Rrp6, the catalytic subunit of the nuclear exosome, and the nucleoporin Nup60.

It is important to note that despite the interaction being post-transcriptional, possibly due to the accumulation of a stalled protein intermediate caused by inefficient 3' end processing, *HSP104* association to the pore also shows requirement for promoter elements and transcription activation. Indeed, despite the interaction always occurring at the 3' end of the gene, regardless of its nature or position, insertion of the promoter region to an ectopic locus is sufficient to trigger its relocation (Brickner et al., 2012; Mouaikel et al., 2013). Moreover, the interaction to the NPC seem to depend on the transcription activator Hsf1 (Mouaikel et al., 2013), which is more consistent with the transcription activation driven model.

Another case of potential double requirement for promoter and gene body determinants is exemplified by the *INO1* locus. While initial SUMOylation events at the promoter-proximal zip-codes are necessary for targeting the inactivated *INO1* gene to the pore, after 3h of inositol starvation, Siz2 occupancy propagates along

the gene body, and the deSUMOylase Ulp1 is found associated with the 3', favouring efficient transcription and NPC association of the activated gene (Saik et al., 2020).

It is therefore plausible that, at least in the cases described above, instead of reflecting a distinct targeting mechanism, the stable association of the 3' end with the NPC may be an event chronologically subsequent to promoter-mediated targeting to the pore, which allows the interaction to be stabilised, particularly in a context such as that of the *tho* mutant, characterised by defects in transcription elongation and mRNA export.

1.2.5 | Consequences of NPC association on gene expression

In line with the “gene gating” hypothesis, it has been widely suggested that the main function of relocation of active loci to the nuclear pore would be the improvement of gene expression.

In light of the multiple mechanisms of activation and relocation described above, nuclear pore vicinity appears to be potentially beneficial in several ways, acting in different moment of the transcription process:

- The presence of a gene in proximity to the nuclear pore can increase accessibility to activating factors facilitating transcription initiation.
- A gene maintained repressed in the nucleoplasmic space can gain access to de-repression mechanisms by migrating to the nuclear periphery.
- Activation of a gene favours association with the nuclear pore to positively promote its expression, possibly facilitating RNA export and 3' end processing.

A few reports in yeast showed direct evidence for an improvement in transcription efficiency after relocation to the pore (Brickner and Walter 2004; Taddei et al. 2006; Saik et al. 2020). Moreover, in human, the recruitment to the nuclear periphery of active *MYC* alleles resulted in improved mRNA export efficiency, specifically in cancer cells (Scholz et al., 2019). Also, a recent study in yeast proposed a role for NPC association in repressing pervasive transcription, promoting transcription directionality (Forey et al., 2021). However, improved transcription efficiency was not systematically detected as a consequence of pore relocation.

Even if the relocation does not translate in a measurable increase in transcription levels, beneficial effects in the dynamics and kinetics of transcription activation, RNA processing or NPC-associated RNA quality control mechanisms should not be excluded. Indeed, the reduced expression of *GAL1* and *GAL10* in absence of the *GRS4* zip-codes resulted from a decrease of the proportion of cells in which the *GAL* locus was active (Brickner et al. 2016), suggesting that the encounter with the nuclear pore increase the probability of proper transcription initiation.

We can therefore imagine a general model according to which the nuclear pore acts as a receiving scaffold equipped with distinct but close 'platforms' that could transiently accommodate certain genes at given times of their activation process. The competence for such locus–pore interaction is determined by the accessibility of the binding domains on the pore side, and the nature of the DNA-binding factors

decorating the locus. Conversely, NPC-binding may also actively modulate the association of factors to repositioned genes, for instance by promoting their eviction and degradation via the proteasome, which has been shown to associate with active genes regulating transcription through proteolytic processing of transcription factors (Auld and Silver, 2006) and to be anchored to the nuclear basket (Albert et al., 2017).

The nuclear pore ability to shelter such interactions can be modulated by post-translational modification: Nup2 phosphorylation by the cyclin dependent kinase Cdk1 causes the loss of peripheral localization of *INO1* and *GAL1* in S-phase and alpha-factor arrested cells (Brickner and Brickner 2010). Moreover, the loss of *HIS4* interaction with the NPC in SAGA or Mediator mutants could be bypassed by the overexpression of the transcription factor Gcn4 (Brickner et al. 2019), indicating that multiple, partially overlapping and redundant mode of interactions could interest the same locus; the access to or choice of a specific pathway depends on the factors associated to the locus itself.

This degree of selectivity would therefore restrict the ability to reside persistently at the pore to only a subset of genes undergoing specific steps of the transcription process. In this scenario, the nuclear pore assumes a pivotal role in controlling the accessibility to protein regulatory factors at the moment of need, and in spatially and temporally separating differentially activated genes to possibly avoid unintentional firing at neighbouring loci, a task even more important in a dense genome such as the one of *S. cerevisiae*.

1.2.6 | Nucleoporins contribution to gene regulation

As already highlighted above, an important contribution to loci residency to the periphery is given by the nucleoporins, which provide the docking site for the relocating loci. Among the nucleoporins identified to be involved in the process, on top of the more intuitive involvement of the nuclear basket nucleoporins (Nup60, Nup1, Nup2, Mlp1, Mlp2), also components of the more internal subcomplexes have been identified, such as the central channel nucleoporins Nup100 and Nup116. It is interesting to note how different nucleoporins has been identified by different studies, or even among the same study, to have different roles or specificity for the same relocating genes. Of note, the ability of several transcription factors to mediate loci relocation to the NPC appeared to have a specific requirement for Nup2, or Nup100, or in some cases both (Brickner et al. 2019). Moreover, *GAL1* and *INO1* interact with distinct sets of nucleoporins when relocating upon the first round of transcription or when engaging in inter-chromosomal clustering for the maintenance of transcriptional memory.

Such active involvement of nucleoporins in the regulation of transcription is even more evident when looking at their behaviour in metazoan nuclei. The role of nucleoporins in directly influencing gene expression of specific subsets of genes has been observed also in *Drosophila* nuclei, but with a twist. While the majority of nucleoporins appear to be rather immobile at NPCs, several nucleoporins, e.g. Nup2 in yeast, and a greater number in metazoan nuclei, show higher dynamicity and

plasticity and can be found detached from the main scaffold and have prolonged residency in the nucleoplasm (Dilworth et al., 2001; Rabut et al., 2004). *Drosophila* nucleoporin Nup98 (homolog of the yeast Nup100/Nup116), has been shown to interact with highly transcribed genes in both its NPC-tethered and nucleoplasmic configuration. However, only the genes associated to the nucleoplasmic pool of Nup98 show an improved transcription and a more open chromatin landscape (Maya Capelson et al., 2010; Kalverda et al., 2010). This points to a shift in the functions of nuclear organization during evolution, probably due to the increasing dimensions and complexity of the genome and the nuclear environment: despite the nuclear pore proximal space still corresponding to regions of open chromatin, it does not appear to be a main hub of transcription activation, and nucleoporin-dependent regulation activity has shifted to the nuclear interior.

Another evidence for the active function of nucleoporins in regulating transcription is the identification, in mammalian leukemic cells, of chromosome translocation leading to the fusion of DNA binding domains of transcription factors with FG-nucleoporins repeats domains, including Nup98 (Wiermer et al., 2007; Parry, 2013). Such chimeric proteins were shown to be able to modulate transcription of genes, possibly by changing the acetylation state of the target loci (Kasper et al., 1999; Bai et al., 2006), or more recently, by triggering phase separation within the nucleoplasm (Ahn et al., 2021; Terlecki-Zaniewicz et al., 2021).

In the nuclei of male flies, their unique X chromosome shows high enrichment of the nucleoporin Nup153 (homolog of yeast Nup1 and Nup60) and Mtor (Mlp1/2) (Mendjan et al., 2006). Interestingly, the two nucleoporins appear to coat the X chromosome at high density, spanning extended chromosomal regions up to 500kb. Such regions, called NAR (Nucleoporin Associated Regions) appear to have predominantly acetylated histones, mark of active transcription, which is necessary for the 2-fold transcription upregulation of the sex-related genes as required for the dosage compensation mechanism characteristic of flies sex determination (Vaquerizas et al., 2010). Moreover, Sec13 and ELYS has been shown to promote chromatin decondensation at target genes, thus favouring gene activation by generating regions of open chromatin in which the transcription machinery can bind (Kuhn et al., 2019). Nucleoporin action on gene expression seems therefore to transcend a mere structural function; modulation of gene expression seems likely to be conducted also by local modulation of the chromatin landscape promoting open chromatin conformations favourable to transcription.

1.2.7 | Nuclear pore association and transcriptional memory

Another mechanism with which the nuclear pore has been implicated in transcriptional regulation consists in the establishment of “transcriptional memory”, a phenomenon promoting faster reactivation upon repression of inducible genes.

This function have been proposed to be exerted through the formation of “gene loops”, in which the 5' and 3' of a gene are maintained in proximity with each other, favouring the loading and recycling of the polymerase to achieve faster RNA accumulation upon reinduction (O'Sullivan et al., 2004; Tan-Wong et al., 2009). In yeast, this phenomenon has been observed for several inducible genes, i.e. *HXK1* and *GAL1*, but not *INO1*, and appeared to be dependent on the nuclear basket nucleoporin Mlp1 (Tan-Wong et al., 2009).

In parallel, several studies from the Brickner lab characterised the prolonged residence of *INO1* and *GAL1* at the nuclear periphery after transcription repression, which can be inherited by subsequent generations (Brickner et al., 2007). *INO1* transcriptional memory does not seem to be mediated by “gene loop” formation but it correlates with the persistent interaction of a poised, unphosphorylated RNA polIII with the promoter, and requires the deposition of the histone variant H2A.Z and the interaction with Nup100 (Light et al., 2010). Moreover, inter-chromosomal clustering of *INO1* alleles, or between *GAL* genes, has also been observed upon transcription activation, and has been shown to require different zip-codes and to contact different nucleoporins than peripheral recruitment upon first activation (Brickner et al., 2016). Transcriptional memory allows faster reactivation upon repression of inducible gene responsive to metabolic changes, providing an advantage in fast-changing, unstable environments to promote adaptation and survival.

Transcriptional memory is a conserved phenomenon, that affects several hundreds of genes in human cells, notably genes responsive to interferon- γ (IFN- γ), is associated with the H3K4me2 chromatin mark at the promoter, and requires Nup98 for association to the NPC (Light and Brickner, 2013). In *Drosophila*, Nup98 has been shown to be required for enhancer-promoter looping at ecdysone-inducible genes, which also exhibit transcriptional memory (Pascual-Garcia et al., 2017). In this context, Nup98 also interact with other structural proteins, including the CCCTC-binding factor CTCF, which, together with cohesin, facilitates interactions between *cis*-regulatory elements by formation of chromatin loops, coordinating higher-order chromatin organisation. In mouse embryonic stem cells, Nup153 similarly interacts with transcription start sites and enhancer, and also associates with TAD boundaries, thus interacting with CTCF and cohesin (Kadota et al., 2020).

All together, these observations highlight an additional role for the nuclear pore in the maintenance of transcriptional memory, a pivotal event for the establishment of cell type specific transcription in response to cellular cues, e.g. growth factors, and the promotion of the rapid response of poised genes to environmental cues. Indeed, peripheral relocation in yeast seem to predominantly occur in response to stress of metabolic changes, when fast adaptation to the new conditions is pivotal for the cell survival. Flies and mammalian nucleoporins seem to preferentially interact with cell cycle and differentiation genes (Kalverda et al., 2010), and their deregulation could be oncogenic. Finally, plants nucleoporins have been shown to participate in signalling of the defence response (reviewed in Wiermer et al. 2007; Parry 2013), and mutation in *A. thaliana* nuclear basket nucleoporin Tpr/NUA have been reported to affect flowering time, seed production and leaf morphology (Xu et al., 2007). All together, these observations point to a conserved, fundamental role of nucleoporins in modulating the expression of several genes, whose regulation may be pivotal for cell survival and the determination of cell identity.

1.3 | The nuclear pore and maintenance of genome stability

In addition to the important roles in exchanges between nucleus and cytoplasm and in the regulation of gene expression, the nuclear pore also emerges as a key player in maintaining genomic stability and coordinating DNA repair processes.

The most intuitive way in which the nuclear pore can protect the genome from damage is by acting as a gatekeeper, isolating the nuclear environment from the cytoplasm. Ensuring the compartmentalisation of the genomic material protects it from the unregulated action of cytoplasmic nucleases, e.g. hSTREX1, whose mislocalisation in the nucleus, observed in cells residing at the edge of tumour masses, can become a source of DNA damage that can promote tumour progression (Nader et al., 2021). Moreover, preventing leakage of genomic DNA in the cytoplasm avoids the triggering of autoimmune inflammation responses normally directed at viral DNA, which can lead to chronic inflammation (Wolf et al., 2016).

In addition to structural safeguarding, nuclear pores seem to have a more active role in coordinating DNA damage repair pathways, as testified by the harmful consequences on genome integrity due to mutations of nuclear pore components. Several studies showed how the deletion of nucleoporins, in particular the nuclear basket and the Nup84 complex, leads to increased sensitivity to genotoxic stress (Galy et al., 2000; Loeillet et al., 2005). Moreover, inactivation of some non-essential subunits of the yeast Nup84 complex (e.g. Nup133, Nup120, Nup84) display synthetic lethality when combined with mutants of the replication-associated flap-endonuclease Rad27 or factors of the Rad52 pathway (Loeillet et al., 2005). An increase in Rad52 foci, hallmark of DNA damage, specifically in S and G2 phase was observed in *nup84* mutants, and to a lesser extent in *nup60* and *mlp1mlp2* mutants but not in mutants impairing mRNA export, protein import, or affecting nuclear pore distribution (Loeillet et al., 2005; Palancade et al., 2007), indicating that the increase in spontaneous DSBs formation is due to the loss of specific functions shared between the Nup84 complex and the nuclear basket (Palancade et al., 2007). In addition, also deletion of the NPC partner Slx5/8 leads to increase sensitivity to HU (Mullen et al., 2001), a significant increase in Gross Chromosomal Rearrangements rates, cell cycle delay and constitutive checkpoint activation (Zhang et al., 2006). Furthermore, numerous studies in mammalian cells identified several nucleoporin abnormalities associated to cancer and diseases, affecting, among others, perturbation of the DDR response and the S-phase, G2/M and mitotic-spindle checkpoints (reviewed in Bermejo et al., 2012).

Notably, also the control of import and localisation of DNA repair related proteins revealed to be important for genome stability. In particular, in human cells Nup153 contribute to regulate the choice between homologous recombination (HR) and non-homologous end joining (NHEJ), likely by regulating the nuclear import of 53BP1, which is predominantly mislocalised in the cytoplasm in *nup153* mutants (Lemaître et al., 2012; Moudry et al., 2012).

Finally, an ever-growing number of studies has highlighted the involvement of NPC and NPC partners in the spatial regulation of repair, in particular concerning the compartmentalisation of damage forming in critical loci, or that cannot be repaired by canonical repair pathways. The following paragraphs will describe the circumstances requiring localised repair at the nuclear pore, the mechanisms driving lesion relocation to the nuclear periphery, and the role of the nuclear pore in directing them to the most appropriate repair pathway.

1.3.1 | Strategies to investigate damage compartmentalisation

The nuclear pore has been shown to be involved in the regulation of the repair of several types of damaged or challenged loci: persistent DSBs (Nagai et al., 2008), challenged replication forks (Su et al., 2015; Kramarz et al., 2020), eroded telomeres (Khadaroo et al., 2009), and DSBs forming in critical loci, i.e. repeat-rich heterochromatin (Chiolo et al., 2013) and the nucleolus (Horigome et al., 2019). These lesions have been shown to relocate to the nuclear periphery, where they seem to be channelled into distinct repair pathways. Notably, such relocation events appear to be conserved from yeast to mammals, indicating the pivotal importance of such mechanisms for the maintenance of genome stability.

The numerous studies aiming at deciphering the mechanisms underlying the relocation of challenged loci to the nuclear periphery took advantage of various strategies to exacerbate the formation of stress or damage to the structures of interest, so that their formation could be controlled, and their evolution monitored.

DSBs generation

Several genotoxic agents can be used to generate a consistent amount of damage in cells, such as chemical compounds (Bleomycin, Zeocin, Methyl Methanesulfonate– MMS) or Ionizing Radiations (IR).

In yeast, the cleavage site of the homothallic (HO) endonuclease (either in its physiological position at the *MAT* locus or ectopically inserted at a position of interest) and the concomitant deletion of the donor loci *HML* and *HMR* can be used to generate persistent DSBs which cannot be repaired by Homologous Recombination (Fig. 8A). Inducible expression of the endonuclease leads to the establishment of repeated cycles of cleavage and ligation, ultimately leading to cell death by chromosome loss (Lee et al., 1998). The rare-cutting *I-sceI* endonuclease can also be used for the same purpose (Marcand et al., 2008).

Position of the lesion can be detected by labelling the cleavage site with an array of operator repeats (see Fig. 6A in paragraph 1.2.1), while detection of fluorophore-tagged repair factors can be used to monitor repair dynamics.

Replication stress induction

Replication forks can be stalled by treatment with hydroxyurea (HU), which inhibits the Ribonucleotide Reductase (RNR) enzyme thus causing dNTPs starvation and blocking DNA synthesis (Slater, 1973). While the reversibility of this treatment makes it a useful tool for cell synchronization, high concentration or prolonged treatment lead to DNA break formation and fork collapse (Musiałek and Rybaczek, 2021). Dysfunctional replication forks can also be artificially generated in a site-specific manner by introducing at a locus of interest the polar Replication Fork Barrier (RFB) naturally positioned close to the *S. pombe mat* locus, which consists

in the 859 bp RTS1 sequence, which is bound by the proteins Swi1 and Swi3, which constitute a physical obstacle for the replication machinery (Fig. 8B), leading to replication fork stalling that results in intrachromosomal recombination and gross chromosomal rearrangements (Lambert et al., 2005).

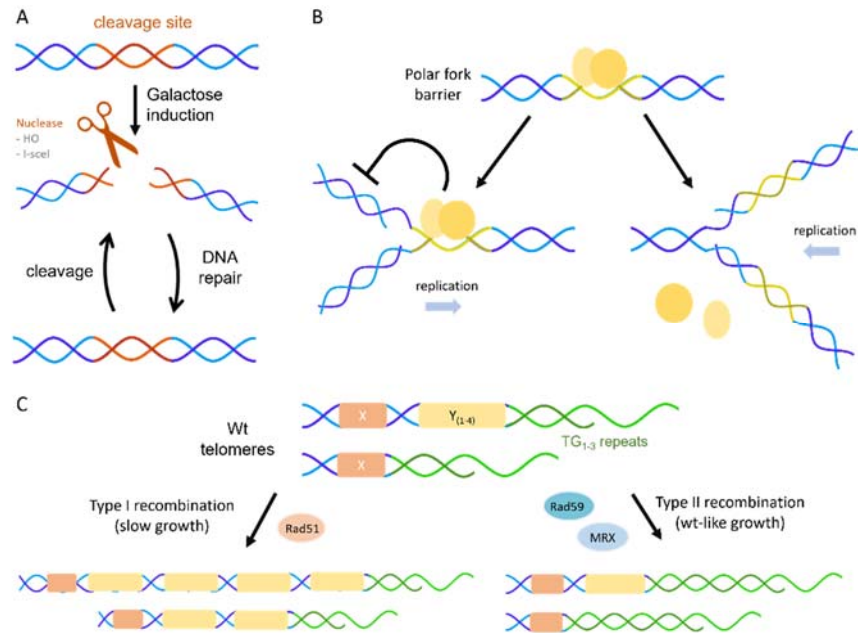


Figure 8. **Strategies to induce genetic instability.** A. DSBs generation: induction of the HO or I-SceI nuclease causes DNA cleavage. Reconstitution of the site by DNA repair leads to repeated cycles of cleavage and ligation. B. A polar replication fork barriers causes the stalling of the replication fork, in only one direction of replication. C. Telomere erosion and appearance of survivors (Adapted from Claussin and Chang, 2015). When critical telomere shortening is reached in absence of telomerase, cells enter replicative senescence. Alternative mechanisms of telomere lengthening by HR between telomere ends can give rise to survivors: type I survivors arise upon HR using Y repeats as template, while the use of TG repeats give rise to the fitter type II survivors.

Telomere erosion

Telomerase is required for the maintenance of telomere length throughout generations. Its depletion allows to better understand the mechanisms underlying telomere maintenance, since it causes the progressive shortening of chromosome ends, until the critical length is reached. This happens after around 50 generations, and the cells enter a state of replicative senescence. Such a critical condition can be overcome, leading to the appearance of survivors, in which the remaining ends lacking the TG₁₋₃ repeats, resembling DSB ends, are repaired by Rad52-dependent recombination, re-establishing a viable telomere. Survivors can be distinguished between type-I and type-II, according to the template used for recombination, respectively the sub-telomeric Y elements or the TG₁₋₃ repeats (Fig. 8C). Rad51-dependent BIR is involved for telomere maintenance in type I survivors, which arise more frequently but grow very poorly, while type II survivors require the MRX complex and Rad59 to arise, and their growth is comparable to telomerase-positive cells. Human alternative lengthening of telomeres (ALT) cancer cells similarly maintain their telomeres using recombination-mediated mechanisms requiring the MRN complex and the Sgs1-homolog, BLM, generating long, heterogeneous-sized telomeres (reviewed in Claussin and Chang, 2015).

1.3.2 | Chromatin dynamics upon DSB formation

Multiple studies in yeast have shown how upon formation of a DSB, DNA motility is increased both locally at the site of damage and more broadly in the genome, possibly favouring homology search for HR mediated repair (Dion et al., 2012; Miné-Hattab and Rothstein, 2012). The increased mobility is accompanied by increased stiffness of the chromatin fibre, which may enhance the ability of the break arms to explore the chromatin during homology search (Herbert et al., 2017; Miné-Hattab et al., 2017). Similarly, increased loci mobility and chromatin relaxation has been observed also in mammalian and fly nuclei upon DSBs formation (Chiolo et al., 2013). Modification on loci mobility following DNA damage has been proposed to depend on changes in higher order chromosomal conformation, e.g. due to histone H2A phosphorylation, which increases the negative charge of the fibre, contributing to the increase in stiffness (García Fernández et al., 2021). Moreover, detection of phosphorylated histone H2AX in mammalian cells (Aten et al., 2004) or Rad52 foci in yeast (Lisby et al., 2003) highlighted the tendency of simultaneously formed DSBs to transiently cluster together at early stages upon damage formation. Such vicinity represents a high risk of gross chromosomal rearrangement if the wrong ends are reunited during repair, particularly in regions enriched in repeated sequences.

Interestingly, in flies, heterochromatic DSBs have been shown to depend on nuclear actin and myosin for their relocation outside of the heterochromatin domain (Caridi et al., 2018). Moreover, F-actin has been shown to be involved in stressed replication foci and telomeric DSBs mobility towards the nuclear periphery in mammalian cells (Lamm et al., 2020; Pinzaru et al., 2020). Conversely, microtubules and the kinesin-14 motor protein complex have a role in promoting the mobility of telomeric DSBs to and away from the Nup84 complex in budding yeast (Chung et al., 2015). DSB mobility has also been shown to depend on the linker of nucleoskeleton and cytoskeleton (LINC) complex and dynamic microtubules also in fission yeast (Swartz et al., 2014) and human cells (Lottersberger et al., 2015). Curiously, although the involvement of the cytoskeleton suggests a directed movement, single particle motion analysis describe a movement which is compatible with random diffusion (Chung et al., 2015; Lottersberger et al., 2015). This paradox has been explained by the observation in budding yeast of the formation of damaged-induced microtubules (DIM) which themselves move within the nuclear space (Oshidari et al., 2018). Taking into consideration both the DNA and DIM mobility, the movement appear to be directional, thus adding an additional layer of complexity to the mechanisms underlying the dynamics of lesions compartmentalisation in the nucleus.

1.3.3 | Mechanisms driving damaged loci relocation to the nuclear periphery

In order to be successfully targeted to the nuclear periphery, lesions have to be recognised by dedicated factors, which determine the fate of the locus both in terms of localisation and choice of repair pathway. Different kind of lesions are sensed by different combination of factors, and changes in localisation can happen at different stages of the repair process. The different dynamics of relocation seem to reflect both the severeness of the lesion and the urgency for the isolation of such challenged loci from their original location, in order to limit their genotoxicity.

Notably, DNA breaks occurring at highly repetitive sequences are rapidly excluded from their compartment, in order to avoid clustering and unwanted recombination events with neighbouring, homologous loci. In metazoan cells, highly repetitive sequences (i.e. satellites and transposable elements) are grouped in a distinct nuclear domain, characterized by canonical heterochromatin markers, i.e. histone H3K9 methylation and Heterochromatin Protein 1 (HP1; HP1a in *Drosophila*). Interestingly, IR induced DSBs have been shown to relocate outside of this domain after γ H2A.X (γ H2Av in *Drosophila*) and ATRIP loading, but prior to Rad51 foci formation (Chiolo et al., 2011; Jakob et al., 2011). Notably, Rad51 appear to be actively excluded from the heterochromatin compartment by interaction with the Smc5/6 complex in *Drosophila* (Chiolo et al., 2011).

Similarly, DSBs forming at the rDNA locus in yeast have been shown to be excluded from the nucleolus prior to recruitment to the break of Rad52, which is excluded from the yeast nucleolus in a Mre11, Smc5/6 and SUMOylation dependent manner (Torres-Rosell et al., 2007). Relocation of both hetero-chromatic and nucleolar lesions outside their respective domain require checkpoint activation (Chiolo et al., 2011; Horigome et al., 2019). In mammalian cells, ribosomal DSB are mobilised at the periphery of the nucleolus, after resection, with a mechanism involving the nuclear envelope-associated LINC complex and the actin pathway (Marnef et al., 2019).

Strikingly, another kind of repeated sequences, i.e. long stretches of triplet repeats, susceptible to form hairpins, have been shown in yeast to relocate to the nuclear pore, independently from DSBs formation (Fig. 9, *center*) (Su et al., 2015). Notably, CAG repeats peripheral localization require ongoing replication at the repeats, but not Rad51 nor the checkpoint kinases Mec1/Tel1, consistent with the fact that their relocation is consequent of replication forks stalling at CAG repeats, but can happen prior to fork collapse (Su et al., 2015). However, the requirement for Sgs1, Exo1 and the exonuclease activity of Mre11 point to a role for long range resection (and possibly the subsequent RPA loading) in mediating its relocation to the periphery (Whalen et al., 2020). This supports the hypothesis that stalled forked relocation would be triggered by the formation of a “stuck” intermediate unable to be resolved, e.g. reversed forks (Freudenreich and Su, 2016).

Consistently, relocation of persistent DSBs lacking homology donor (Fig. 9, *left*) have been shown to require checkpoint activation (Nagai et al., 2008; Dion et al., 2012), indicating that relocation interests stalled repair intermediates that require

an adequate takeover to be resolved. Similarly, in telomerase negative yeast cells, short telomeres colocalise with both checkpoint and recombination factors; Mre11 and Rad52 coated telomere ends have been shown to colocalise with nuclear pores (Fig. 9, right), likely as a result of extended resection or replication stress at eroded ends (Khadaroo et al., 2009). Conversely, stalled forks at unique loci in *S. pombe* have been shown to require Rad51 dependent remodelling of the stalled intermediate to trigger anchoring to the NPC (Kramarz et al., 2020), pointing to the existence of differentiated relocation mechanisms depending on the nature and severity of the lesion.

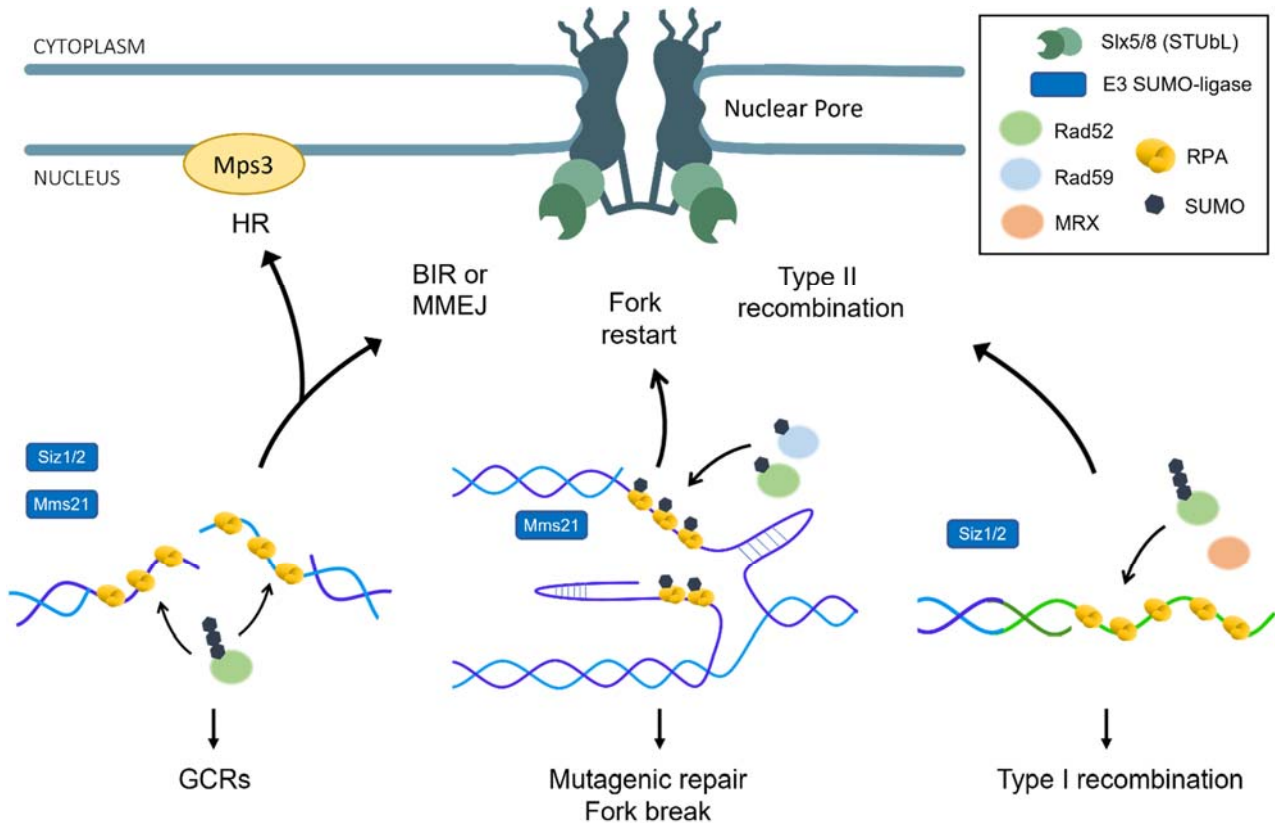


Figure 9. Schematic representation of challenged loci relocation to the nuclear periphery described in budding yeast. Unrepairable DSBs, stalled replication forks spanning TNRs, or short telomeres in telomerase negative cells are coated by RPA (in yellow). Mono- or poly-SUMOylation of repair factors, i.e., RPA and/or Rad52, by different SUMO-ligases is necessary for targeting the lesions to the periphery. Anchorage at the nuclear pore is mediated by the SIM-containing NPC-associated STUbL Slx5/8. Targeting to the nuclear pore channels the lesions towards non canonical repair pathways, and can promote fork restart, avoiding genetic instability. DSBs can also relocate to the INM protein Mps3, where more conservative repair pathways are promoted. Adapted from (Freudenreich and Su, 2016).

Nucleoporins involved in lesion relocation

Once sensed by the dedicated factors, lesions are relocated to the nuclear periphery, where they can interact with the nuclear pore and nuclear pore associated factors. Ribosomal DNA peripheral relocation involve interaction with the nucleoporins Nup84 and Nup120 (Horigome et al., 2019); the Nup84 complex is also involved in the peripheral localisation of persistent DSBs (Nagai et al., 2008) and DSBs forming at telomeric ends (Chung et al., 2015). Heterochromatic DSBs, instead, are stabilized at the periphery through interactions with the Smc5/6

complex and the NPC associated STUbL Dgrn/dRad60 (Ryu et al., 2015). In yeasts, the STUbL Slx5/8 is required for the anchoring of replication forks stalled at fork barriers or at CAG repeats (Kramarz et al., 2020; Su et al., 2015). The latter further depend on the Nup84 complex and the C-terminal domain of Nup1 (Su et al., 2015; Aguilera et al., 2020). Interestingly, Nup84 mutants were also reported to have increased genetic instability (Therizols et al., 2006; Palancade et al., 2007).

Persistent DSBs have multiple NE docking sites

Interestingly, the relocation of persistent HO-induced DSBs to the nuclear periphery is not restricted to anchoring to the nuclear pore, but unrepairable lesions can also be targeted to other nuclear envelope proteins, notably the inner nuclear membrane protein Mps3 (Horigome et al., 2014). Targeting to either site is cell cycle dependent: Mps3 anchoring is limited to S-phase while association to the nuclear pore is possible at every cell cycle stage. Chromatin remodellers contribute to determine the choice of the target: G1 recruitment of DSBs to the nuclear pore require Htz1 deposition by the SWR-C complex, while INO80 favour targeting to Mps3 (Horigome et al., 2014). While LexA-mediated targeting of Swr1, Arp6, and Htz1 was sufficient to trigger peripheral localization, the targeting of INO80 did not shift the locus to the periphery, indicating that while SWR-C mediated Htz1 deposition has a more structural role in loci-NPC interaction, INO80 function in the relocation may rather implicate its catalytic activity in chromatin remodelling by nucleosome eviction to promote resection (Horigome et al., 2014).

1.3.4 | Role of SUMOylation in lesion relocation and repair activation

By comparing the various mechanisms identified for the peripheral repositioning of stalled intermediates and lesions, the importance of SUMOylation in orchestrating their spatial redistribution and ensure their proper takeover becomes immediately evident. Indeed, SUMOylated repair factors or SIM-containing nuclear pore partners are directly involved in each of the characterised pathways, but each time in a slightly different flavour.

The SUMO E3 ligases dPIAS recruitment to heterochromatic DSBs precedes H2Av spreading, and SUMOylation from both Nse2 and dPIAS are required for heterochromatic DSBs relocation outside of the heterochromatin domain (Ryu et al., 2015, 2016). In this context, SUMOylation at early stage of DSBs signalling is important to prevent Rad51 loading on the lesion, which can occur only once the locus reach the periphery, to avoid premature engagement in the repair process while the DSBs ends are still close to heterologous homologous repeats (Ryu et al., 2015, 2016).

Similarly, a requirement for concurrent Mms21-mediated mono-SUMOylation of Rad52, Rad59 and RPA was scored for CAG spanning forks relocation to the NPC (Whalen et al., 2020). Again, SUMO-RPA binding seems to prevent Rad51 accessibility to the locus, preventing HR when the locus is far away from the pore, where the stalled intermediate can be properly processed once the SUMO-lock has been removed (Whalen et al., 2020).

In telomerase negative cells, SUMOylated telomere-bound proteins accumulate progressively as the telomere shorten, peaking during telomere-erosion-driven

crisis, concomitantly with the telomere shift from its heterochromatic environment to the nuclear pore, to then gradually decrease during appearance of type-II survivors (Churikov et al., 2016). Telomere relocation to the pore correlates with increased RPA SUMOylation, and is prevented by the inactivation of the Siz1 and Siz2 E3 ligases (Churikov et al., 2016). It is important to note, however, that the involvement of Mms21 in the process has not been investigated, so its contribution cannot be formally excluded; moreover, there is no reported evidence indicating that RPA is the direct, unique substrate of Siz2, leaving open the hypothesis that the two ligases could contribute together to telomere relocation.

The three E3 ligases have been shown to cooperate in the targeting of loci to the pore of persistent, unrepairable DSBs. Strikingly, the choice of target between the pore and Mps3 appear to be determined by the degree of SUMOylation of the lesion associated factors: while NPC association require poly-SUMOylation, mono-SUMOylation is sufficient to target a DSB to Mps3 in S phase cells. Consistently, Siz2 deletion affected predominantly association to the pore regardless of cell cycle phase, while association to Mps3 remained unaltered (Horigome et al., 2016), supporting a model in which the two ligases sequentially intervene on the substrate, with Mms21 responsible for the initial SUMOylation event, while the SUMO-chain is only subsequently deposited by Siz2.

Interestingly, polySUMOylation from the PIAS family E3 ligase Pli1 is required also for relocation to the NPC of forks stalled by protein fork barriers (Kramarz et al., 2020). Its blocking effect on HR-dependent DNA synthesis, which is necessary for efficient fork restart, is relieved post-anchoring, thanks to the eviction of the SUMO chain by Ulp1, stabilized at the pore by the nucleoporin Nup132 (Kramarz et al., 2020).

**Relevance of STUbL
in peripheral
anchoring and
repair efficiency**

As already hinted, SUMOylated lesion-associated factors are recognised by SIM-containing NPC partners, which fulfil both a structural role as docking site bridging the interaction with the nuclear pore, to which they are stably associated, but that also have a functional role in the repair pathway. The main actor in this respect is the STUbL Slx5/8 (Slx8 in *Sc. pombe*, Dgrn in *D. melanogaster*, RNF4 in mammals). Indeed, DSBs relocation is lost in *slx5/8* mutants in G1 cells, while DSBs are still partially able to relocate to the periphery in S phase in *slx5* cells (Horigome et al., 2016). Such residual relocation is accounted by the binding to Mps3, which remains unchanged in *slx5/8* mutants despite the requirement for SUMOylation, indicating the presence of other factors in the docking of SUMO-coated lesions to the periphery (Horigome et al., 2016). Incidentally, in *Drosophila*, localisation of Dgrn appear to be stabilised by both the nuclear pore and the Mps3 homologs Koi and Spag4 (Ryu et al., 2015).

Proximity to NPC-associated STUbLs seems to be essential for the removal of the lesion associated SUMOylated factors, freeing the locus for the following steps of HR mediated repair, that were hindered during the locus residency in repeats-rich compartments (Ryu et al., 2015, 2016). Consistently, eroded telomeres proximity to the NPC could determine the choice of repair pathways by modulating Rad52 SUMOylation. Indeed, SUMOylation of Rad52 promote its interaction with Rad51,

leading to type I recombination; Slx5/8 or Ulp1 mediated deSUMOylation of Rad52 at the nuclear pore may contribute to favour its interaction with Rad59, which in turns mediates type-II recombination, generating fitter survivors (Charifi et al., 2021). Further insight on the mode of action of STUbLs in mediating lesion relocation is given by experiments of artificial recruitment of the DSBs through a LexA-Slx5 fusion. In this context, the requirement of Siz2 for relocation is bypassed independently from the presence of Slx8, indicating that is substrate recognition, not substrate ubiquitylation, the main mode of action of slx5/8 in the mediation of the interaction in this context (Horigome et al., 2016).

The ability of NPC associated STUbL to recognise SUMOylated factors, with a preference for poly-SUMO-chains (Mullen and Brill, 2008) and its role in the remodelling of the lesion site to prime it for the following steps of repair is conserved among all the organisms characterised. STUbL mediated anchoring functions as a partition between the pre-anchoring stage, in which the stalled intermediated must be protected from aberrant repair, and the post-anchoring phase, in which the lesion gains access to the appropriate machinery to be repaired in the most conservative manner possible.

1.3.5 | Physiological consequences of lesion compartmentalisation at the nuclear periphery

Altogether these data point to a model according to which association of SUMOylated proteins to DNA lesions occurring at critical loci prevent aberrant recombination by triggering relocation to the nuclear periphery, where STUbL mediated eviction of such factors allow the accessibility of the problematic locus to the appropriate repair mechanism, promoting repair. Notably, error-prone recombination-based repair pathways can be promoted at the periphery, such as Break Induced Replication (BIR) and Microhomology Mediated End Joining (MMEJ) when the canonical conservative pathways are unavailable, in order to preserve chromosome integrity, at the expenses of information integrity (Freudenreich and Su, 2016). Consistently with a role for the nuclear pore to limit fragility, deletion of Nup84, leads to increased instability at CAG repeats containing loci, and GCRs in a repeat-length proportional fashion (Su et al., 2015). Interestingly, the length of the CAG repeat stretch is directly proportional to the percentage of peripheral loci and to the extent of repeat fragility in mutants impairing relocation, indicating that the harder it is to solve the non-canonical DNA structure, the more peripheral localisation is necessary to prevent repeat instability (Su et al., 2015).

Remarkably, each type of lesion appears to have its own dedicated pathway, with distinct, but partially overlapping, mechanisms for sensing, targeting and anchoring to the periphery depending on the nature, severity, and repair needs of the lesion. This level of fine control is possible thanks to the plasticity and versatility of SUMOylation, which acts as a recognition signal for lesions, ensuring their correct treatment and governing accessibility to certain factors at each step of the process. DSBs targeting to either the pore or Mps3 has different outcomes for the repair of the lesion: in G1, where NHEJ is the predominant repair pathway, lesions lacking

the homology donor are targeted to the pore independently from resection (Horigome et al., 2014). Consistently, nucleoporin mutants exhibit NHEJ defects (Palancade et al., 2007).

In S-phase, when HR become progressively predominant, binding to both targets is possible: to this regard, the increase in DSBs binding to the pore scored in *mps3Δ65-145* cells (lacking Mps3 nucleoplasmic protruding anchorage domain) points to a competition between the two possible anchoring sites, whereas the additive effect in the rates of unequal sister chromatid recombination in the double *nup120Δ mps3Δ65-145* indicate that the two anchoring sites have distinct role in repair pathway (Horigome et al., 2014). Notably, the nuclear pore microenvironment appears more permissive to error-prone recombination-based pathways, which are instead repressed in Mps3 proximity (Horigome et al., 2014; Oza et al., 2009). Indeed, when eroded telomeres, whose physiological peripheral localisation depends also from Mps3 anchoring, are prevented to relocate to the NPC, this decreases the occurrence of type-II recombination, necessary for the appearance of survivors (Churikov et al., 2016).

Spatial compartmentalisation of lesions away from their native location is particularly important in the context of highly repeated regions, i.e. heterochromatic or ribosomal DSBs. Unequal sister chromatid recombination in such environment can lead to chromosome translocation or even formation of dicentric or acentric chromosomes, thus compromising cell division and survival. Rapid relocation of these lesions outside of their compartment may be therefore essential to isolate them from the surrounding chromatin and avoid unwanted recombination events to preserve genome integrity and repeat stability.

It is also important to note that not all kind of challenged loci undergo repositioning, and the one that relocate do not move in the same way and towards the same location. Stalled replication forks due to dNTPs starvation, or that encounter a nick, do not relocate to the nuclear periphery unless collapse of the fork occur upon prolonged HU exposure or additional MMS treatment (Nagai et al., 2008; Dion et al., 2012). Similarly, breaks reparable by Single Strand Annealing do not relocate if short-range resection is sufficient for repair (Oza et al., 2009). Moreover, telomeric DSBs mobility in mammalian cells, is not systematically linked to peripheral localisation (Lottersberger et al., 2015), although an ATR-dependent relocation mechanism has been described in retinal pigment epithelial cells (Pinzaru et al., 2020). Finally, in mammalian cells alternative destination are possible for mobilised loci: PML bodies, for example, have been proposed as alternative hosts for HR dependent telomere elongation in ALT cells (Amaral et al., 2017).

This suggests that the decision on whether to relocate a locus or not may be determined by the establishment of a threshold of “hazardousness” of the lesion, which is signalled by the nature of the proteins associated with the locus itself, and their degree of SUMOylation. Once again, the nuclear pore is revealed to be a dynamic and versatile platform hosting critical processes for genome stability, ensuring their proper completion to preserve genome integrity and cell survival.

Chapter2

Biogenesis, regulation and functional relevance of R-loops

The first structural description of DNA:RNA hybrids dates back in 1967, when X-ray diffraction patterns of the heteroduplex allowed the description of its helical conformation, which is intermediate between the homoduplex RNA A-form and the DNA B-form (Milman et al., 1967). A decade later, the ability of two complementary single strand DNA and RNA molecules to establish stable Watson and Crick interactions was observed *in vitro* (Thomas et al., 1976).

Since then, DNA:RNA hybrid formation has been recognised as a major step in all essential processes required to ensure the faithful inheritance of genetic information. Indeed, DNA polymerases requires a short RNA primer, synthesised by a primase, in order to initiate the new strand replication (Murakami et al., 1992) or resume DNA synthesis at stalled or collapsed replication forks (reviewed in Yeeles et al., 2013). Recently, an increasing amount of evidence has highlighted the formation of DNA:RNA hybrids at sites of DNA damage. However, whether damage induced hybrids play a positive or negative role on the repair process is still a matter of debate (Marnef and Legube, 2021).

R-loops are three-stranded nucleic acid structures composed of a DNA:RNA hybrid and an exposed ssDNA filament (Fig. 10). They can form co-transcriptionally when the nascent transcript has the opportunity to reanneal to its template DNA displacing the non-template strand, thus preventing the DNA to reform the double helix.

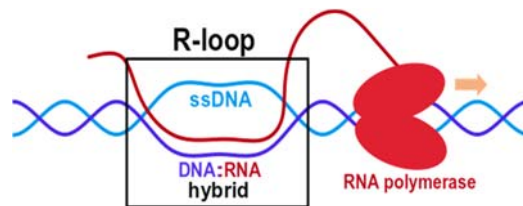


Figure 10: schematic representation of an R-loop

Their size ranges around 150–500 bp as observed by electron microscopy (Duquette et al., 2004) and confirmed by single molecule R-loop foot-printing (García-Pichardo et al., 2017; Malig et al., 2020). R-loops occupy up to 5, 8, and 10% of the human, yeast, and Arabidopsis genomes, respectively (Sanz et al., 2016; Wahba et al., 2016; Xu et al., 2017), and can be found in protein coding genes but also in rDNA and tRNA genes, indicating that they can form during transcription regardless of the polymerase involved (Santos-Pereira and Aguilera, 2015). Notably, R-loop formation is not due to the accidental extension of the short hybrid forming within the polymerase as a primer for transcription: the DNA and the RNA exit the polymerase from two distinct channels (Westover et al., 2004), and annealing happen only after, upon changes in DNA or RNA conformation that favour RNA invasion of the duplex.

Thanks to the development of increasingly sophisticated technologies for detecting R-loops (discussed in paragraph 2.2), great progress has been made in understanding their properties and roles in genome homeostasis. General features

of R-loops are strand specificity, determined by the co-directionality with transcription, co-dependence with transcription levels, and a propensity to form at sequences with peculiar features favouring hybrid stability.

Despite the necessity and pervasiveness of DNA:RNA hybrid formation, excessive or unregulated hybrid generation can compromise genome integrity. In this respect, R-loops represent a major threat for genetic stability, although in some instances their controlled formation has been shown to be required for physiological processes.

In the next paragraphs, I will detail the main features of R-loops, the circumstances favouring their formation, and the tools currently available for their detection. An overview of the physiological roles of R-loops will be followed by the description of the mechanisms and tools that cells evolved to regulate the formation of these structures. Finally, I will provide a brief description of the ways in which R-loop formation can be dangerous to cellular homeostasis, as well as contributing to the aetiology of numerous human diseases.

2.1 | Biogenesis of R-loops

R-loop formation is a result of the competition between the RNA and the non-template DNA for the reannealing to the template strand. Several factors have been proposed to favour the RNA annealing, to the detriment of the double helix reconstitution.

2.1.1 | R-loop prone sequences

Thermodynamically, a DNA:RNA hybrid appear to be more stable than a DNA homoduplex, especially if the RNA is purine-rich (Roberts and Crothers, 1992). Indeed, R-loops tend to form at sequences rich in GC, and characterised by an asymmetry in the GC distribution between the two DNA strand (GC-skew). Consistently, in immunoglobulin (Ig) class switch sequences, R-loop formation seems to originate at G clusters (Roy et al., 2008). Moreover, *in vitro* and *in vivo* observations in G-rich plasmid templates suggest that co-transcriptional formation of G-quadruplexes in the G-rich non template strand could stabilise the opening of the DNA and consequently the presence of the DNA:RNA hybrid (Duquette et al., 2004). Moreover, high-GC content and G-quadruplex formation have been found to correlate with longer R-loop lifetimes (Crossley et al., 2020).

R-loops and torsional stress

R-loop formation is also tightly linked with the topological conformation of the DNA molecule. During transcription, negative supercoiling forms upstream of the RNA polymerase: the consequent increase in DNA breathing may favour R-loop formation by facilitating the invasion of the DNA molecule by the RNA. Consistently, mutants for topoisomerases, i.e. the enzymes in charge of relieving topological stress, have been shown to accumulate R-loops in several species (Drolet et al., 1995; El Hage et al., 2010). Strikingly, R-loop maps in human cells depleted of topoisomerase I, showed not only R-loop gain at long, highly transcribed, isolated genes, but a surprising loss of R-loop at loci associated to early firing replication origins (Manzo et al., 2018), highlighting the importance of the chromatin context in the determination of hybrid fate.

Interestingly, R-loop formation has been proposed to absorb the DNA rotational stress thus relieving DNA super-helicity and potentially preventing further hybrid formation. Regions characterised by intense topological stress may therefore form R-loops regardless of the sequence composition (i.e. GC content/skew) of the locus, thus emerging as alternative R-loop formation determinant (reviewed in Chedin and Benham, 2020).

ssDNA lesions

Moreover, the presence of nicks in the non-template strand can also serve as an efficient R-loop initiation site: the subsequent conformational variation of the DNA molecule may contribute to promote the hybridisation of RNA to the template DNA, to the detriment of the non-template strand reannealing to its complementary template strand, as observed *in vitro* (Drolet, 2006; Roy et al., 2010).

2.1.2 | R-loop 3D structures

Atomic Force Microscopy observations of murine R-loop forming loci revealed the potential of R-loops to adopt different 3D conformations, denominated blobs, spurs and loops (Neaves et al., 2009; Carrasco-Salas et al., 2019). The differential R-loop architecture seems to depend on the 3D conformation of the ssDNA displaced filament, and to have an influence on the conformation of the adjacent dsDNA by causing kinks with specific angles (Carrasco-Salas et al., 2019). Notably, while the differential R-loop architecture observed at the murine immunoglobulin Sy3 switch region was shown to depend on G-quadruplexes forming at the ssDNA (Neaves et al., 2009), such structures were not observed to be required for the same architecture forming at the *Airn* locus (Carrasco-Salas et al., 2019). The formation of such 3D structures may have an influence in the R-loop fate and stability, since its resolution would require an additional step of unfolding prior to unwinding or degradation. Moreover, different 3D structures may lead to different consequences, with some potentially being more genotoxic than others.

2.1.3 | Post-transcriptional R-loops

Although the majority of R-loops are believed to be co-translational, forming in *cis* through the reannealing of the nascent mRNA to its DNA template, several examples of post-transcriptional formation of R-loops have also been observed across organisms. Notably, R-loop formation in *trans* seems to require strand exchange mechanisms performed by proteins of the recombination machinery. Indeed, bacterial protein RecA (ortholog of eukaryotic Rad51) has been shown *in vitro* to promote homology-dependent RNA incorporation in a DNA duplex, thus forming an R-loop (Kasahara et al., 2000; Zaitsev and Kowalczykowski, 2000). This observation supported a model of RecA-dependent hybrid formation that had been postulated as an alternative mechanism to initiate DNA replication (Cao and Kogoma, 1993; Hong et al., 1995). Moreover, *in vivo* evidence in budding yeast highlighted the ability of Rad51 to stimulate R-loop formation by promoting DNA-RNA strand exchange, thus increasing genetic instability (Wahba et al., 2013).

TERRA R-loops Rad51 also seems to be involved in promoting the post-transcriptional hybridization of the long noncoding RNA TERRA at telomeres in human cells (Feretzaki et al., 2020). TERRA (Telomeric Repeat-containing RNA) is transcribed by RNA polymerase II from promoters residing in subtelomeric regions and terminates at the telomeric repeats tract. In telomerase negative HR-proficient yeast cells, formation of DNA:RNA hybrids at shortened or damaged telomeres is crucial to prevent telomere shortening, possibly by promoting homology directed repair, contributing to the maintenance of telomere homeostasis and delaying senescence (Balk et al., 2013).

lncRNA form R-loops in trans Moreover, in *S. cerevisiae*, two long noncoding RNAs have been shown to be transcribed at the *GAL* locus in repressed conditions. Both lncRNA originates from the 3' end of *GAL10*, one encompasses *GAL10* and *GAL1* ORFs, while the other overlaps with *GAL7* promoter. Such lncRNAs have been shown to form R-loop post-transcriptionally at the *GAL* locus, thus promoting faster activation of the galactose-

inducible genes upon carbon source shift, conferring a competitive advantage for metabolic adaptation (Cloutier et al., 2016).

Furthermore, the plant lncRNA APOLO (AUXIN REGULATED PROMOTER LOOP) has been shown to coordinate the transcription of a network of 182 auxin responsive genes by R-loop formation in *trans* at the promoter of the target loci, recognised through sequence homology (Ariel et al., 2020).

CRISPR-Cas system Another mechanism of controlled hybrid formation in *trans* is carried out by the CRISPR-Cas9 system (clustered regularly interspaced short palindromic repeats) system, a bacterial defence mechanism against phages and plasmid transfer, widely exploited in molecular biology as a high-precision tool for gene editing. This system exploits the dual DNA endonuclease Cas9 protein, which recognises its target thanks to the presence of a short Protospacer Adjacent Motif (PAM). Once the target sequence is identified, a DNA:RNA hybrid is formed within the protein between the target DNA and its complementary Cas9 associated guide RNA, displacing the non-complementary ssDNA strand. Formation of the hybrid lead to conformational changes conferring competence for cleavage to Cas9, which operates a cut in both DNA strands generating a blunt-end DSB (Jiang and Doudna, 2017).

2.2 | Strategies for R-loop detection

Pioneering observations of R-loops relied on the use of electron microscopy, with the first map obtained for the *Drosophila* 18S and 28S rDNA almost half a century ago (Glover and Hogness, 1977; White and Hogness, 1977). A major breakthrough for R-loop detection has been the production and the characterization of the S9.6 monoclonal antibody, which presents a high affinity for DNA:RNA hybrids (Boguslawski et al., 1986) of a size of as little as 6 base pairs (Phillips et al., 2013). Although the specificity of this antibody is still debated (discussed more in detail below), it remains the gold standard for R-loop detection up to date. In the last two decades, a wide range of techniques have been developed to allow the precise, specific mapping and quantification of R-loops. In addition to the antibody-based techniques, new strategies have been introduced to map ssDNA exposure, which exploit either bisulfite deamination activity or the binding domain of RNase H1 ribonuclease, which specifically recognise DNA:RNA hybrids and degrade the RNA moiety.

2.2.1 | S9.6 antibody-based methods

The S9.6 antibody can be used in a variety of assays, both *in vivo* and *ex vivo*, to detect and quantify both locus-specific and global changes in R-loop levels. The latter can be achieved by dotblot or immunofluorescence, while DNA:RNA hybrids Immunoprecipitation (DRIP) is the most widely used technique to detect R-loop levels in a locus specific manner.

DNA:RNA hybrid immunoprecipitation

Typical DRIP experiments consist in the extraction of deproteinised genomic DNA followed by fragmentation by either sonication or enzymatic restriction. DNA:RNA hybrids are pulled down using the S9.6 antibody, adopting a standard IP protocol, followed by DNA analysis by qPCR, tiling microarray or high throughput sequencing (Fig. 11A). Over the years, this technique underwent several optimizations that significantly improved the specificity, resolution and signal-to-noise ratio, allowing to obtain reliable genome-wide R-loop profiles in several organisms, from yeast (Chan et al., 2014; Wahba et al., 2016), to mammals (Ginno et al., 2012; Chen et al., 2015) and plants (Xu et al., 2017). Removal of the ssDNA through S1 nuclease treatment, for example, has been used to stabilize the hybrids during sonication, improving the yield of the IP (Fig. 11B; Wahba et al., 2016). A significant improvement for DRIP approaches has been the optimization of the library preparation step in order to acquire information about the orientation of the DNA:RNA hybrid. Strand specificity have been achieved either by incorporating dUTPs during the second-strand synthesis step (RDIP-seq; Nadel et al., 2015) or by introducing the first sequencing adaptor specifically at the 3' end of the DNA fragments (ssDNA-seq; Xu et al., 2017). Finally, the DRIPc-seq method involve sequencing of the RNA fragment retrotranscribed to cDNA (fig. 11C; Sanz et al., 2016). Acquiring information about R-loops strand directionality allowed a more precise correlation between R-loop maps and genomic features and improved the understanding of R-loop relationships with transcription and chromatin organization.

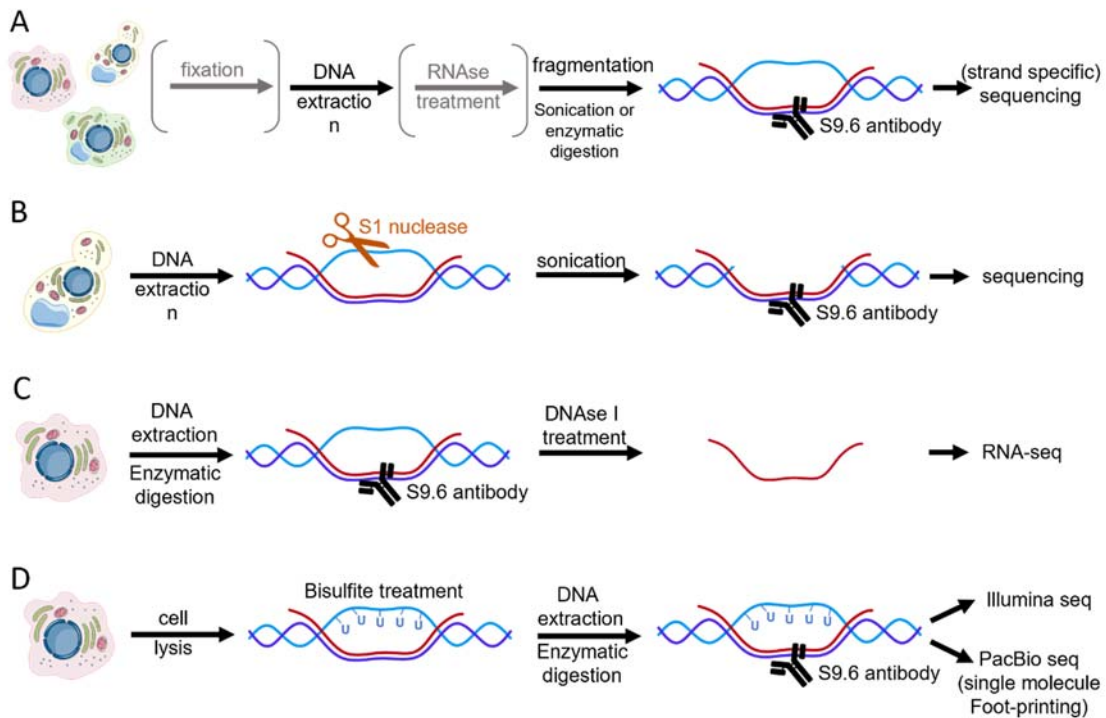


Figure 11. **Schematic representation of S9.6-based R-loop detection approaches.** A. Schematic representation of a generic protocol for DNA:RNA hybrid Immunoprecipitation using the S9.6 antibody: several variations have been applied over the years including or not recommending intermediate steps, indicated in grey (further discussed in paragraph 2.2.3). B. DRIPc-seq protocol: strand directionality is achieved by sequencing the RNA moiety of the hybrid after R-loop IP (Sanz et al., 2016). C. S1-DRIP-seq protocol: ssDNA degradation by the S1 nuclease prior to fragmentation preserve hybrids integrity during sonication, increasing the yield of the IP (Wahba et al., 2016). D. Bisulfite sequencing/SMRF-seq schematic protocol: bisulfite treatment during cell lysis allows the deamination of the cytosin on the displaced ssDNA filament, which can be mapped by sequencing (Dumelie and Jaffrey, 2017); the employment of long reads sequencing allows single molecule footprinting (Malig et al., 2020).

Bisulfite treatment Another method to map R-loops is native bisulfite treatment, which was used to obtain the first *in vivo* evidence of R-loop formation in eukaryotic nuclei (Yu et al., 2003). This approach, classically used for detection of DNA methylation, exploits the ability of sodium bisulfite to convert Cytidine to Uracil by deamination. Since this chemical acts exclusively on ssDNA, bisulfite treatment on non-denatured genomic DNA leads to C to U conversion only at already exposed ssDNA stretches, e.g. the displaced non-template DNA strand of an R-loop (Fig. 11D). This technique is still widely used to obtain strand-specific R-loop maps, also combined with S9.6 IP to increase signal-to-noise ratio (bis-DRIP-seq) (Dumelie and Jaffrey, 2017). The recent development of single molecule sequencing technology, in particular PacBio real time sequencing, has improved the throughput and the resolution of this approach, allowing Single Molecule R-loop Footprinting (SMRF-seq) (Malig et al., 2020). Bisulfite sequencing constitutes an optimal tool to determine distribution, position, and size of R-loops in relation to the DNA sequence, with near nucleotide resolution, deepening the understanding of R-loop structure and heterogeneity and its relationships with the genomic landscape. A major limitation of this genre of technique is its dependency on the Cytosines distribution along the DNA sequence, which may lead to underrepresentation of shorter R-loops, limit the information on the hybrid boundaries and introduce a bias towards GC-rich sequences (Castillo-Guzman and Chédin, 2021).

2.2.2 | RNase H based methods

As an alternative to the S9.6 antibody, several groups exploited RNase H ability to bind DNA:RNA hybrids: a catalytic inactive but binding competent RNase H1 mutant has been used to detect R-loops by pull-down assays (DRIVE-seq - Ginno et al., 2012), fluorescence microscopy (Bhatia et al., 2014) or Immunoprecipitation (R-ChIP - Legros et al., 2014; Chen et al., 2017). R-loop IPs using RNase H1 as bait have been used to map R-loops in fission yeast (Legros et al., 2014) and mammalian genomes (Ginno et al., 2012; Chen et al., 2017). Recently, a new set of approaches relying on the ability of the RNase H hybrid binding domain (HBD) to efficiently recognise DNA:RNA hybrids has been developed to further increase the resolution of R-loop detection and improve the protocol feasibility.

Detection of native R-loops (*in vivo* detection in absence of fixation and subsequent DNA shearing) was achieved by adapting CUT&RUN (Cleavage Under Targets and Release Using Nuclease) and CUT&TAG (Cleavage Under Targets and Tagmentation) protocols in order to “tether” a cleaving enzyme to RNase H-labelled chromatin loci, thus inducing the cleavage of the hybrids which are then retrieved for sequencing.

mapR R-loop CUT&RUN exploits the catalytically inactive RNase H to target the micrococcal nuclease (MNase) at hybrid forming loci (Fig. 12A). The deadRNaseH-MNase fusion is delivered by diffusion inside the nucleus after cell permeabilization, and the nuclease activity is induced by addition of calcium. Binding of the fusion protein to a DNA:RNA hybrid leads to its cleavage and consequent release from the nucleus by diffusion. This strategy allows to bypass the need of stable *in vivo* expression of RNase H and the chromatin manipulation steps required for the affinity purification of DNA:RNA hybrid in DRIP or R-ChIP methods. However, possible bias due to hypersensitivity of chromatin to cleavage by MNase, which shows a slight preference for AT rich regions (Dingwall et al., 1981), has to be taken into consideration for experiment design. Moreover, MapR requires a lower amount of starting material, which may become highly beneficial in medical applications, notably for assessment in patient derived material (Yan et al., 2019). A recent development of the technique consisting in the combination of MapR with bisulfite sequencing allowed to introduce strand specificity, thus increasing the accuracy of the map (Wulfridge and Sarma, 2021). The mapR-based map obtained in human cells shows significant overlap with previous published maps, predictably stronger with RNase H based methods, but also detected a great amount of hybrids never identified by the other techniques, notably at the level of enhancers (Yan et al., 2019).

R-loop CUT&TAG R-loop CUT&TAG targets the Tn5 transposase to R-loop using an artificial DNA-RNA hybrid sensor consisting of tandem repeats of the HBD labelled with a glutathione S-transferase and an hexahistidine tag (GST-His6-2×HBD; Fig. 12B). The Tn5 transposase ability to cleave and tag with sequencing adaptors can be exerted to both dsDNA and DNA:RNA hybrids (Lu et al., 2020). R-loop maps obtained with this strategy significantly overlap with the other RNase H based methods and display the higher signal-to-noise ratio. In addition to the methodological advantages

already listed for CUT&RUN, the use of tagmentation simplify the library preparation, making the protocol even more rapid (Wang et al., 2021). Major drawbacks of this technology are the lack of strand specificity, and the potential aspect specific cleavage of open chromatin regions by the enzyme regardless of the presence of R-loops, which can be controlled by RNase A overexpression.

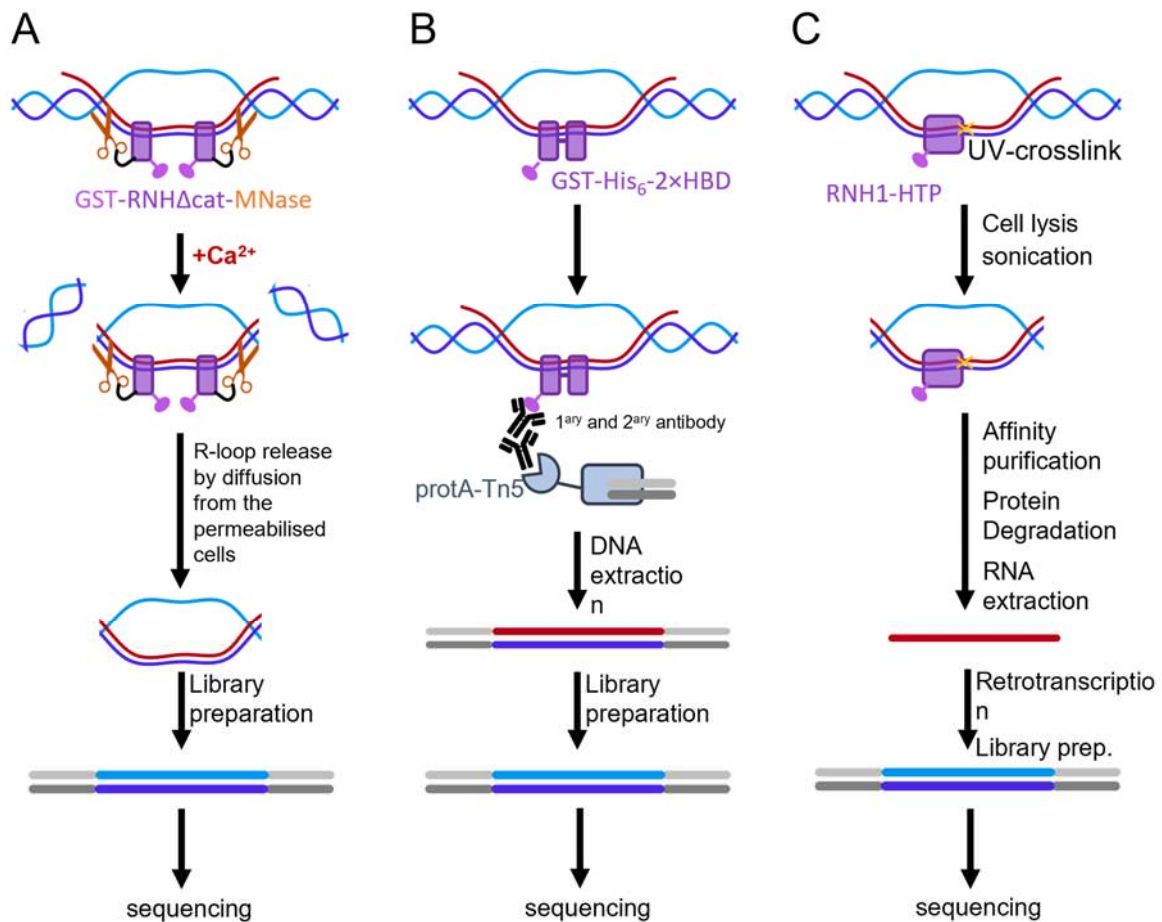


Figure 12. Schematic representation of RNase H-based R-loop detection meeting. A mapR protocol: micrococcal nuclease is targeted to R-loop by a fusion with a catalytic dead RNase H. Cell permeabilization allows the delivery of the fusion construct and the retrieval of the hybrids (Yan et al., 2019). B. CUT&TAG protocol: the Tn5 transposase is targeted to R-loops through an artificial DNA:RNA hybrid sensor (Wang et al., 2021). C. H-CRAC protocol: UV-crosslinking is exploited to bind RNase H to R-loops. The protein-nucleic acid complex can be retrieved by affinity purification; sequencing of the cDNA provides strand directionality (Aiello et al., 2022).

RNase H CRAC

A major limitation of R-loop detecting strategies is the ability to achieve strand specificity, which requires additional steps in the library preparation that are often complicated to implement and are sensitive to contamination from other RNA present in the preparation. To overcome this issue and obtain highly sensitive strand-specific R-loop mapping in yeast, the Libri group employed the CRAC technique (Crosslinking Analysis of cDNAs), which exploits *in vivo* UV crosslinking between the RNA and the protein of interest, followed by IP, retrotranscription, and sequencing (Bohnsack et al., 2012; Fig. 12C). H-CRAC using either RNaseH1 or RNase H2 as bait provided high resolution, strand specific R-loop maps, which appear similar to each other and showed a significant overlap with published R-loop maps in yeast (Aiello et al., 2022).

2.2.3 | Caveats of R-loop detecting strategies

The technological advancement occurred in the last decade has brought significant improvement in the precision and high resolution of R-loop detection. However, R-loop profiling methods still suffer from a large number of criticalities, and the R-loop profiles available up to date display inconsistencies and low reproducibility; it is therefore crucial to be mindful of strengths, limitations and specificities of each technique in order to identify the best strategy for investigating the biological question at hand. Importantly, the use of cell-based or synthetic spike-in as internal standard for cross-condition normalisation has been proposed to provide a more reliable, reproducible and comparable quantification of R-loops (Crossley et al., 2020). The implementation of such practice as standard procedure in future applications will further reduce the variability between samples and allow more precise quantitative analysis.

Differences between S9.6 and RNase H based strategies

A careful comparison between the various human R-loop maps available bring a better understanding not only on the source of differences between studies but also on the nature and characteristics of the R-loops.

The most evident controversy is the difference between the profiles obtained exploiting the S9.6 antibody and RNase H based strategies. Catalytic dead RNase H1-based or HBD-based maps detect relatively short hybrids mapping in GC-rich and GC-skewed promoter-proximal regions. Intergenic enhancer regions and tRNA loci are also identified as R-loop hotspots. Conversely, S9.6 antibody maps identify R-loops predominantly distributed on gene bodies, with higher signal downstream of CpG island promoters. The peaks identified with this technique appear longer (median of 1,5kb in human cells), although long hybrid stretches has been shown to be the sum of shorter R-loops forming all along the gene body, as detected by single molecule R-loop foot-printing (reviewed in Castillo-Guzman and Chédin, 2021). Although such differences were initially attributed, and could be partially explained, by different specificity of RNase H or the antibody for hybrid recognition, the main source of these discrepancies seems to rely on the different context in which the encounter between the probe and the hybrid occurs, and the different manipulations to which chromatin is subjected during hybrid enrichment and purification. While S9.6 binding to the DNA:RNA hybrid occur *ex vivo*, after DNA extraction and fragmentation, RNase H based methods rely on the *in situ*, *in vivo* tethering of the probe to the hybrid prior to eventual fixation, cell lysis and DNA shearing. Hybrid detection in more native conditions would therefore allow the detection of more transient R-loops, that would be lost upon the harsher chromatin treatment employed in S9.6 based methods, i.e. promoter proximal or enhancer-associated R-loops. In support of this, a recent comparative analysis between R-loop profiles obtained with either a CUT&RUN or a DRIPc-seq adapted protocol employing either the S9.6 antibody or an HBD-base artificial R-loop probe showed how the specificity of the R-loop map obtained was dependent on the protocol applied and not on the type of probe. Strikingly, adapted DRIPc-seq protocol using GST-His6-2×HBD as bait produce an R-loop map preferentially correlating with standard S9.6-based profiles, while S9.6 CUT&TAG succeeded in retrieving

promoter proximal DNA:RNA hybrid signal (Wang et al., 2021). Consistently, R-loop foot-printing, which also rely on *ex vivo* treatment of the DNA with bisulfite, show higher correlation with other *ex vivo* protocols rather than *in vivo* capture strategies.

R-loops fate during sample preparation

On top of this, the impact of sample manipulation on the stability of the R-loops is still a source of debate in the field: consensus on the best practices to ensure the preservation of the hybrids is not yet unanimous regarding chromatin fixation and DNA fragmentation for DRIP protocols. On the one hand, sonication of fixed chromatin is believed to lead to loss of R-loops (up to 80%), due to the harshness of the treatment, and S1 mediated ssDNA degradation has been shown to reduce hybrid loss (Wahba et al., 2016); on the other hand the distribution of fragment size obtained by sonication has been proposed to increase the resolution of R-loops detection and mapping compared to enzymatic restriction (Halász et al., 2017).

Finally, the deproteinization of the chromatin and the consequent dissociation of the polymerase could lead to *ex vivo* formation of R-loops during the sample preparation process, that would be otherwise unlikely due to the strong energy barriers that disfavour RNA invasion. If the short internal DNA:RNA hybrid survives the detachment of the polymerase, further RNA invasion of the DNA duplex could occur at loci already prone to R-loop formation due to their sequence composition or topological tension, thus generating artifacts that could bias the quantification particularly of R-loops mapping along the gene body (Belotserkovskii and Hanawalt, 2022).

Aspecificity of the S9.6 antibody

The aforementioned observation seems to point to the experimental conditions as the main source of variability between experiments. However, the probes used for hybrid capture are also susceptible to introduce significant biases during R-loop detection. The S9.6 antibody is not completely sequence aspecific (König et al., 2017) and a preference for GC-rich sequences has been recently reported, along with the inability of the antibody to bind short AU/AT R-loops (Bou-Nader et al., 2022). Such preference in binding is attributed to geometric requirements in the antigen binding site to accommodate the double helix that would not be fulfilled by specific subsets of probes. Moreover, qDRIP experiment identified a subset of DNA:RNA hybrids partially resistant to *ex vivo* RNaseH treatment, which correlated with GC-skewed regions mapping downstream of the TSS of genes (Crossley et al., 2020).

Another source of concern regarding the specificity of R-loop mapping methods is the ability of both S9.6 and RNase H to recognise and bind dsRNA (Phillips et al., 2013; Nowotny et al., 2008), which can introduce strong biases in particular in genome wide mapping techniques relying on the sequencing of the RNA moiety of the hybrid, or for immunofluorescence experiments. Although both display a higher affinity for DNA:RNA hybrids than dsRNA, this becomes rather irrelevant in *in vivo* situations where the amount of hybrids is relatively negligible compared to the amount of dsRNA present in the cells (Bou-Nader et al., 2022). This can therefore lead to identification of false positives, or to a bias in R-loop level quantification due

to the titration of the antibody by off-targets. This become particularly relevant when DNA:RNA hybrid isolation is followed by RNA-based sequencing techniques, i.e. DRIPc-seq, which was shown to yield RNase III sensitive, RNase H insensitive strand aspecific signal in *Sc. Pombe*, an organism in which dsRNA are physiologically produced at significant levels (Hartono et al., 2018). The binding of the antibody to dsRNA can also explain the predominant cytoplasmic signal that is detected in experiments of Immunofluorescence. This, together with the strong, RNase H insensitive signal detected at the nucleolus, make this technique unreliable and unsuitable for R-loop quantification (Smolka et al., 2021). Despite its comparable dual affinity for hybrids and dsRNA, catalytic inactive RNase H1 has been deemed preferable to the antibody for immunofluorescence applications, in which it performs with higher specificity and sensitivity (Crossley et al., 2021).

Best practices to ensure detection of bona fide R-loops

RNase H treated controls are crucial in any application to be able to discriminate between *bona fide* R-loops from other RNase H – resistant nucleic acid conformations. Moreover, pre-treatment of samples with RNase A and RNase III has been proposed as an additional step to decrease RNA contamination (Zhang et al., 2015; Hartono et al., 2018), although not all studies agree on the effect of this procedure (Halász et al., 2017).

Finally, since the minimal requirement for hybrid recognition being 4bp for RNase H and 6-8bp for the S9.6 antibody, DNA:RNA hybrid mapping strategies can potentially indiscriminately detect any kind of DNA:RNA hybrid, including primers of Okazaki fragments or DNA:RNA hybrid generated at damage sites. Checking the transcription, cell cycle dependency and strand specificity of the output could help to determine the nature of the hybrid detected in order to identify *bona fide* three stranded R-loops.

2.2.4 | R-loop levels manipulation strategies

In order to characterise R-loop features and the factors involved in their metabolism, a great variety of strategies is used to manipulate R-loop levels *in vivo*. However, many of these approaches interfere with key steps of the transcription process, i.e. splicing, elongation, RNA export, thus influencing R-loop levels or generating concomitant R-loop independent genetic instability. Such important biases need therefore to be carefully considered for the interpretation of the results. An example of this is exemplified by a recent study available on Biorxiv investigating the involvement of splicing in R-loop metabolism in human cells. Drug-induced splicing inhibition was shown to cause R-loop gain at a subset of stress-induced upregulated genes, which however was not directly linked to splicing but arose from termination defects consequent of the Pladienolide B treatment. Global R-loop levels instead experienced a general decrease, but that was also accompanied by transcription elongation defect and increased promoter proximal pausing due to the drug treatment, which are also significant determinant of R-loop levels and may have a great impact on the outcome regardless of intron retention (Castillo-Guzman et al., 2020). Notably, the result of this study contrasted with previous published observations in yeast, where local splicing perturbations by

insertion of an artificial intron in a model gene, which did not have such conspicuous effects on transcription dynamics, showed a preventing effect of splicing against R-loop formation (Bonnet et al., 2017).

The most widely used tool to manipulate R-loop levels is either depletion or overexpression of RNase H. Notably both overexpression of *E. coli* RnhA or deletion of endogenous RNase H in *Sc. Pombe* lead to transcriptome changes even at genes not forming R-loop (Hartono et al., 2018). Moreover, changes in the proteome, in particular downregulation of Top1 and factors of the NHEJ pathway, have been detected upon RNase H overexpression in Hela cells, with a consequent increase in DNA damage (Shen et al., 2017).

Mutant of the THO complex or the helicase Sen1/hsSETX are also used as R-loop forming mutants, with the caveat of significant impact on transcription elongation and RNA export for the former, and transcription termination for the latter. In this case, normalising for the actual transcription levels, e.g. by normalising to the polII ChIP or nascent mRNAs signal (Bonnet et al., 2017) could allow a more fair comparison between wt and mutant conditions, which otherwise would have a different R-loop forming potential determined by transcription efficiency.

To conclude, a plethora of strategies are available for high resolution, high throughput detection of R-loops. Different strategies contributed to highlight different properties and classes of R-loops, shedding additional light on the characteristics of these peculiar structures. While RNase H based methods seem the most suitable for the detection of highly dynamic, transient R-loops associated with transcription initiation events, S9.6 based strategy allow the detection of possibly rarer but more stable R-loops associated with transcription elongation. Finally, non-quantitative bisulfite foot-printing shed additional light on the size and position of R-loops molecules.

Despite all the limitation highlighted above, the array of available R-loop detection approaches constitutes a powerful set of tools for R-loop investigation. The more and more common practice of combining multiple techniques in parallel allows to overcome individual limitations. Moreover, the recent technological advances in single cell analysis and high throughput sequencing will provide great benefits for the advancement of the field.

2.3 | Physiological roles of R-loops

Initially considered as rare, short-lived by-products of transcription, R-loops have gained more and more relevance in many aspects of nuclear homeostasis. Indeed, they are associated to several fundamental biological processes like antibody class switch recombination, replication initiation, transcription regulation, and more.

2.3.1 | R-loop in immunoglobulin class switch recombination

The first *in vivo* evidence of R-loop formation in eukaryotic cells was in the context of their regulatory role in immunoglobulin Class Switch Recombination (CSR - Yu et al., 2003, reviewed in Yu and Lieber, 2019), the process underlying the switching of Ig isotype from IgM to IgG, IgE or IgA.

The mammalian *IGH* locus contains a tandem of C regions coding for the heavy chain of the different immunoglobulin isotypes, located downstream of the rearranged variable region (VDJ segment). Interspersed between the C regions are large, repetitive Switch regions, characterized by high GC content and GC skew and G clusters. CSR consist in the rearrangement of the locus through the excision of exons between S regions, placing the new C gene directly downstream of the exons coding for the Ig variable regions (Matthews et al., 2014).

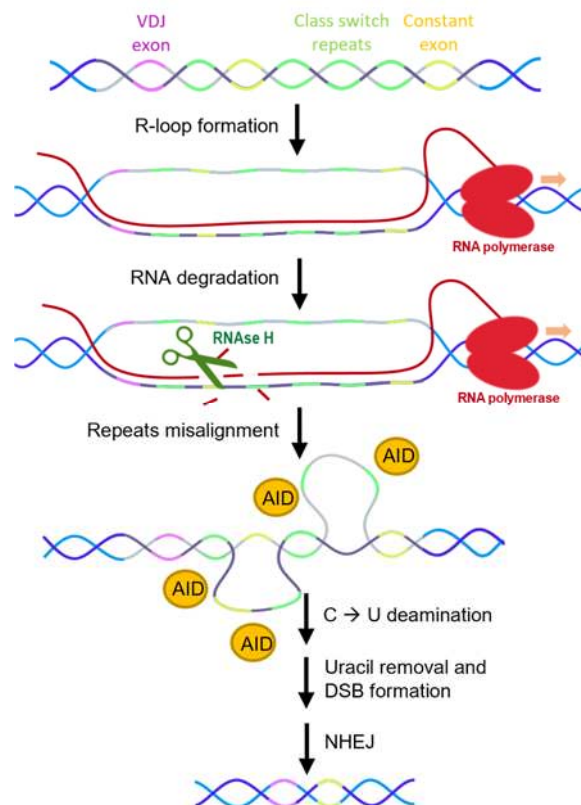


Figure 13. Schematic representation of the “collapsed R-loop” model explaining the mechanism of transcription dependent AID-mediated CSR.

The occurrence of CSR has been shown to require transcription at the Switch regions; multiple models have been proposed to explain the mechanism leading to

CSR. In the “collapsed R-loop” model (Fig.13), long stable R-loops form at the S region after generation of long noncoding RNA from cytokine-induced promoters. According to Lieber and colleagues, formation of the R-loop would be followed by its degradation by RNase H which would cause the reformation of the DNA bubble. Due to repetitive nature of the Switch region, misaligned reannealing of dsDNA would leave extruded ssDNA stretches that can be further targeted by the cytidine deaminase AID (activation-induced deaminase; Yu et al., 2003). AID binds preferentially at branched DNA structures like bubble extremities or DNA overhangs near G-quadruplexes (Qiao et al., 2017), driving Cytidine to Uracil conversion, preferentially at WGCW motifs (where W can be either A or T). Uracils can then be targeted by either the Base Excision Repair (BER) or Mismatch Repair (MMR) pathways, to generate nicks that, if close to each other and in opposite strands, lead to formation of DSBs (Petersen-Mahrt et al., 2002). Two DSBs formed in distinct S regions can then be bridged together (reviewed in Kotnis et al., 2009), deleting the exons in between and generating the new Ig isotype coding gene. Although alternative mechanisms have been proposed, e.g. post-translational R-loop formation by an RNA containing G4 quadruplexes, which would promote the recruitment of AID at the locus (Ribeiro de Almeida et al., 2018), the “collapsed R-loop” model remains the prevalent model to explain CSR up to date.

2.3.2 | R-loop and Replication initiation

R-loop formation has been shown to represent a necessary intermediate during phagic and bacterial DNA replication initiation. Notably, a possible mechanism of DNA replication in the bacteriophage T4 is the formation of a persistent R-loop, whose processing by RNase H generate the primer for lagging strand synthesis. Moreover, bacterial ColE1-type plasmids require the formation of a 550bp DNA:RNA hybrid on the leading strand, which is also processed by RNase H to generate the free 3'-end necessary for replication initiation (Aguilera and García-Muse, 2012).

Interestingly, *rnhA E. coli* mutants have also been shown to be able to replicate their genome in a DnaA and oriC- independent mechanisms, with an R-loop acting as a primer for DNA synthesis. While this is the default mechanism for E1 plasmid replication (Itoh and Tomizawa, 1980), R-loop-initiated replication cycles of the bacterial genome are a less frequent alternative mechanism that can lead to genetic instability (Drolet and Brochu, 2019). R-loop formation and R-loop dependent replication have been proposed to be relevant in stress condition to favour mutagenesis and chromosome rearrangements in non-dividing bacteria to accelerate adaptation, promoting survival (Wimberly et al., 2013). Unscheduled origin-independent replication events have also been observed in eukaryotic genomes: in yeast cells depleted of the R-loop preventing enzyme Top1 and both RNases H, persistent R-loops can prime DNA synthesis at the 35S ribosomal locus (Stuckey et al., 2015).

R-loops are also important for vertebrate mitochondria strand-asynchronous replication: formation of an R-loop at the site of replication initiation allows the

opening of the DNA to generate the replication bubble. The RNA acts as a primer for the unidirectional replication of the heavy strand and is then rapidly processed by RNase H1 (Holt, 2019).

2.3.3 | R-loops and transcription regulation

R-loop formation is also intrinsically linked to the regulation of transcription. This relationship is reciprocal and works both ways: on the one hand, high levels of transcription and factors involved in the transcription process influence the formation or stability of R-loops. On the other hand, certain R-loops have the ability to influence the transcription process by participating in initiation and termination events, leading to either positive or negative outcomes.

Transcription initiation

Formation of R-loop by lncRNAs seems to be highly frequent in plants, where, differently than other organisms, R-loop are mostly found to form in antisense orientation than the annotated CDS, at the vicinity of the promoter region (Xu et al., 2017). Several examples have been described of association of long-noncoding RNAs with promoters leading to R-loop dependent transcriptional repression by promoter occlusion. R-loop formation by the lncRNA *COOLAIR* at Flowering Locus C (*FLC*) in *Arabidopsis* is promoted and stabilized by the ssDNA binding protein AtNDX, which ultimately leads to the downregulation of *COOLAIR* contributing to the derepression of *FLC*, promoting flowering after vernalization (Sun et al., 2013). R-loop dependent gene downregulation was also observed in rice, where R-loops have been shown to naturally form at the promoters of Auxin signalling and transport genes. Tissue specific expression of Topoisomerase I prevents R-loop accumulation at these loci favouring their expression in the root tip promoting root development and gravitropism (Shafiq et al., 2017).

In vertebrate genomes, instead, R-loop formation confers protection from DNA methylation at CpG islands associated with promoters of active genes, notably housekeeping genes, preventing transcriptional silencing (Deaton and Bird, 2011). R-Loop formation on episomal templates has been shown to protect at least partially from DNMT3B1-mediated DNA methylation (Ginno et al., 2012), while reduced R-loop formation in ALS4 patient cells showed increased methylation and decreased expression of genes regulated by CpG-rich promoters (Grunseich et al., 2018). The protective action of R-loop against DNA methylation seems to be ascribable to the lower affinity of the DNA methylase enzymes for the DNA:RNA hybrids compared to dsDNA (Grunseich et al., 2018). Alternatively, R-loop formation has been shown to decrease methylation through the recruitment of the demethylase TET1 by the R-loop reader GADD45A in mouse embryonic stem cells (Arab et al., 2019).

Another way in which R-loop formation at promoter proximal region can influence transcription is by modulating the binding of trans-acting DNA binding factors, e.g. key regulators of ESC pluripotency Tip6—p400 and PRC2 (Chen et al., 2015; Skourti-Stathaki et al., 2019). Finally, artificially constructed R-loops have been shown to perform *de novo* transcription acting as promoters *in vitro*. Such activity is proposed to account for the generation of many known antisense lncRNA transcribed from

the mammalian genome, but further research will be necessary to characterise the extent of this phenomenon *in vivo* (Tan-Wong et al., 2019).

Transcription termination

Transcription termination has also been associated with R-loop formation. In HeLa cells, RNase H overexpression leads to transcription readthrough, suggesting a role of R-loops in promoting polymerase pausing downstream of poly(A) sites. Their subsequent resolution by SETX would then liberate the RNA making it the substrate of Xrn1 to complete the termination process (Skourti-Stathaki et al., 2011). Moreover, N⁶-methyladenosine (m⁶A), the most abundant reversible RNA modification, have been recently linked to R-loop formation and transcription termination. Notably, R-loops forming at termination sites in m⁶A⁺ genes were majorly affected by the absence of the methyltransferase METTL3. These observations point to a role of m⁶A (and METTL3) in promoting R-loop formation at transcription termination sites to prevent polymerase readthrough (Yang et al., 2019).

2.3.4 | R-loops in cell differentiation and development

Mapping the location and abundance of R-loops during drosophila embryogenesis or at different stages of cell differentiation of plant tissues and human pluripotent stem cells (hPSC) has brought to light interesting new aspects of the metabolism of R-loops and their potential relevance during development. By comparing the profiles of different developmental stages in drosophila embryos (Munden et al., 2022) and plant tissues (Xu et al., 2020), it was noted that both global and local levels of R-loops vary depending on the stage of differentiation. In hPSCs and their derivatives, while global hybrids levels remain similar among cell types, R-loops are differentially detectable at cell type-specific genes. These R-loops are also often associated with epigenetic signatures and in some cases are retained following reprogramming (Yan et al., 2020). In light of this, it is proposed that the controlled formation and resolution of R-loops constitutes a further level of regulation in differentiation and development, with possible repercussions in the establishment and maintenance of epigenetic memory. The mechanisms by which these functions might be exerted have yet to be determined. Interestingly, the fluctuation of R-loops levels during differentiation have been found to be independent of the expression levels of the corresponding genes, indicating that such fluctuations are not merely the result of changes in the transcription programme, but could require the modulation of factors that actively favour or remove R-loops at designated positions at the appropriate time (Xu et al., 2020; Munden et al., 2022).

2.4 | Factors involved in R-loop control

Despite the numerous biological processes in which R-loops have been shown to be actively and positively involved, their presence in the cell is far from harmless. Excessive accumulation and unscheduled formation of R-loops can lead to significant disruptions in the transcription and replication programs, as well as threatening the survival of the cell. To this end, cells are equipped with multiple mechanisms to deal with the formation of these structures, which I will detail below.

2.4.1 | *Sed quis custodiet ipsos custodes?*²

How cells are able to discriminate between regulatory, neutral or genotoxic R-loops is not entirely clear. On the one hand, the formation of R-loops must be allowed, if not encouraged, and the fulfilment of their function may even require the programmed formation of DNA damages. On the other hand, hybrid formation outside these controlled contexts must be immediately detected and suppressed from the outset, to minimise their impact on cell homeostasis.

Therefore, strict regulation is necessary to finely regulate the timing, amount and persistence of physiological R-loop formation in order to avoid abnormal accumulation or unscheduled formation at ectopic sites, or collateral activity of the factors promoting or processing R-loop in these contexts. Given the wide variety of factors that have been identified as being involved in R-loop formation or resolution, it is plausible to hypothesize that different factors could be involved in the regulation of specific subpopulations of R-loops. The attributes by which each one can mark its “territory”, however, are still largely unknown.

R-loop associated chromatin marks

A possible strategy to flag genotoxic R-loops in eukaryotes has been proposed to rely on the chromatin landscape surrounding the hybrid. Indeed, specific chromatin marks, i.e. the histone modification H3S10 phosphorylation has been shown to correlate with R-loop formation in both yeast, *C. elegans* and human cells (Castellano-Pozo et al., 2013) and to be required for R-loop dependent genetic instability in yeast (García-Pichardo et al., 2017). H3S10P has been associated with transcription and is compatible with active chromatin modifications. This is consistent with the observed tendency of R-loops to generally associate with an open, H3K4/H3K36-methylated chromatin state (Sanz et al., 2016). Moreover, DNA:RNA hybrids do not wrap around nucleosomes (Dunn and Griffith, 1980), thus likely defining a region of more open chromatin behind the running polymerase. Conversely, R-loop formation has also been linked to increased deposition of H3K9me2/me3 and heterochromatin formation; given the transient H3K9me2/me3 deposition observed at sites of DSBs repair, such association with silenced chromatin may reflect R-loop induced fragility (Chédin, 2016). The mechanism with which chromatin modification are implicated in the modulation of R-loop genotoxicity is still a matter of investigation.

² Latin for “but who will guard the guards themselves?” (Juvenal, *Satire*, VI, O31-O32)

2.4.2 | R-loop prevention

When dealing with such unpredictable and dangerous structures, prevention is likely better than cure. Indeed, several strategies are in place to minimize the formation of R-loops, from the tight regulation of transcription dynamics and mRNP packaging, to the prevention of torsional stress that can arise during transcription (Fig. 14; El Hage et al., 2010). A major strategy to prevent R-loop formation is disfavoring the RNA ability to anneal to the template DNA by its co-transcriptional coating with mRNA associated factors, to sterically prevent hybridization and favour its eviction from the transcription site.

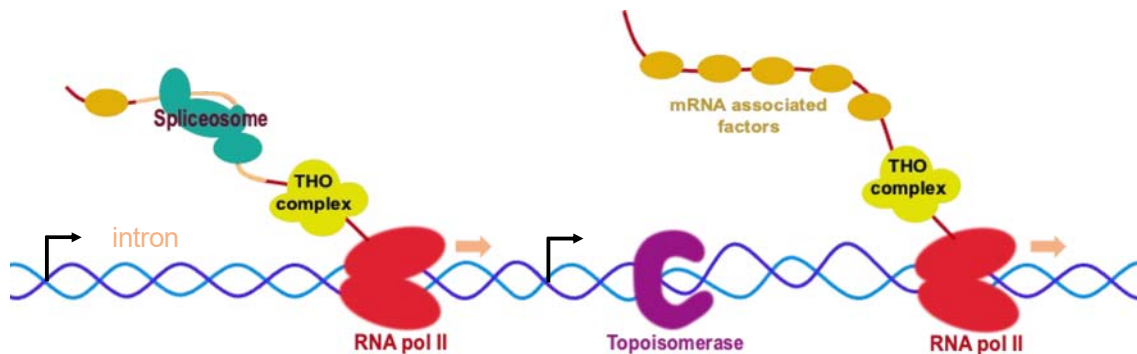


Figure 14. **Factors implicated in the prevention of R-loop formation.** Coating of the nascent transcript by mRNA associated factors, i.e. the THO complex, and the assembly of the spliceosome at intron-containing genes sterically prevent the RNA to reanneal to the template DNA. Topoisomerases are essential in relaxing the negative supercoiling forming upstream of the transcribing polymerase, avoiding RNA invasion of the looser alpha-helix.

THO/TREX The THO complex (Tho2, Hpr1, Mft1, Thp2 in yeast; THOC1,-2,-5,-6,-7 in metazoans) is a conserved eukaryotic factor which together with other components of the mRNA export machinery (Tex1/TEX1, Sub2/UAP56-DDX39 and Yra1/ALY) form the TREX (TRanscription and EXport) complex (Strässer et al., 2002, Katahira 2012). Together, these factors are implicated in the formation of export-competent mRNPs to allow polIII transcripts to reach the cytoplasm to be translated. The THO complex was initially identified in yeast through a screen looking for hyper-recombinant mutants, where the *hpr1Δ* mutant stood out for the dependency on transcription of its genetic instability phenotype (Chávez and Aguilera, 1997; Prado et al., 1997). Characterization of a similar phenotype followed shortly after for the other subunits (Piruat and Aguilera, 1998; Chávez et al., 2000), and have been later found to be shared also by mutants for the other components of the TREX complex Yra1 and Sub2, mRNA export factors Mex67 and Mtr2, and the TREX2 subunits Thp1 and Sac3 (Jimeno et al., 2002; Gallardo and Aguilera, 2001; Gallardo et al., 2003). Interestingly, *tho* dependent transcription associated recombination (TAR) impacts only a subset of genes (Chávez et al., 2000), notably long, GC rich loci (Chávez et al., 2001). Moreover, the transcription impairment observed in *hpr1Δ* mutants can be suppressed by self-cleavage of the RNA by a hammerhead ribozyme (Huertas and Aguilera, 2003), further supporting that the passage of the polymerase *per se* is not sufficient to trigger instability. This observation highlights the role of the RNA as a source of mitotic recombination during transcription elongation and the protective function of RNA associated factors against fragility. Formal demonstration of R-loop

formation in *tho* mutants was achieved by means of recombination assays, which show the sensitivity of *hpr1Δ* hyperrecombination to both transcript cleavage and RNase H overexpression, and dotblot analysis of DNA intermediates, which showed an increase in hybrid formation in *hpr1Δ* mutants at the 3' end of the probed locus (Huertas and Aguilera, 2003).

In light of this, the THO/TREX complex has been proposed as a major protector from R-loop formation by ensuring the coating of the mRNA, promoting the correct assembly and export of the mRNP, thus sterically preventing the RNA to anneal to the template DNA. However, the THO/TREX complex could also prevent R-loop formation by recruiting dedicated interactors, such as the Sub2/UAP56 R-loop unwinding helicase (Pérez-Calero et al., 2020), or the histone deacetylase Sin3A (Salas-Armenteros et al., 2017). Consistently, human THOC1 mutants showed histone hyper-acetylation correlating with R-loop accumulation and R-loop dependent genetic instability, highlighting the importance of physical and functional connections between RNA associated proteins and chromatin modifier in coping with R-loop formation (Salas-Armenteros et al., 2017).

Several other RNA associated factors have been suggested to be involved in R-loop metabolism; subunits of the mRNA cleavage and polyadenylation subunits have been identified in genetic screens assessing Rad52-foci accumulation among mutants with elevated chromosome instability, whose phenotype appeared to be transcription dependent (Stirling et al., 2012). Moreover, R-loop dependent genetic instability has been characterised for the mRNA related factors Trf4, a noncanonical polyA-polymerase involved in RNA surveillance, the catalytic subunit of the nuclear exosome Rrp6, and the 3' end processing factors Rna14 and Rna15 (Luna et al., 2005; Gavaldá et al., 2013).

Spliceosome

Similarly to the THO complex, a protective function through RNA coating has also been proposed to be exerted by the spliceosome, prompted by the observation that depletion of the SR splicing factor ASF/SF2 display increase RNase H sensitive genomic instability and R-loop formation in the chicken B cell line DT40 (Li and Manley, 2005). In yeast, a small proportion of genes contains an intron, which is predominantly positioned at the 5' of the coding sequence, just few bp from the start codon, and the spliceosome is assembled co-transcriptionally. Interestingly, highly transcribed intronless genes appear more prone to form R-loops than their equivalent intron-containing counterparts, pointing to the protective effect of splicing against R-loop formation. Strikingly, insertion of an artificial intron in a *bona fide* R-loop forming reporter gene leads to a decrease of R-loop levels and R-loop dependent genetic instability. The assembly of the spliceosome was sufficient to prevent R-loop formation independently of the occurrence of the splicing reaction, confirming the importance of RNA coating in preventing DNA:RNA hybrid formation (Bonnet et al., 2017).

Consistently, bisulfite mapping of R-loop in human cells highlighted a correlation of R-loop size and position with the first exon of intron-containing genes, identifying the 3' R-loop boundaries in correspondence of the first exon-intron junction (Dumelie and Jaffrey, 2017). However, this observation was not reproduced by

single molecule R-loop footprinting studies, which mapped R-loop at intronic regions or spanning exon-intron junctions (Malig et al., 2020)

The fact that both the THO/TREX complex and components of the spliceosome travel with the transcription machinery (Meinel et al., 2013; Saldi et al., 2016) and are transferred to the nascent transcript as soon as it exits the polymerase is a strong indicator of the importance of RNA coating in preventing R-loop formation, especially at the 5' end of the transcript, which may be the portion most likely to invade the double helix (Belotserkovskii and Hanawalt, 2022).

Topoisomerases

As already hinted in the previous paragraphs, Topoisomerases exert an important role in preventing R-loop formation by relieving torsional stress that physiologically form around the RNA polymerase during transcription. Type I and II topoisomerases are responsible to relieve negative supercoiling by operating a temporary single or double strand break respectively, allowing the DNA to freely rotate favouring the relaxation of the torsional tension (reviewed in Mackay et al., 2020). The absence of Topoisomerase I leads to the accumulation of negative supercoiling upstream of the RNA polymerase, which could facilitate strand invasion and the annealing of the RNA to the loosely wounded DNA template. Such phenotype of topoisomerase mutants is conserved in bacteria (Drolet et al., 1995), *S. cerevisiae* (El Hage et al., 2010) and human cells (Tuduri et al., 2009). Topoisomerase 2 has also been shown to participate in R-loop prevention. In yeast, Top2 seems to act in a redundant manner with Top1, while in multicellular organisms TOP2 role on R-loop prevention has not been directly investigated, although it has been found to be physically present at sites of R-loop formation in proteomic screens (Patel et al., 2022). Finally, another member of the subfamily, TOP3B, has been shown to also prevent R-loop formation by reducing negative supercoiling. Notably, TOP3B recruitment at mouse and human loci appear to require arginine-methylated histones (Yang et al., 2014), thus highlighting one more time the intertwined connections between the chromatin landscape and R-loop fate.

2.4.3 | R-loop Removal

Once the preventing mechanisms fail to impede R-loop formation, or once the physiological R-loops have fulfilled their role, several factors can intervene for the removal of the hybrid, either by the degradation of the RNA moiety or its unwinding from the DNA by helicases. In budding yeast, the main actors in R-loop removal are the ribonucleases RNase H1 and H2 and the helicase Sen1 (Fig. 15).

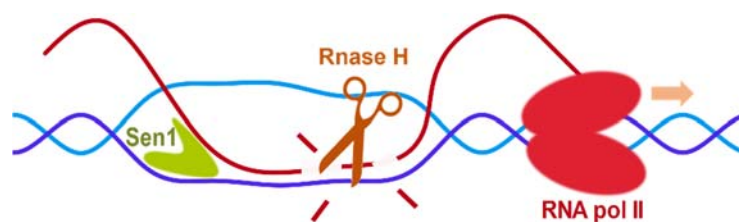


Figure 15. Schematic representation of the main factors involved in R-loop removal.

RNases H RNase H are endoribonucleases that recognise DNA:RNA hybrids and specifically cleave the RNA moiety. RNase H enzymes have been classified in two main groups: Type 1 (prokaryotic RNase HI, eukaryotic RNase H1, and viral RNase H) and Type 2 (prokaryotic RNase HII and HIII, and eukaryotic RNase H2).

RNase H1 The typical organisation of eukaryotic RNase H1 consists in a N-terminal hybrid binding domain (HBD) and a C-terminal catalytic H-domain connected by a poorly conserved linker domain that provides flexibility to facilitate binding to the substrate (reviewed in Cerritelli and Crouch, 2009).

The HBD binds preferentially to DNA:RNA hybrid, and with lower affinity to dsRNA. Protein-nucleic acid binding interaction occurs along the minor groove of the double helix, and requires 4 specific residues (W43, K59, K60 and Y29 in human). This domain, which is absent in prokaryotic enzymes and present in two copies in yeast, is important to increase the processivity of the enzyme (Nowotny et al., 2008).

The catalytic domain is responsible for the cleavage of the RNA moiety of the hybrid. Hybrid cleavage requires the interaction of the enzyme with at least 4 ribonucleotides, which makes it compatible with the resolution of R-loops but also the processing of shorter hybrids, such as Okazaki primers. Notably, despite its ability to bind dsRNA, the homoduplex cannot be cleaved since the cleavage require a specific distortion that is possible only for the hybrid double helix (Cerritelli and Crouch, 2009).

Mammalian RNase H1 is predominantly localised in the nucleus but a mitochondrial isoform also exists, which has proven to be essential for mouse embryonic development. The mitochondrial targeting signal is encoded in between two in frame AUG initiation codons, and the expression of one or the other isoform is due to leaking scanning translation initiation (Cerritelli et al., 2003). *S. cerevisiae* RNase H1 lacks mitochondrial targeting signal: however, an increase in R-loop in mitochondrial DNA was observed in *rnh1Δ* yeast mutants, suggesting a possible function of the enzyme in this organelle (El Hage et al., 2014).

RNase H2 RNase H2 exist as a trimer (RNase H 201, 202, 203 in yeast, RNase H 2A, 2B, 2C in mammals), consisting of a catalytic subunit (RNase H 201/A) and 2 accessory subunits. RNase H2 has the unique ability of recognizing and cleaving single ribonucleotides misincorporated into the DNA duplex through the Ribonucleotide Excision Repair (RER) pathway (Rydberg and Game, 2002), on top of the removal of longer hybrids. It has also been shown to participate in Okazaki primers processing, although it requires additional factors to complete their removal since the cleavage leave one residual ribonucleotide (Qiu et al., 1999). Notably, the increased levels and activity of RNase H2 observed in *rnh1Δ* mutants (Arudchandran et al., 2000), together with the more elevated levels of genetic instability scored in the double *rnh1Δrnh201Δ* mutant compared to each of the single deletants (O'Connell et al., 2015) imply the redundancy of the function of these proteins in the regulation of R-loop levels. Interestingly, RNase H2 has been shown to take on a more prominent role than RNase H1 in DNA:RNA hybrid removal and maintenance of genome stability (Arudchandran et al., 2000; O'Connell et al., 2015): a recent study on the

cell cycle regulation of RNase H1 and 2 expression suggested a more prominent housekeeping role for RNase H2 particularly in G2, while RNase H1, whose expression is independent from the cell cycle, seems to behave more like a stress-response factor (Lockhart et al., 2019).

While the deletion of both RNases H is viable in yeast (Arudchandran et al., 2000), deletion of any of the two enzymes is embryonic lethal in mouse cells, due to mitochondrial abnormalities, while their depletion at later developmental stages leads to accumulation of misincorporated ribonucleotide and R-loop accumulation similarly of what is observed in *rnh1Δrnh201Δ* yeast mutants (reviewed in Cerritelli and Crouch, 2019).

RNases H, in particular H1, are widely used as controls in R-loop detection strategies to verify that the retrieved signal corresponds to *bona fide* DNA:RNA hybrids. However, the extent of their involvement in R-loop regulation *in vivo* is still matter of study. Notably, RNases H could be implicated in the regulation of only a subset of R-loops, an hypothesis supported by the detection of RNase H1-resistant hybrids (Crossley et al., 2020). Moreover, given the pervasiveness of R-loop formation in the cell, R-loop removal through RNA degradation may be a costly strategy, particularly in mammalian cells where transcripts are several kb long.

Senataxin Senataxin (Sen1/hsSETX), is a 5'-3' helicase involved in transcription termination: as a subunit of the essential NNS complex (Nrd1, Nab3, Sen1), it promotes eviction of RNA polymerase II from the DNA. Sen1/hsSETX has long been considered one of the main factors involved in R-loop resolution. Yeast *sen1-1* loss-of-function mutant displays R-loop accumulation, transcription-dependent hyper-recombination, and is synthetic lethal with several DNA repair genes (Mischo et al., 2011). Analogous phenotypes have been recently recapitulated using an auxin inducible degron (San Martin-Alonso et al., 2021). Similarly, depletion of human SETX leads to R-loop formation at transcription termination sites (Skourti-Stathaki et al., 2011) and SETX has been shown to associate with replication forks, promoting their progression across actively transcribed genes (Alzu et al., 2012). Interestingly, a recently characterised separation-of-function mutant, *sen1-3*, which loses the interaction with the replisome but still performs efficient termination, do not display R-loop accumulation but is synthetic lethal with *rnh1Δrnh201Δ* and *hpr1Δ* R-loop accumulating mutants (Appanah et al., 2020). Moreover, investigations on the cell cycle dependency of R-loop genotoxicity highlighted how Sen1 depletion causes R-loop accumulation exclusively in S-phase, suggesting a major role for the helicase in R-loop resolution at transcription replication conflicts (San Martin-Alonso et al., 2021). Sen1 activity in limiting R-loop dependent genetic instability does not seem to rely on its hybrid unwinding activity, but rather on its role in evicting RNA polymerases both at transcription replication conflicts and at termination sites (Aiello et al., 2022).

Fanconi Anaemia Pathway Other important players in the recognition and removal of R-loops, albeit partially indirectly, are factors of the Fanconi Anaemia pathway. The Fanconi Anaemia (FA) pathway, or BRCA pathway, is composed of at least 19 proteins, and exerts a crucial role in the repair of DNA inter-strand crosslinks (ICLs), whose formation can stall

the replication and transcription processes. Notably, some of these factors also play additional roles beyond the canonical recognition and repair of ICLs, thus creating a broad network interconnecting the different DNA repair pathways to safeguard genome integrity (Ceccaldi et al., 2016). In the context of DNA:RNA hybrids, the FA proteins BRCA1, BRCA2 and FANCD2 have been shown to be recruited at actively transcribed, R-loop prone regions, and to have a protective role against R-loop-dependent damage (Bhatia et al., 2014; García-Rubio et al., 2015; Hatchi et al., 2015; Schwab et al., 2015). Consistently, FANCI-FANCD2 has been shown to interact with R-loops *in vitro* (Liang et al., 2019), while FANCM and its yeast ortholog Mph1 display branchpoint translocase activity, unwinding R-loops (Hodson et al., 2018). BRCA1 and BRCA2 functions at R-loop formation sites are instead exerted through their interactions with ulterior R-loop resolving enzymes, i.e. RNase H2 (D'Alessandro et al., 2018), and the unwinding helicases SETX (Hatchi et al., 2015) and DDX5 (Sessa et al., 2021). Moreover, the human helicase BLM has also shown to interact with several FA factors, and their collaboration is important in mitigating R-loop genotoxicity (Chang et al., 2017 and references therein). The presence of these interconnections demonstrates how numerous proteins possessing the ability to recognise or process lesions can act in concert to signal potential genotoxic situations and direct the most suitable factors to resolve these insults.

2.4.4 | R-loop sensing

The majority of the interactions analysed so far have in common the use of the DNA:RNA hybrid as the feature recognised and targeted by the proteins involved in the resolution of R-loops. However, another specific component of the R-loop structure, that distinguishes them from other types of DNA:RNA hybrids, is the displaced single strand DNA filament. Binding of proteins to the ssDNA can have an important role in the sensing of R-loop presence and in mediating the recruitment of R-loop resolving proteins, but also in the protection of such vulnerable structure waiting to have the conditions to reform the double helix. Moreover, binding of ssDNA proteins to R-loops have been shown to stabilize them, as is the case for AtNDX action at the Arabidopsis FLC locus (Sun et al., 2013). Also in plants, the ssDNA binding AtALBA2, in concert with the DNA:RNA hybrid binding protein AtALBA1, has been shown to localize at a subset of Arabidopsis R-loop forming loci, where it prevents damage formation. Deletion of the two proteins do not influence R-loop levels, indicating that the dimer act as an R-loop reader, preserving genome stability with a mechanism that has yet to be deciphered (Yuan et al., 2019).

RPA Replication protein A is a widely conserved protein complex constituted of three subunits of 70, 34 and 14 kDa, respectively called Rfa1-2-3 in yeast and RPA1-2-3 in mammals. The complex contains six DNA binding domains, four of which located in the major subunit Rfa1, which allow the binding to ssDNA in a sequence independent manner. Remarkably, long considered to be a complex that binds primarily single-stranded DNA, RPA was recently re-evaluated for its binding affinity for RNA, which turns out to be only an order of magnitude lower than that of ssDNA (Mazina et al., 2020). RPA is implicated in most of the fundamental processes

concerning genome homeostasis: DNA replication, checkpoint signalling, and DSBs repair (Dueva and Iliakis, 2020). Strikingly, RPA also physically interacts with RNA polII and has been shown to associate with highly-transcribed genes in ChIP assays, suggesting an additional role of RPA in transcription elongation (Sikorski et al., 2011), and raising the hypothesis of the involvement of this multipurpose protein in preventing transcription dependent genetic instability. RPA colocalization with R-loops have been observed by proximity ligation assay in HeLa cells; the signal was sensitive to R-loop modulation through knockdown of the helicase AQR or RNase H overexpression. In the same study, RPA was shown to physically interact with RNase H1, and to stimulate its hybrid degradation activity *in vitro*. It is noteworthy that human RPA specifically binds human RNase H1, while no interaction has been shown with *E. coli* RNase H1, and vice versa, indicating that this function may be species specific (Nguyen et al., 2017). RPA is also involved in the targeting of AID at Ig Switch regions (Chaudhuri et al., 2004), and can interact with and stimulate the activity of the BLM helicase (Brosh et al., 2000), further supporting its role in R-loop sensing and recruitment of R-loop resolving factors.

However, an opposite role on R-loop metabolism has also been proposed: RPA displays *in vitro* hybrid formation activity with a forward-strand exchange mechanism (Mazina et al., 2020). This, together with the re-evaluated affinity of RPA for RNA, prompted the hypothesis that RPA could rather be involved in R-loop formation or extension rather than mediating their resolution. Moreover, ssDNA coating by RPA, if not timely removed, may stabilise the hybrid, and prevent the reannealing of the DNA, even after RNA eviction.

Further studies will be necessary to better elucidate the contribution of such an ubiquitous protein to R-loop metabolism, which has the potential to play an important role in the sensing of the three-stranded molecules and in coordinating their take over by the most appropriate resolution mechanism.

2.4.5 | The expanding R-loop interactome

Although many hybrids-associated factors have been identified and their contribution to R-loop metabolism thoroughly investigated, the picture is still not complete. In addition to the most characterized factors already mentioned, over the years a considerable and ever-increasing number of proteins has been added to the list, which is, however, still not exhaustive. Several helicases, for example, have been associated with R-loop resolution, i.e. the RECQ-like helicases BLM and WRN (Sgs1 in yeast - Chang et al., 2017; Marabitti et al., 2019), DDX19 (Hodroj et al., 2017), DDX21 (Song et al., 2017), FANCM (Mph1 in yeast - Schwab et al., 2015; Lafuente-Barquero et al., 2017), UAP56/DDX39B (yeast Sub2 - (Pérez-Calero et al., 2020), and many others (reviewed in Petermann et al., 2022).

R-loop proteome studies

In recent years several groups have tried to systematically investigate the R-loop interactome in order to identify novel, unknown factors involved in R-loop metabolism. Several strategies have been applied, such as hybrid pull down coupled to mass spectrometry in human (Cristini et al., 2018; Wang et al., 2018) or mouse cells (Wu et al., 2021), or proximity biotinylation assays using catalytic dead

RNase H1 or its hybrid binding domain as bait in human cells (Mosler et al., 2021; Yan et al., 2022).

The implementation of such methods has proved useful in expanding the knowledge about the chromatin landscape surrounding R-loops and allowed the identification of novel R-loop binding proteins, notably helicases. However, the most known R-loop regulator, i.e. RNase H1 and Sen1/SETX do not appear as *bona fide* R-loop interactants in two out of the three pull down assays (Cristini et al., 2018; Wu et al., 2021), while other R-loop preventing factors, i.e. subunits of the spliceosome machinery, are identified by all the R-loop proteome investigations. Moreover, an integrative analysis of such available datasets shows a very poor overlap between the different studies (Kumar et al., 2022). Although the weak consensus among studies could be explained by the differences in model organisms and strategies applied, such variability, together with the greatly elevated number of hits identified by each screen, raises some perplexity about the specificity and stringency of these methods.

Indirect effects of R-loop resolving proteins

The ever-growing number of factors identified up to now as putative regulators of R-loop metabolism raises the question whether their implication in R-loop modulation is direct and carried out *in vivo*. Notably, *in vitro* studies on the ability of a protein to act on DNA:RNA hybrids have sometimes contrasted with *in vivo* findings. SETX, AQR and WRN, for example, have been proposed to resolve R-loops *in vivo*, i.e. their inactivation leading to R-loop accumulation (El Hage et al., 2014; Sollier et al., 2014; Marabitti et al., 2019), but little is known about their substrate specificity or regarding the dynamics of their activity (Chakraborty, 2020). Conversely, DNA:RNA hybrid formation or unwinding activities established *in vitro* have not always been corroborated by *in vivo* proofs (yet), as is the case for the RNA-DNA strand exchange activity of the *Drosophila* polycomb complex PRC2 (Alecki et al., 2020).

Furthermore, proteins can have multiple abilities which could result in different outcomes facing the same substrate. The DHX9 helicase, for example, was identified by R-loop proteomic pull-down as a major R-loop interactant, and was shown to prevent R-loop accumulation upon CPT treatment and R-loop dependent damage, possibly through its role in transcription termination (Cristini et al., 2018). Consistently, DHX9 can unwind DNA:RNA hybrids *in vitro* (Chakraborty and Grosse, 2011). However, another study attributed to DHX9 an R-loop promoting role *in vivo* through its ability in removing RNA secondary structures which would otherwise prevent hybrid formation (Chakraborty et al., 2018). Finally, as our knowledge of yeast Sen1's *modus operandi* in yeast deepens, it seems less and less plausible that it would act by directly unwinding hybrids through its helicase domain, but rather exerts its protective function through its interactions with the DNA and RNA polymerases (Aiello et al., 2022).

Particular attention should therefore be paid to determine the direct or indirect nature of the actions on R-loop by the proteins already known and those that will be identified by the ongoing and future investigations of the R-loop proteome.

2.5 | R-loop as a threat for genome stability

R-loops form pervasively in the genome, and despite having regulatory roles, they can constitute a dangerous threat for genome homeostasis. While R-loop processing can happen in a regulated manner in the context of specific physiological properties, their unscheduled accumulation or the deregulation of their processing enzymes can lead to DNA damage with a negative impact on the cell homeostasis. Formation of R-loops can become therefore a major cause of genetic instability, both linked to specific structural weaknesses of these structures and to the genomic context at which they are formed or stabilised.

2.5.1 | ssDNA modifying enzymes and nucleases

The exposed ssDNA, a characteristic feature of the R-loop structure, can have an increased susceptibility to both mechanical solicitations and enzymatic modification that could lead to DNA breakage (Fig. 16). The activation-induced deaminase AID, for example, required for CSR at Immunoglobulin loci, deaminates deoxycytidine residues to deoxyuridine, acting at ssDNA in a sequence and transcription dependent manner. Such lesions are then processed through Uracil excision by the uracil-DNA glycosylase UNG1, which generates an abasic site (Di Noia et al., 2007) that can be processed into a single strand break by the Base Excision repair (BER) pathway. Indeed, yeast reporters carrying murine Ig S regions or the human *c-MYC* oncogene display increased occurrence of AID-dependent DSBs and subsequent chromosomal rearrangements, which were exacerbated in the R-loop accumulating THO null mutants (Ruiz et al., 2011). Similarly, the yeast deaminase Fcy1 has been shown to act on expanded CAG/CTG tracts leading to the instability of repeats engaged in a co-transcriptional R-loop. In addition, the nuclease activity of the mismatch repair (MMR) protein MutLγ (Mlh3 in yeast) was also shown to contribute to CAG repeat fragility and contractions likely by nicking R-loop substrates (Su and Freudenreich, 2017). Finally, following the formation of R-loop 3D structures or supercoiling, stalled topoisomerase I or G4 specific nucleases could lead to the formation of potentially genotoxic DNA lesions (Hamperl and Cimprich, 2014). In all these instances, SSBs formed at the site of R-loop formation, if left unrepaired, become hazardous for DNA integrity due to their subsequent conversion in DSBs by either ulterior processing events, or at the passage of the replisome, causing chromosomal rearrangements and compromising the replication program.

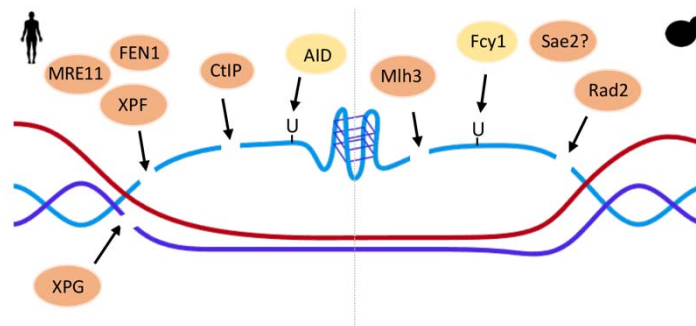


Figure 16. Schematic representation of the factors involved in R-loop processing into DSBs.

2.5.2 | Structure-specific nucleases

The dsDNA-ssDNA junctions present at the extremities of the R-loop bubble are also R-loop specific structures vulnerable to nucleolytic attack. The nucleases of the Transcription Coupled – Nucleotide Excision Repair pathway (TC-NER), XPF and XPG (Rad1 and Rad2 in yeast), have been shown to recognise and cleave the flaps at these junctions both *in vitro* (Tian and Alt, 2000) and *in vivo*, causing R-loop dependent genetic instability in human cells upon AQR knockdown (Sollier et al., 2014). The nicks and gaps generated by the concomitant cleavage of the two nucleases at both sides of the hybrid would result in DSBs formation and fork collapse at the subsequent passage of the replisome. In addition, XPG/XPF together with the Flap endonuclease FEN1 (Rad27 in yeast) and the 5' endo- 3' exonuclease MRE11 have been shown to generate DNA breaks at R-loops forming at transcription blocking DNA lesions at stalled TOP1 cleavage complexes (Cristini et al., 2019). Finally, ssDNA break formation in mammals have been attributed also to the CtIP endonuclease, a function possibly conserved by its yeast orthologue Sae2 (Makharashvili et al., 2018).

2.5.3 | Transcription replication conflicts

R-loop genotoxicity is not limited to the physical damage endured by the nucleic acid structure itself, but also relies on disturbances caused to processes occurring at the vicinity of their site of formation, and to the transcription process itself. R-loop dependent genetic instability arising in *E. coli* (Gan et al., 2011; Lang et al., 2017), yeast (Gómez-González et al., 2011; San Martin-Alonso et al., 2021) and human cells (Gan et al., 2011; Hamperl et al., 2017) has been shown to require active replication. Moreover, yeast *tho* mutants display an increase in transcription replication conflicts as scored by the increased occupancy of the helicase Rrm3 at active loci during S-phase (Gómez-González et al., 2011).

The severity of the consequences of transcription-replication conflicts depends on the orientation of the polymerases at the moment of the collision (reviewed in Petermann et al., 2022). If the two polymerases are travelling in the same direction, a codirectional (CD) collision has mild if any consequences. In this context, the replicative DNA helicase, CMG, has been shown *in vitro* to be able to unwind or translocate DNA:RNA hybrids, allowing fork progression (Kumar et al., 2021). Head-on (HO) collisions between polymerases travelling in opposite direction on the same DNA strand, instead, have a much more severe impact, leading to transcription and fork stalling. Replication blockage can lead to fork reversal and subsequent checkpoint activation, and can result in DNA breakage and genetic instability; different checkpoint pathways are activated depending on the nature of the collisions: in human cells, HO and CD conflicts induce distinct DDRs, activating either ATR or ATM respectively (Hamperl et al., 2017). R-loop formation can trigger transcription-replication conflicts in several ways: RNA polymerase pausing or accumulation of positive supercoiling can interfere with replication fork progression (Promonet et al., 2020), or trigger formation of 3D structures at the

ssDNA, e.g. G-quadruplexes in regions encompassing G clusters, or hairpins in triplet repeat sequences. Chromatin compaction around sites of R-loops formation has also been proposed to contribute to transcription-replication interference (Castellano-Pozo et al., 2013). *In vitro* reconstruction of eukaryotic R-loop-replisome collisions revealed how formation of R-loops on the leading strand in both codirectional or head-on orientation has the potential to perturb replication progression leading to fork uncoupling. Strikingly, while removal of the hybrid by RNase H is sufficient to rescue the replication efficiency in CD collisions, the same is not observed for HO conflicts: replication stalling persists even after degradation of the hybrids, likely due to secondary structures forming at the ssDNA, i.e. G-quadruplexes, that constitute a further physical obstacle for proper restart (Kumar et al., 2021).

In line with the hazardous nature of HO collisions, bacterial genes transcribed in head-on orientation with replication display higher mutagenesis rate (Paul et al., 2013; Merrikh and Merrikh, 2018). Interestingly, certain bacterial genomes are characterised by a gene organisation bias that privileges the codirectional orientation for essential genes. This, on top of favouring chromosome segregation by preventing conflicts between transcription and the translocating bacterial condensin complexes (Gruber, 2018), also minimises the occurrence of dangerous and potentially lethal transcription-replication conflicts (Rocha and Danchin, 2003). Interestingly, transcription replication conflicts, in particular in HO orientation, have also been shown to cause further accumulation of R-loops, (Hamperl et al., 2017; Lang et al., 2017), possibly due to the negative supercoiling building up behind the two stalled polymerases, favouring RNA hybridisation. The fact that R-loop formation could be both a cause and a consequence of transcription-replication conflicts adds a further level of complexity to the study of these phenomena. The development of approaches determining the chronological order of these events would be crucial for the determination of causality.

2.5.4 | R-loops in DNA repair

As mentioned at the very beginning of this chapter, DNA:RNA hybrid formation has been shown by an ever-increasing amount of studies to be connected to DNA damage and DNA repair processes. However, many aspects of this relationship remain to be elucidated, notably whether R-loops and/or DNA:RNA hybrids formation happens prior or post DSB formation, and if the presence of RNA at DSBs sites has positive or negative effects on the efficiency of DNA repair.

Possible models for DNA:RNA hybrid or R-loop formation at DSBs

DNA:RNA hybrids formation can be detected at DSBs generated by reactive oxygen species, laser irradiation or site-specific endonucleases in several organisms, from yeast to metazoans (Palancade and Rothstein, 2021). Different models have been proposed to explain R-loop formation at DSBs sites. The decrease of repair efficiency following transcription inhibition or R-loop removal suggested that R-loops could be required intermediates of the repair process. It was therefore proposed that *de novo* transcription could occur at damage sites, possibly arising from pre-initiation complexes assembled at DSB ends in a process dependent on MRN, resulting in bi-directional transcription (Pessina et al., 2019; Sharma et al.,

2021). However, this model is not coherent with the observed deposition of repressive chromatin marks at DSBs (Ayrapetov et al., 2014), and the predominant detection of hybrids at breaks generated at sites of ongoing transcription (Cohen et al., 2018; Bader and Bushell, 2020; Teng et al., 2018). In line with this, an alternative model supports the formation of R-loop as a consequence of DSBs-induced transcriptional repression. Alternatively, reannealing of pre-existing mRNAs may be favoured by DSBs-induced chromatin remodelling, extensive resection, or polymerase backtracking (reviewed in Marnef and Legube, 2021).

Physiological relevance of DSB DNA:RNA hybrids

The physiological relevance of DNA:RNA hybrids at DSBs is also still a matter of debate. Notably, both positive and negative outcomes have been detected in this context, painting a rather blurry picture of this phenomenon.

Modulation of R-loop levels has been shown to have an impact on repair pathway choice, e.g. favouring NHEJ over HR by counteracting the formation of Rad51 filaments. Moreover, formation of hybrids could have an impact on resection by interfering with the recognition of DNA ends, thus altering the dynamics of the process (reviewed in Palancade and Rothstein, 2021). Furthermore, R-loops have been proposed to act as a platform to recruit damage related factors, i.e. BRCA1/2 (D'Alessandro et al., 2018) and Rad52 (Teng et al., 2018; Yasuhara et al., 2018), thus promoting faster repair.

Further research is necessary to better elucidate the contribution of R-loops to DNA damage pathways. In addition to the development of strategies allowing to better determine the causality and chronology of the observed events, particular attention should also be put in determining the three-stranded or double stranded nature of the hybrids identified.

2.5.5. | Pathological R-loops

We have just seen various possible mechanisms of R-loop dependent genome instability and listed the high number of factors playing a pivotal role in governing R-loop existence and containing the aftermaths of their unscheduled accumulation.

In light of this, it is easy to deduce how formation of R-loops or deregulation of R-loop associated factors can be associated with pathological situations. However, in all these situations, it is not clearly established whether R-loop deregulation is the primary cause or a consequential symptom of the pathology.

R-loop and cancer

There are many mechanisms with which oncogenic events can increase R-loop levels and consequently genetic instability: oncogene induced increase in transcription levels can promote R-loop formation, as is the case for the EWS-FLI1 fusion in the paediatric cancer Ewing sarcoma (Gorthi et al., 2018). Mutations affecting R-loop preventing factors, many of which act as tumour suppressor genes, can also lead to unscheduled R-loop formation. Expression of mRNA export factors THOC1 and ALY have been found to be differentially perturbed in cancer cells: in particular, THOC1 mRNA and protein levels are up-regulated in ovarian and lung tumours but down-regulated in those of testis and skin (Domínguez-Sánchez et al., 2011). In a few cases, oncogenes have been shown to promote R-loop suppression, thus enhancing cancer cells resilience to transcription-related stresses, allowing

them to tolerate high levels of instability while proliferating, increasing cancer invasiveness (reviewed in Petermann et al., 2022). Finally, R-loop dependent chromosomal rearrangements could by itself lead to oncogenesis: AID mediated translocations between the immunoglobulin locus and c-MYC, which are enhanced in absence of the THO complex, have been found to be associated with Burkitt's lymphomas (Ruiz et al., 2011).

R-loops in genetic diseases

Mutations in R-loop resolving factors have instead been identified as aetiologic to several human diseases: RNase H2 is mutated in the neuro-inflammatory disorder Aicardi-Goutières syndrome (Lim et al., 2015), while mutations of SETX have been associated with Ataxia Oculomotor Apraxia type 2 and Amyotrophic Lateral Sclerosis (ALS) type 4. In ALS4 patients, SETX mutation leads to decrease levels of R-loops and consequent changes in the methylation state of the affected promoters, impacting the expression of several genes (Grunseich et al., 2018).

R-loops in repeat expansion disorders

R-loop deregulation has also been observed in several neurologic disorders, e.g. C9orf72-mediated ALS, Friedreich Ataxia, Fragile X syndrome and Huntington's disease. All these syndromes are characterized by expansion of GC-rich repeats, which form long stretches of GC-rich sequences, prone to R-loops formation and possibly further stabilization by 3D structures forming at the ssDNA, leading to transcription perturbation, epigenetic changes and increased repeat instability (Richard and Manley, 2017).

R-loops as therapeutic targets

Once a better understanding on the mechanisms underlying R-loop contribution to pathologies will be achieved, their use as therapeutic targets could be envisioned for a large panel of severe human diseases. DNA:RNA hybrids binding compounds could represent a powerful tool for R-loop based therapies, together with targeting proteins physiologically involved in R-loop metabolism (Perego et al., 2019). However, for this to become a reality, further research is necessary to improve the mechanistic understanding of R-loop biology and their multifaceted characteristics, in order to better decipher how such ubiquitous structures are differentially regulated based on the time, location and context at which they form.

Formation of R-loops is a frequent, pervasive event, ubiquitously observed throughout the phylogenetic tree. The conservation of factors involved in their metabolism is testimony to the fact that their regulation is crucial, both for their controlled formation in physiological processes, and for their removal in order to safeguard nuclear homeostasis. Although research has made considerable progress in recent years towards a better understanding of these remarkable structures, there is still much room for exploration. Refining our knowledge on the biogenesis, functions and turnover of R-loops, and reaching a consensus on the best practices for their detection, will allow a better understanding of the mechanisms underlying numerous biological events, both physiological and pathological.

OBJECTIVES OF THE STUDY AND BIOLOGICAL QUESTIONS

The nuclear pore is emerging as a key element in cell homeostasis, with important roles in molecular trafficking, genome stability maintenance, and, in multicellular organisms, in cell division and the determination of cell identity. This varied repertoire of functions is made possible by the large number of interactions that the subunits of this remarkable structure are able to establish with multiple factors, which result in regulation and spatio-temporal coordination of numerous fundamental biological processes. The nuclear pore thus constitutes an important hub towards which a great variety of molecules converge and in which numerous biological processes are concentrated.

During my PhD, I devoted myself to the study of biological processes spatially restricted to the nuclear pore, with a particular interest in the targeting of messenger RNA-containing structures from both the nucleoplasmic and cytoplasmic sides.

In the context of a study on the co-translational regulation of NPC assembly, my group identified two mRNAs, coding for two nucleoporins of the nuclear basket, showing a stable interaction with the pore once engaged in the translation process. This interaction is in fact dependent on the mRNAs' association with ribosomes, and the interaction of karyopherins with the N-terminal of the nascent protein. We therefore wondered what the functional role of localised translation of this mRNA might be. To answer this question, I devised genetic systems to favour the translation of such mRNAs away from their preferred translation site, in order to be able to assess the consequences of such a perturbation on cell homeostasis.

On the nuclear side, instead, I have focused on the mechanisms of interaction of the nuclear pore with transcriptionally active chromatin. Numerous studies have already been devoted to the characterisation of mechanisms of transcription-dependent or DNA damage-dependent relocation of loci to the nuclear pore. In this frame, we set out to investigate the existence of a similar relationship between the nuclear pore and loci forming R-loops, which constitute an important cause of transcription-dependent genetic instability. Some evidence is already available in the literature supporting this hypothesis: in a cytologic screen performed in budding yeast for the identification of factors involved in R-loop formation, a mutant for the nucleoporin Nup133 was found to accumulate hybrids, suggesting a possible role of the nuclear pore in R-loop metabolism (Chan et al., 2014). Furthermore, a genetic screen for mutants showing hyper-recombination phenotype upon overexpression of the cytidine deaminase AID, identified the nuclear basket nucleoporin Mlp1 (García-Benítez et al., 2017). Increased R-loop levels in this mutant were found to be decreased at loci artificially anchored to the nuclear pore, further proving that mechanisms of R-loop prevention and/or resolution could also be active in the NPC microenvironment.

In light of this, my work was aimed at characterising the mechanism of targeting and interaction of R-loops with the nuclear pore. In particular, I investigated whether the formation of R-loops could indeed modify the localisation of an active gene by favouring its association with the nuclear pore. I then proceeded with the identification of the protein factors involved in the sensing of these structures and their targeting and anchoring to the nuclear pore. Finally, I sought to understand the physiological consequences of R-loop residence in the pore microenvironment, and how these effects are executed.

In the next section I report the results of these investigations in the form of two articles. The first paper presents the study of the co-translational regulation of NPC assembly and the characterization of the molecular mechanisms and functional consequences of localized translation at the nuclear pore. This work, that I sign as co-first author, was published in *Molecular Cell* in 2021. The second paper, instead, comprises the study of R-loop dependent genes relocation to the nuclear pore, the characterization of the pathway of R-loop targeting to the NPC, and the physiological consequences of this phenomenon. This work, that I sign as unique first author, will be submitted for publication in the next months.

RESULTS

ARTICLE 1

Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis

The following manuscript investigates the co-translational events involved in the biogenesis of the nuclear pore complex, identifying several binary interactions between NPC subunits which are established in the cytoplasm while one of the identified partners is engaged in translation. Moreover, we unveiled a mechanism mediating localised translation at the nuclear pore concerning the nucleoporins Nup1 and Nup2, together with a subset of nuclear proteins. Co-translational recruitment of the nascent peptides to the nuclear pore requires the interaction of the nascent N-terminus with nuclear transport receptors.

My contribution to the work has consisted in the determination of the physiological relevance of such spatial restriction of translation at nuclear pore complexes. To this aim, I set up genetic systems to be able to uncouple translation from NPC association and assembly (Fig. 4, S4). Through single molecule Fluorescence In Situ Hybridisation (FISH) and live imaging, I was able to show how translation of the proteins of interest away from the nuclear pore generate the formation of cytoplasmic protein aggregates. I also contributed to several experiments during the revision of the manuscript (Fig. 2, S2, S3).

Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis

Ophélie Lautier^{1,7}, Arianna Penzo^{1,7}, Jérôme O. Rouvière^{1,2}, Guillaume Chevreux³, Louis Collet¹,
Isabelle Loïodice⁴, Angela Taddei⁴, Frédéric Devaux⁵, Martine A. Collart⁶ & Benoit Palancade^{1,8,*}

(1) Université de Paris, CNRS, Institut Jacques Monod, F-75006, Paris, France.

(2) Present address: Department of Molecular Biology and Genetics, Aarhus University, C.F. Møllers Allé 3, DK-8000 Aarhus C, Denmark.

(3) ProteoSeine@IJM, Université de Paris, CNRS, Institut Jacques Monod, F-75006, Paris, France.

(4) Institut Curie, PSL Research University, CNRS, Sorbonne Université, UMR3664 Nuclear Dynamics, Paris, France.

(5) Sorbonne Université, CNRS, Institut de biologie Paris-Seine (IBPS), UMR 7238, Laboratoire de biologie computationnelle et quantitative, LCQB, 4 place Jussieu, 75005, Paris, France.

(6) Department of Microbiology and Molecular Medicine, Faculty of Medicine, University of Geneva, Geneva, Switzerland.

(7) These authors contributed equally.

(8) Lead contact.

* Correspondence to: B. Palancade. Email: benoit.palancade@ijm.fr

Summary

mRNA translation is coupled to multiprotein complex assembly in the cytoplasm, or to protein delivery into intracellular compartments. Here, by combining systematic RNA immunoprecipitation and single molecule RNA imaging in yeast, we have provided a complete depiction of the co-translational events involved in the biogenesis of a large multiprotein assembly, the nuclear pore complex (NPC). We report that binary interactions between NPC subunits can be established during translation, in the cytoplasm. Strikingly, the nucleoporins Nup1/Nup2, together with a number of nuclear proteins, are instead translated at nuclear pores, through a mechanism involving interactions between their nascent N-termini and nuclear transport receptors. Uncoupling this co-translational recruitment further triggers the formation of cytoplasmic foci of unassembled polypeptides. Altogether, our data reveal that distinct, spatially-segregated modes of co-translational interactions foster the ordered assembly of NPC subunits, and that localized translation can ensure the proper delivery of proteins to the pore and the nucleus.

Keywords

mRNA translation ; mRNA localization ; multiprotein complex assembly ; nuclear pore complex ; protein localization ; proteome homeostasis.

Highlights

- Co-translational interactions contribute to the assembly of NPCs.
- A subset of nuclear pore and nuclear proteins are translated at NPCs.
- Recognition of nascent Nup1/Nup2 by karyopherins is required for translation at NPCs.
- Localized translation of Nup1 prevents its aggregation in the cytoplasm.

Introduction

As an essential step in the gene expression process, the assembly of multiprotein complexes from individual polypeptides is tightly connected to the fate of mRNAs. While mRNA-associated factors primarily impact the yield and the subcellular localization of protein synthesis, mRNAs engaged in translation can further serve as platforms on which subunits are sequentially recruited onto their nascent protein partners (Halbach et al., 2009; Duncan and Mata, 2011; Shiber et al., 2018; Panasenko et al., 2019). Alternatively, mRNAs can join by virtue of the simultaneous interaction between polypeptides emerging from the ribosomes (Kamenova et al., 2019). Such co-translational assembly events have been proposed (i) to stabilize nascent chains, or chaperone their folding, (ii) to facilitate the ordered and stoichiometric association of subunits, (iii) to prevent the accumulation of orphan, potentially toxic polypeptides, or (iv) to serve as targets for translational regulations (Natan et al., 2017; Schwarz and Beck, 2019). However, despite methodological advances in the characterization of co-translational interactions, little is known about their physiological relevance for proteome homeostasis, essentially because it has proven difficult to interfere with co-translational assembly without impacting the functional integrity of fully-formed complexes. In bacteria, where translation occurs in the context of chromosomal operons, the subunits of a luciferase complex were successfully prevented from interacting during their synthesis by moving apart the corresponding genes (Shieh et al., 2015), further revealing that the production of functional complexes can be fostered by co-translational assembly.

Among the eukaryotic machineries whose biogenesis is also spatially restricted are the nuclear pore complexes (NPCs), which form either at the surface of post-mitotic chromatin in conjunction with nucleus reformation, or within the existing nuclear envelope during interphase, the latter being the one and only pathway in species with closed mitosis, e.g. fungi (Weberruss and Antonin, 2016). These megadalton-sized proteinaceous complexes are composed of multiple copies of ~30 subunits (the nucleoporins, or Nups) which assemble into structural modules and further delineate a transport channel perforating the nuclear envelope (**Fig. 1A, left panel**). In this place, NPCs provide the unique route for nucleo-cytoplasmic trafficking by virtue of reversible interactions between nucleoporins harboring FG (Phenylalanine Glycine) repeat domains and shuttling nuclear transport receptors (NTRs) such as karyopherins (Kaps), bound to protein or RNA cargoes (Beck and Hurt, 2017). While comprehensive genetic and structural studies have highlighted the robustness of these sophisticated complexes (Fischer et al., 2015; Lin et al., 2016; Kim et al., 2018; Onischenko et al., 2017; Allegretti et al., 2020), their biogenesis does not seemingly involve dedicated assembly cofactors or chaperones, beyond nucleoporins and NTRs themselves (Ryan et al., 2007; Walther et al., 2003), raising the possibility that co-translational interactions could contribute to shape the assembly and prevent the inappropriate targeting of the most aggregation-prone subunits (e.g. FG-Nups).

Results

A systematic screen reveals distinct modes of co-translational interactions between NPC subunits

To investigate a potential coupling between translation and assembly at NPCs, we used a library of budding yeast strains expressing nucleoporins as C-terminal fusions with three repeats of the IgG-binding motif of protein A (Alber et al., 2007; Rout et al., 2000), systematically immunoprecipitated such individually-tagged subunits and further analyzed their association with the full complement of Nup-encoding mRNAs (**Fig. 1A**, *right panel*; **Table S1**). Western-blot analysis confirmed the enrichment of the tagged proteins following immunoprecipitation (as exemplified in **Fig S1A**). RT-qPCR-based quantification of the co-associating RNAs first revealed that nuclear basket-associated Mlp1 and Mlp2 are in contact with most Nup mRNAs (**Fig. 1A**, **a**), in line with their reported function in docking protein-coding transcripts committed to nuclear export (Bonnet and Palancade, 2014). In addition, we observed that a large fraction of the nucleoporins (16/33) associate with their own transcripts in this assay (**Fig. 1A**, **b**). Such interactions were confirmed in individual RNA immunoprecipitation (RIP) experiments and found to be lost in the presence of EDTA (**Fig. S1B**), a treatment leading to polysome dissociation (Duncan and Mata, 2011; Halbach et al., 2009), supporting the view that our procedure pulls down the nascent proteins being synthesized from their cognate mRNAs and exposing the first protein A repeats before translation ends (as schematized in **Fig. 1D**, *top*). While the differential accessibility of this C-terminal tag could account for the fact that not all baits are enriched with their own mRNAs, in agreement with an earlier systematic study (Duncan and Mata, 2011), these observations demonstrate that the integrity of translation complexes is preserved in our experimental conditions. Remarkably, our screen also scored a number of non-reciprocal protein-RNA interactions (indicated by open circles and stars in **Fig. 1A**) that were similarly validated in independent RNA immunoprecipitation (RIP) experiments (**Fig. 1B-C**). EDTA treatment confirmed that these protein-RNA interactions depend on polysome integrity (**Fig. 1B-C**; **Fig. S1C**), thereby supporting that they reflect sequential, co-translational assembly events between the immunoprecipitated protein (referred to as Nup-X in **Fig. 1D** and below) and its nascent partner (Nup-Y), emerging from the ribosome decoding the *NUP-Y* mRNA.

On the one hand, we found that subunits of the same NPC sub-complexes could assemble during translation (indicated by open circles in **Fig. 1A**; **Fig. 1B**), these events matching direct, characterized protein-protein contacts within the inner ring (Nup53-Nup170), the central channel (Gle2-Nup116), the cytoplasmic filaments (Nup159-Nup82) and the nuclear basket modules (Nup60-Nup2), respectively (**Fig. 1E**; **Fig. S1D**). On the other hand, we detected the co-translational association of a subset of FG-Nups with partners within distinct NPC sub-complexes (Nup116-Nup82, Nup157-Nup145 and Nup192-Nup100; indicated by stars in **Fig. 1A**; **Fig. 1C**), in agreement with the reported role of these nucleoporins as linkers between NPC modules (**Fig. 1F**; **Fig. S1E**) (Fischer et al., 2015; Kim et al., 2018; Lin et al., 2016; Onischenko et al., 2017). Of note, the Nup157-Nup145 interaction reportedly

involves the C-terminal domain of Nup157 and the N-terminal moiety of Nup145, which is a FG-nucleoporin arising from the autocatalytic cleavage of the Nup145 precursor (Kim et al., 2018; Lin et al., 2016). Similarly, the interaction domains that were previously characterized for all protein-protein pairs (**Fig. 1E-F**; **Fig. S1D-E**) mapped in most cases to the C-terminus of the immunoprecipitated subunit (Nup-X, **Fig. 1D-F**) and to the N-terminus of its nascent partner (Nup-Y, **Fig. 1D-F**), supporting an ordered assembly pathway where the former is fully translated and folded before its recruitment during the early translation of the latter. Notably, the multivalent nucleoporin Nup82 was found to recruit co-translationally Nup159, Nup116 and Gle2 (**Fig. 1A, E-F**), the latter interaction being likely bridged by Nup116 (**Fig. 1F**), suggesting that formation of this NPC building block could be driven by the translation of its centerpiece subunit (Yoshida et al., 2011).

Finally, we also noticed that *NUPI* and *NUP2* mRNAs, which both encode nuclear basket-associated FG-Nups, associate with several nucleoporins belonging to distinct sub-complexes (e.g. Nup133, Nup59, Pom152; indicated by ‘#’ in **Fig. 1A**), suggesting a close proximity of these transcripts with fully-assembled NPCs.

Translation-dependent targeting of mRNAs encoding NPC-bound proteins

To specifically explore such mRNA localization events, we immunopurified the NPC scaffold upon cryolysis of cells expressing tagged versions of Nup59, Nup82 or Nup116, as previously described (Alber et al., 2007). Silver stain analysis revealed a specific and similar banded pattern in the purifications performed from these three tagged strains, as compared to a control pull-down obtained from untagged cells (**Fig. S2A**). Quantitative mass spectrometry analysis further confirmed that these three isolates are enriched by more than one order of magnitude for the complete set of nucleoporins, with little contaminants beyond nuclear transport factors (**Fig. S2B-D**; **Table S1**), as previously observed (Kim et al., 2018). We then analyzed the mRNAs co-isolating with such affinity-purified NPCs. Strikingly, microarray analysis revealed that *NUPI* and *NUP2* are the most abundant mRNAs in association with Nup59-purified NPCs, besides *NUP59* mRNA itself (**Fig. 2A**). This same subset of mRNAs was confirmedly found in association with the isolated NPCs obtained using Nup82 or Nup116 as alternative baits (**Fig. S2E**). Furthermore, these NPCs isolates were also generally associated with mRNAs encoding nuclear proteins (e.g. *ULP1*, *ESCI*, *TAF1*, *CBP80*; **Fig. 2A**, light blue dots; **Table S1**; **Fig. S2F**), as revealed by gene ontology analysis (GO: “nucleus”: $p=0.001$). In contrast, our assay did not identify other control, abundant RNAs in interaction with NPCs (e.g. *ACT1*; **Table S1**; **Fig. S2G**), supporting that the above-mentioned hits do not represent mRNAs caught in transit during nuclear export. Importantly, the co-translational formation of NPC building blocks and their association with linker subunits as described above (**Fig. 1B-F**) do not appear to similarly occur at assembled pores since the corresponding mRNAs (*NUP170*, *NUP116*, *NUP82*, *NUPI45*, *NUP100*) were not enriched with NPCs (**Table S1**).

To independently confirm the association of *NUP1* and *NUP2* mRNAs with NPCs *in vivo*, we analyzed their subcellular localization by performing single molecule fluorescence *in situ* hybridization (smFISH). While most of these transcripts were randomly distributed in the cytoplasm, a significant fraction of them was detected at the nuclear periphery when compared to another Nup-encoding mRNA which is similarly expressed yet not enriched in the NPC pull-down (*NSP1*; **Fig. 2B-C**), or to another unrelated transcript (*RNR3*, **Fig. S2H**). To refine the localization of this perinuclear mRNA sub-population, we further performed the smFISH experiment in $\Delta N-nup133$ mutant cells, a genetic background previously reported to display NPC clustering within the nuclear envelope and classically used to resolve pores from the remaining nuclear envelope (Doye et al., 1994). *NUP1* and *NUP2* mRNAs significantly colocalized with GFP-labeled clustered pores in this context, confirming their association with NPCs (**Fig. S2I**).

To investigate whether the association of these mRNAs with NPCs is also co-translational, we further interfered with protein synthesis through the following independent approaches (**Fig. 2D**) (Blobel and Sabatini, 1971; Duncan and Mata, 2011; Halbach et al., 2009): (i) polysomes were dissociated in the presence of EDTA, as described above ; (ii) polysomes were disassembled by treatment with puromycin at increased potassium concentrations (**Fig. S2J**) ; (iii) *NUP1* or *NUP2* start codons were deleted from GFP-fused transgenes to prevent their specific translation (**Fig. S2K**). All tested conditions virtually abolished the association of *NUP1*, *NUP2* or other target mRNAs with NPCs (**Fig. 2E-F**; **Fig. S2F**). Our findings thereby imply that a subset of NPC-bound proteins is targeted to assembled NPCs during the course of their translation.

Karyopherin-mediated recognition of nascent proteins mediates their co-translational association with NPCs

Mechanism-wise, the requirement for translation prompted us to question the role of nascent polypeptides in the association of these mRNAs with NPCs. Comparison of N-terminal domains of the proteins encoded by NPC-associated mRNAs revealed a systematic occurrence of nuclear localization determinants, including NLS (nuclear localization signals) or other validated interaction domains (ID) for NTRs of the karyopherin (Kap) family (**Fig. 3A**; **Fig. S3A**). In the unique case for which such information was not available (i.e. Esc1), our deletion analyses similarly demonstrated the existence of a nuclear targeting determinant in the N-terminal region of the protein (**Fig. S3B**). These observations raised the possibility that Kaps could bridge the interaction between NPCs and mRNAs undergoing translation for which N-termini have already emerged from the ribosome. Supporting this hypothesis, mass spectrometry identified a subset of Kaps within our NPC isolates enriched for *NUP1* and *NUP2* mRNAs (**Fig. S2B-D**). In addition, deletion of Nup1 and Nup2 N-terminal Kap-IDs precluded the association of the corresponding mRNAs to NPCs (**Fig. 3B-C, left panels**), although the truncated proteins were properly translated (**Fig. S3C**). Importantly, the targeting signal is carried by the encoded

protein domain rather than the corresponding mRNA sequence in view of the lack of NPC association scored for ΔAUG mutant transcripts (see above, **Fig. 2E-F**).

To complement this finding, we interfered with karyopherin activity as follows. In line with the reported interaction of both Kap60-Kap95 and Kap121 with Nup1 Kap-ID (Mészáros et al., 2015), we analyzed *NUP1* mRNA association to NPCs upon over-expression of the *kap123 ΔN* dominant negative mutant, which leads to the dissociation of both karyopherins from NPCs (Panse et al., 2003; Schlenstedt et al., 1997). A short induction of the *kap123 ΔN* mutant, while mislocalizing a *bona fide* Kap60-Kap95/Kap121 target (i.e. Ulp1; Panse et al., 2003; **Fig. S3D**) without irreversibly impacting cell growth (**Fig. S3E**), strongly affected *NUP1* mRNA association to NPCs (**Fig. 3B, right panel; Fig. S3F**). Similarly, in agreement with the specificity of Nup2 Kap-ID for the Kap60-Kap95 dimer (Dilworth et al., 2001), rapid relocalization of Kap60 away from NPCs through the anchor away approach (Haruki et al., 2008) led to the dissociation of *NUP2* mRNAs from NPCs (**Fig. 3C, right panel**). Yet, ectopically fusing the Nup2 N-terminal Kap-ID to a reporter mRNA (*GFP*, **Fig. 3D**) was not sufficient to recapitulate this association with NPCs (**Fig. 3E, left panel, arrow**), demonstrating that additional features of NPC-associated mRNAs contribute to their targeting.

In view of the recently reported links between ribosome pausing and co-translational assembly (Panasencko et al., 2019), we thereby analyzed the translational features of these transcripts. By examining the association with translational regulators in published datasets (Hogan et al., 2008; Wolf et al., 2010; Rouviere et al., 2018), we observed that a large subset of the NPC-bound mRNAs (e.g. *NUP1*) are targeted by the hnRNP-K-like Hek2/Khd1 protein, a repressor of translation initiation (**Fig. S3G**). However, *NUP1* mRNAs were still detected in association with NPCs in the absence of Hek2 (**Fig. S3H, top panel; Fig. S3I**), showing that the translational attenuation mediated by this RNA-binding protein is not required for the localization of its targets at the nuclear envelope. To further investigate how translational elongation could impact mRNA targeting to NPCs, we analyzed the distribution of ribosomes onto NPC-associated mRNAs in ribosome footprinting assays, and detected a major ribosome pausing event within *NUP1* CDS, occurring at a typical pause-inducing Pro-Pro di-codon (**Fig. S3J-L**). Yet, mutation of the Pro-Pro pause site within *NUP1* mRNAs did not impact their targeting to NPCs ('*PPAA*' mutant in **Fig. S3H, bottom panel; Fig. S3I**). Finally, analysis of the sequence features of NPC-associated mRNAs revealed that the length of the translated region, a proxy of translational duration, was significantly higher as compared to the rest of the transcriptome (**Fig. 3F**). To assay the importance of CDS length for NPC targeting, we expressed Nup2 Kap-ID in the context of a lengthened reporter mRNA (Nup2-KapID-4xGFP, **Fig. 3D**). Strikingly, this long chimeric mRNA was efficiently targeted to NPCs, as opposed to its shorter counterpart (Nup2-KapID-1xGFP, **Fig. 3E**). This finding establishes that a nascent Kap-ID is sufficient to target a heterologous mRNA to NPCs, and supports that a prolonged translation allows this recruitment to occur prior to the release of the complete polypeptide from the ribosome.

Taken together, our data support a model in which karyopherin-mediated recognition of nascent polypeptidic chains target specific mRNAs (e.g. *NUP1*, *NUP2*) to NPCs before their terminating translation, ensuring the co-translational import and assembly of the proteins (**Fig. 3G**).

Uncoupling co-translational interactions impacts nucleoporin homeostasis

The spatially-restricted interaction between these nascent FG-nucleoporins (Nup1/Nup2) and NPCs provided a unique situation to decipher the physiological significance of such co-translational association events. To uncouple translation from assembly, we first took advantage of the fact that *HEK2* inactivation substantially increases the fraction of polysome-associated *NUP1* mRNAs (Rouviere et al., 2018) while barely impacting their association with NPCs (**Fig. S3I**; **Fig. S4A**). Strikingly, in this situation where the cytoplasmic pool of *NUP1* mRNAs is likely translated, we detected the appearance of cytoplasmic foci of unassembled Nup1 polypeptides in a small subset of the cells (**Fig. S4B**), in line with previous observations (Rouviere et al., 2018).

To independently confirm the functional relevance of Nup1 translation at NPCs, we fused the full *NUP1* coding sequence to the 3' untranslated region of the *ASH1* mRNA, a motif sufficient to localize heterologous transcripts to the bud tip of dividing yeasts (Long et al., 1997), and used as a control a fusion harboring the 3' region from the housekeeping *ADHI* mRNA (**Fig. 4A**). smFISH experiments confirmed that *NUP1* mRNAs were successfully redirected to the bud in a fraction of the cells carrying the *NUP1^{ASH1 3'}* construct as compared to those expressing the *NUP1^{ADHI 3'}* fusion (**Fig. 4B-C**). Strikingly, when we further monitored the consequences of *NUP1* mRNA mislocalization on the recruitment of Nup1 polypeptides to NPCs, we detected increased levels of Nup1(-GFP) foci in the cytoplasm and the bud of *NUP1^{ASH1 3'}* cells as compared to *NUP1^{ADHI 3'}* cells (**Fig. 4D-E**; **Fig. 4F, left column**), although both constructs allowed to accumulate similar levels of Nup1 polypeptides (**Fig. S4C**). These Nup1 foci did not contain other scaffold or FG-nucleoporins (e.g. Nic96 or Nup49), supporting that they do not represent NPC assembly intermediates, but trapped the Kap60 karyopherin (**Fig. 4F**), suggesting that their accumulation could interfere with nuclear import at NPCs. Of note, the presence of the *ASH1* 3'UTR does not cause a complete mislocalization of *NUP1* transcripts (**Fig. 4B-C**), possibly explaining why Nup1 proteins can still get access to NPCs in these cells (**Fig. 4D**). Our data thus reveal that coupling translation with assembly can prevent the deleterious cytoplasmic aggregation of an FG-nucleoporin.

Discussion

Altogether, our data reveal that distinct, spatially-segregated modes of co-translational interactions partake in NPC biogenesis: while the co-translational assembly of NPC sub-complexes and their association with linker FG-Nups occur in a sequential manner, away from the NPC, local

translation of a subset of NPC-bound proteins ensure their proper delivery at the nuclear envelope. Importantly, the low fraction of *NUP1* and *NUP2* mRNAs found at NPCs (**Fig. 2B**; **Fig. S2I**) is compatible with their supplying the bulk of polypeptides required for interphase pore biogenesis (**Fig. S4D**).

Yet, the situation in dividing yeast cells strikingly differs from what has been recently observed in conditions of cell cycle arrest during *Drosophila* oogenesis, where Nup mRNAs are translated away from the nucleus, at the surface of precursor condensates which further progress into NPCs (Hampoezl et al., 2019). However, a common feature of both systems is the pivotal role of NTRs in NPC assembly, in line with earlier findings (Ryan et al., 2007; Walther et al., 2003). While the molecular mechanism proposed here positions NTRs as molecular anchors for translating mRNAs at NPCs, their binding to nascent polypeptides could actively prevent their aggregation, consistent with the proposed chaperone functions of karyopherins (Jäkel et al., 2002; Padavannil et al., 2019). In this respect, our specific experimental conditions allowed to detect co-translational assembly events for a large fraction of the NPC subunits (13/33), notably FG-Nups. Remarkably, the localization of interaction surfaces within the C-terminal regions of other scaffold nucleoporins such as most subunits of the outer ring complex (Kim et al., 2018) could preclude their co-translational association. Nonetheless, it is likely that coupling translation with assembly or import would be particularly relevant for (i) aggregation-prone subunits (e.g. FG-Nups), (ii) misfolding-prone, long polypeptides, which are preferentially targeted by NPC-associated translation in our assays (**Fig. 3E-F**), or (iii) other nuclear proteins whose cytoplasmic accumulation could be toxic (e.g. the SUMO-protease Ulp1 or the nuclear Cap-binding complex subunit Cbp80; **Fig. 2A**). Whether such polypeptides are exclusively translated at NPCs remains however to be determined. Interestingly, only a minor fraction of *NUP1* mRNAs (<25%) undergoes translation in *wt* cells (as scored by our polysome profiling analysis, Rouviere et al., 2018), in line with the detected subset translated at NPCs (15%, **Fig. 2C**). In this situation, the cytosolic, translationally-inactive mRNAs could serve as a reservoir whose translation could be rapidly reactivated upon demand, leading to their repositioning at the nuclear periphery.

In the absence of interactions with motor proteins in systematic analyses (Casolari et al., 2012), mRNAs encoding NPC-bound proteins do not appear to follow well-described cytoskeleton-dependent RNA localization routes (Singer-Krüger and Jansen, 2014). Rather, our described pathway for NPC-bound proteins is strikingly reminiscent of the signal recognition particle-mediated targeting of mRNAs encoding membrane and secreted proteins to the endoplasmic reticulum (ER). In both cases, the recognition of an N-terminal signal in the nascent chain by a soluble receptor brings the translating mRNA for local delivery of the protein at a membrane-bound compartment (**Fig. S4E**). In view of the independent evolutionary trajectories of the NPC and the ER translocon, such functional convergence may underlie a common requirement in protecting the cytoplasm from toxic or insoluble proteins.

Limitations. While we have scored here a number of co-translational events involving nucleoporins, it is possible that additional binary interactions, also impacting NPC biogenesis, could not

be identified because of their stability or half-life in our experimental conditions. Similarly, it is likely that other, smaller karyophilic proteins are also translated at NPCs but could not be captured here due to the shorter duration of their translation, that would be terminated in the time frame of our experiments, to the lower number of translating ribosomes that could partake in their NPC association, or to the size of their nascent interacting-domain, whose binding to the NPC would be sterically precluded. Finally, while our data establish that recognition of the N-terminal Kap-ID by karyopherins is necessary to direct *NUP1* and *NUP2* mRNA translation to the NPC, we do not exclude the possibility that additional interactions, beyond the recognition of the nascent chains, participate to the anchoring of the ribosomes at the nuclear envelope. In the future, improving the spatio-temporal resolution of such translational analyses will likely solve these opened questions.

Acknowledgements:

We are very grateful to Jeffrey Gerst, Won-Ki Huh, Ed Hurt, Sébastien Léon and Elmar Schiebel for reagents ; to members of the ProteoSeine@IJM facility (Université de Paris, CNRS, Institut Jacques Monod) for mass spectrometry analyses ; to Mike Rout and Svetlana Dokudovskaya for providing the tagged nucleoporin strains ; to George Allen for analysis of the ribosome profiling data ; to Vedrana Andric, Michel Werner and other members of the Palancade and Libri lab for fruitful discussions ; and to Bertrand Cosson, Valérie Doye, Vincent Géli, Anahi Molla-Herman and Mathieu Rougemaille for their critical reading of the manuscript.

This work was supported by Agence Nationale pour la Recherche (ANR-18-CE12-0003, to B.P.), Fondation ARC pour la recherche contre le Cancer (projet ARC, to B.P.), Ligue Nationale contre le Cancer (comité de Paris, to B.P.) and the Swiss National Science Foundation (grant 31003A_172999, to M.A.C.).

Author contributions:

Conceptualization, B.P.; Methodology, A.P., A.T., G.C., F.D., M.A.C., B.P.; Investigation, O.L, A.P., J.O.R, G.C., L.C., I.L., A.T., F.D., M.A.C., B.P.; Visualization, O.L., A.P., I.L., B.P.; Writing - Original Draft, B.P.; Writing - Review & Editing, A.P., J.O.R., A.T., F.D., M.A.C.; Funding Acquisition, M.A.C., B.P.; Supervision, B.P.

Declaration of interests: The authors declare no competing interests.

Figure legends

Figure 1 – Distinct modes of co-translational interactions between NPC subunits.

A, *Left panel*, schematic representation of the yeast nuclear pore complex (NPC) showing sub-complexes of nucleoporins (Nup) as colored circles. FG-Nups are underlined. *Right panel*, the association of Nup-encoding mRNAs (columns) to every protein A (pA)-tagged Nups (lines) was analyzed by systematic RIP followed by RT-qPCR, and represented as the \log_2 ratios between immunopurified and input RNAs (relative to an untagged control strain) using the indicated color code (blue to yellow for values between 1 and 3; see **Table S1** for the complete dataset). **B-C**, The association between different mRNAs and the indicated pA-tagged Nups (baits) was analyzed by RIP followed by RT-qPCR and represented as the ratios between immunopurified and input RNAs (mean \pm SD, n=4, relative to untagged). When indicated, the purification was performed in the presence of 40 mM EDTA. *, p<0.05 (Mann-Whitney-Wilcoxon test). **D**, Schematic representation of the co-translational interactions between Nups as scored by RIP. **E-F**, Characterized interaction domains (colored boxes) for the corresponding pairs of proteins. The *NUP145* mRNA encodes a nucleoporin which is autocatalytically cleaved into Nup145-N (FG-Nup, *plain line*) and Nup145-C (outer ring, *dashed line*). The Gle2-Nup82 interaction is most likely bridged by Nup116, in line with the reported identification of the corresponding trimeric complex (Alber et al., 2007). See also **Fig. S1**.

Figure 2 – Translation-dependent targeting of mRNAs encoding NPC-bound proteins.

A, NPC-associated mRNAs were immunoprecipitated using Nup59-pA as a bait and analyzed by microarray. The levels of association (average \log_2 ratios between immunopurified and input RNAs; n=2) and significance ($-\log_{10}$ p-value) are represented for the whole transcriptome. Enriched mRNAs (\log_2 IP/input>1; p<0.001) are displayed according to the indicated color code with those encoding nuclear proteins (GO: cellular compartment = nucleus) appearing in light blue. Other mRNAs (black dots) are *FLO11*, *MTL1*, *YPS3* and *TPO4* (see also **Table S1** for the complete dataset). **B**, Single-molecule FISH was performed on *wt* cells using sets of probes specific for *NUP1*, *NUP2* or *NSP1* mRNAs. The same cells were hybridized with *NUP1* and *NSP1* probe sets labeled with distinct fluorophores. SmFISH images (all pseudo-colored in red), as well as merged images with the DAPI channel (nuclear staining) are shown. Z-projections are displayed, except for Differential Interference Contrast (DIC) images. Arrowheads point to nuclei exhibiting perinuclear localization for *NUP1* but not for *NSP1* mRNAs. Scale bar, 5 μ m. **C**, *Left panels*, the distribution of mRNA particles in the smFISH experiments (% of total; from **B**.) is represented according to their distance to the nuclear staining (pixels, px). *Right panels*, Plotted is the fraction of particles localized in the perinuclear area (1px < distance < 3px=200 nm, highlighted in blue). The total number of counted mRNA particles and cells (n₁/n₂) from three independent experiments for each probe set is as follows: *NUP1*: 353/55; *NUP2*: 503/55; *NSP1*: 618/61. *, p<0.05; ***, p<0.001 (Fisher exact test). **D-F**, The association between *NUP1* (**E**) and *NUP2* (**F**) mRNAs and NPCs (Nup59-pA) was analyzed by RIP followed by RT-qPCR and

represented as the ratios between immunopurified and input RNAs (mean±SD, n=4, relative to untagged and normalized to *ACT1* mRNA). When indicated, the purification was performed in the presence of 40 mM EDTA (*left panels*) or 2 mM puromycin / 0.2M [K⁺] (*middle panels*), or following expression of *wt* or Δ *AUG* variants of *NUP1-GFP* or *NUP2-GFP* (*right panels*). Transgene-derived mRNAs were detected using *GFP*-specific primer pairs. Note that puromycin is expected to be ~50% active in these experimental conditions (see STAR Methods). *, p<0.05 (Mann-Whitney-Wilcoxon test). See also **Fig. S2**.

Figure 3 – Karyopherin-mediated recognition of nascent proteins mediates their co-translational association with NPCs.

A, Domain organization of Nup1 and Nup2 proteins featuring N-terminal karyopherin interaction domains (Kap-ID) and FG-repeats. The associated Kaps and the mutants interfering with their activity are represented. **B-C**, The association between *NUP1* (**B**) and *NUP2* (**C**) mRNAs and NPCs (Nup59-pA or Nup82-pA) was analyzed by RIP followed by RT-qPCR and represented as the ratios between purified and input RNAs (mean±SD, n=4, relative to untagged and normalized to *ACT1* mRNA). *Left panels*, the purification was performed following expression of *wt* or Δ *Kap-ID* variants of *NUP1-GFP* or *NUP2-GFP*. Transgene-derived mRNAs were detected using *GFP*-specific primer pairs. *Right panels*, Kap activities were inhibited upon galactose-induced over-expression of *kap123 Δ N* for 1.5 hrs (**B**) or upon a short (5 min) treatment with 10 μ g/mL rapamycin in the *KAP60-AnchorAway* (*KAP60-AA*) background (**C**). Anchor away of the Kap60-Kap95 dimer leads to a complete inhibition of nuclear import in these conditions (Haruki et al., 2008). **D**, Schematic representation of the Nup2 constructs used in RIP assays. **E**, The association between NPCs (Nup59-pA) and chimeric mRNAs encoding fusions of Kap-ID^{Nup2} to the indicated number of GFP moieties was analyzed similarly (mean±SD, n=4, relative to untagged and normalized to *ACT1* mRNA). Chimeric mRNAs were detected using *Kap-ID^{Nup2}*-specific primer pairs. **F**, Distribution of CDS lengths for mock mRNAs (most enriched mRNAs in a mock immunoprecipitate, see STAR Methods), NPCs-enriched mRNAs (from **Fig. 2A**) or unbound mRNAs (log₂ IP/input<1 or p>0.001 in **Fig. 2A**). **G**, Model of karyopherin function in NPC-associated mRNA translation. *, p<0.05; ***, p<0.001 (Mann-Whitney-Wilcoxon test). See also **Fig. S3**.

Figure 4 – Uncoupling co-translational interactions impacts nucleoporin homeostasis.

A, Plasmid-borne constructs used to interfere with *NUP1* mRNA targeting to NPCs. **B**, smFISH was performed on *NUP1^{ASH1 3'}* or *NUP1^{ADHI 3'}*-expressing cells using a set of *GFP*-specific probes. Budding cells are outlined and the positions of the mother cells (m) and buds (b) are indicated. **C**, The fraction of cells exhibiting *NUP1* mRNAs mislocalization to the bud is represented. The total number of counted cells from two independent experiments is as follows: *NUP1^{ADHI 3'}*, 403; *NUP1^{ASH1 3'}*, 271. **D**, Live imaging of cells carrying *NUP1-GFP^{ASH1 3'}* or *NUP1-GFP^{ADHI 3'}* constructs. Note that the slight Nup1 over-expression driven by both constructs can trigger nuclear envelope abnormalities. **E**, The fraction

of cells exhibiting cytoplasmic or bud-localized foci of Nup1-GFP is represented. The total number of counted cells from seven independent experiments is as follows: *NUP1^{ADH1 3'}*, 1130; *NUP1^{ASH1 3'}*, 1005. **F**, Live-cell imaging of cells carrying the *NUP1-GFP^{ASH1 3'}* construct and expressing Kap60-mCherry, Nic96-mRFP or Nup49-mCherry. For all panels, z-projections are shown for the GFP or mCherry channels, except for the *NIC96-mRFP* strain, where single plane images are displayed. Merged images with DIC are shown, as well as arrows indicating Nup1-GFP foci. Scale bar, 5 μ m. ***, p<0.001 (Fisher exact test). See also **Fig. S4**.

STAR Methods

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit IgG (for coupling)	Sigma	Cat#I5006
rabbit polyclonal IgG-HRP (to detect protein-A-tagged proteins)	Sigma	Cat#P1291
monoclonal anti-Dpm1	ThermoFisher Scientific	Cat#A6429
monoclonal anti-GFP (clones 7.1 & 13.1)	Sigma	Cat#11814460001
Chemicals, Peptides, and Recombinant Proteins		
rapamycin	LC laboratories	Cat#R-5000
protease inhibitor cocktail, complete, EDTA-free	Sigma	Cat#11836170
pepstatin A	Sigma	Cat#P5318
PMSF	Sigma	Cat#P7626
antifoam B	Sigma	Cat#A5757
RNAsin	Promega	Cat#N251A
puromycin	Invivogen	Cat#Ant-pr-1
heparine	Sigma	Cat#H3149
MS-grade trypsin	Promega	Cat#V511
random hexamers (P(dN)6)	Sigma	Cat#11034731001
Critical Commercial Assays		
magnetic beads (Dynabeads M270 epoxy)	ThermoFisher Scientific	Cat#14302D
Nucleospin RNAII kit	Macherey Nagel	Cat#740955
Superscript II reverse transcriptase	ThermoFisher Scientific	Cat#18064014
LightCycler 480 system	Roche	Cat#04887352001
C18 analytical column	Evosep	Cat#EV1106
Deposited Data		
microarray source data for RIP-chip experiments (Fig. 2A, Table S1)	ArrayExpress	E-MTAB-9521
original data for gels, microscopy images and growth assays (Fig. 2, 4, S1-4)	Mendeley Data	http://dx.doi.org/10.17632/gt852xk7cf.1
NGS source data for ribosome footprinting experiments	SRA	PRJNA512900
Experimental Models: Organisms/Strains		
Yeast strains used in this study	see the detailed list in Table S2	
Oligonucleotides		
qPCR primers	see the detailed list in Table S3	
smFISH probes	see the detailed list in Table S3	
Recombinant DNA		
Plasmids	see the detailed list in Table S4	
Software and Algorithms		
Prism v8.0.2	Graphpad	
Maxquant v1.6.3.4	Tyanova et al, 2016	
Perseus v1.6.2.3	Tyanova et al, 2018	
Metamorph v6.3	Molecular Devices	
ImageJ v1.48	https://imagej.nih.gov/ij/	
LIMMA	Ritchie et al., 2015	
cutadapt	Martin, 2011	
HISAT2	Kim et al, 2015	
Bowtie	Langmead and Salzberg, 2012	
Ribowaltz	Lauria et al., 2018	

RESOURCE AVAILABILITY

Lead Contact. Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Benoit Palancade (benoit.palancade@ijm.fr).

Materials Availability. Yeast strains and plasmids generated in this study are available upon request, without restrictions.

Data and Code Availability. The complete microarray data generated during this study are available in the ArrayExpress database under accession number E-MTAB-9521. Mass spectrometry quantifications are provided in **Table S1** and raw data are available upon request. Source data for gels, microscopy and growth assays images are available through Mendeley Data (DOI: 10.17632/gt852xk7cf.1).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Yeast Strains and Growth. All *S. cerevisiae* yeast strains used in this study (listed in **Table S2**) were obtained by homologous recombination and/or successive crosses. Cells were grown at 30°C in standard yeast extract peptone dextrose (YPD) or synthetic complete (SC) medium supplemented with the required nutrients. Experiments requiring *GAL-kap123ΔN* induction were performed in *GAL2⁺* backgrounds (W303 or BY4741) by adding galactose (2%) for 1.5 hrs to cells grown in glycerol-lactate (GGL: 0.17% YNB, 0.5% ammonium sulfate, 0.05% glucose, 2% lactate and 2% glycerol) supplemented with the required nutrients. For Kap60 anchor away experiments, rapamycin (10 μg/mL, LC laboratories) was added to the medium for 5 min at 30°C before harvesting cells (Haruki et al., 2008). Growth assays were performed by spotting serial dilutions of cells on SC medium and incubating the plates at 30°C.

Plasmids. The construction of the plasmids used in this study (listed in **Table S4**) was performed using standard PCR-based molecular cloning techniques and was checked by sequencing. Expression of *NUP1* and *NUP2* variants from their cognate promoters was achieved in the presence of the endogenous *wt* proteins (see **Fig. S3C**).

METHOD DETAILS

RNA immunoprecipitation (RIP) and RIP-chip. Nucleoporins were all individually expressed from their genomic loci as C-terminally tagged fusions with three tandem repeats of the IgG-binding motif of *Staphylococcus aureus* protein A, using previously characterized strains (Rout et al., 2000; Alber et al., 2007; Chadrin et al., 2010). To investigate the RNA partners of the Nup145 nucleoporin, which is autocatalytically cleaved into Nup145-N and Nup145-C, the protein A tag was inserted either following the cleavage site (*NUP145N-protA* strain, referred to as “*NUP145N*” in **Fig. 1A**), or at the end of the *NUP145* CDS (*NUP145-protA* strain, referred to as “*NUP145C*” in **Fig. 1A**). The Sec13 nucleoporin bait exhibited non-specific RNA binding and was excluded from the analysis. Untagged isogenic strains, together with a strain expressing the protein A tag alone, were used as specificity controls.

Except for experiments aimed at preserving the entire NPC scaffold (**Fig. 2A; Fig. S2A-E**), small-scale RNA immunoprecipitations were performed as described (Rouviere et al., 2018). Magnetic beads (Dynabeads M270 epoxy, ThermoFisher) were conjugated with rabbit IgG (Sigma) according to the manufacturer's instructions. Exponentially growing cells (50 mL at $0.5 \leq OD_{600} \leq 1$) were harvested by centrifugation for 5 min and pellets were frozen in liquid nitrogen. Of note, translation was not stalled with cycloheximide since we observed that this treatment induces the transcriptional upregulation of *NUP1* and *NUP2*, a confounding effect already documented in ribosome profiling analyses (Santos et al., 2019). Cells were further lysed by bead-beating with a Fastprep device (Qbiogene) in the following extraction buffer: 20 mM Hepes pH 7.5, 110 mM KOAc, 2 mM MgCl₂, 0.1% Tween-20, 1% Triton X-100, 1 mM DTT, 1X protease inhibitor cocktail (complete, EDTA-free, Roche), 4 μg/mL pepstatin A (Sigma), 180 μg/mL PMSF (Sigma), antifoam B (Sigma, 1:5,000) and 40U/mL RNasin (Promega). From the soluble extracts further recovered upon centrifugation at 10,000g for 5 min at 4°C, an input fraction (5%) was kept for total RNA and protein analysis and the remaining sample was immunoprecipitated for 30 min at 4°C with IgG-conjugated magnetic beads. The beads were washed 3 times with Extraction Buffer, an aliquot was further eluted with SDS Sample Buffer for protein analysis and the remaining sample was directly used for RNA extraction. Experiments involving EDTA-mediated dissociation of polysomes were performed similarly, except that when indicated, the extraction buffer was supplemented with 40 mM EDTA, as described (Fleischer et al., 2006). For experiments involving *in vitro* puromycin treatment, the extraction buffer composition was adjusted to support puromycin activity while preserving interaction of *NUP1/NUP2* mRNAs with NPCs (see **Fig. S2J**). Since puromycin-mediated polypeptide release requires high [K⁺] and low [Mg²⁺] concentrations *in vitro* and polysome dissociation further necessitates heating (Blobel and Sabatini, 1971), the extraction buffer contained 200 mM KOAc, 1 mM MgCl₂, and extracts were heated for 15 min at 30°C prior to immunoprecipitation. When indicated, puromycin (2 mM, Invivogen) was added before heating. In these buffer conditions, puromycin treatment releases ~50% of polypeptides from polysomes (Blobel and Sabatini, 1971). For experiments involving puromycin treatment or *KAP60-AA* strains, 1 mg/ml heparin (Sigma) was included in the extraction buffer and the duration of immunoprecipitation was reduced to 10 min to reduce non-specific RNA binding. In all cases, pull-down efficiency and specificity was controlled by detecting both bait and control proteins in input and eluates samples, as exemplified in **Fig. S1A**.

For large-scale RNA immunoprecipitations (**Fig. 2A; Fig. S2A-E**), NPCs were isolated by affinity purification using tagged versions of Nup59, Nup82 or Nup116 in conditions preserving the interactions of these bait subunits with the NPC scaffold (Alber et al., 2007). Exponentially-growing cells were harvested, frozen in liquid nitrogen and cryogenically lysed using a MM300 mixer mill (Retsch). Frozen grindates (2 g) were resuspended in 9 volumes of extraction buffer (as above) using a Polytron homogenizer (Kinematica). The resulting extract was clarified by centrifugation at 2,500 g for 5 min followed by filtration through 1.6 μm GD/X Glass Microfiber syringe filters (25 mm, Whatman),

and further incubated for 30 min at 4°C with IgG-conjugated magnetic beads. Beads were washed 3 times with extraction buffer, once with 0.1 M NH₄OAc, 0.1 mM MgCl₂, 0.02% Tween-20 and four times with 0.1 M NH₄OAc, 0.1 mM MgCl₂. A fraction of the sample (30%) was directly used for RNA extraction while the bound proteins were eluted with 0.5M NH₄OH, 0.5 mM EDTA, lyophilized and resuspended in 25 mM NH₄HCO₃ for mass spectrometry analysis. The efficiency of NPC purification in the eluates samples was further controlled by SDS-PAGE/silver staining. Stained gels were imaged using a GelDoc EZ system (BIORAD).

Mass spectrometry. Eluates obtained as above were incubated overnight at 37°C in 20 µL of 25 mM NH₄HCO₃ buffer containing 0.2 µg of sequencing-grade trypsin (Promega). The resulting peptides were loaded and desalted on evotips provided by Evosep according to the manufacturer's procedure. Samples were analyzed on an Orbitrap Fusion mass spectrometer (ThermoFisher Scientific) coupled with an Evosep one system operating with the 30SPD method developed by the manufacturer. Briefly, the method is based on a 44-min gradient and a total cycle time of 48 min with a C18 analytical column (0.15 x 150 mm, 1.9 µm beads, Evosep) equilibrated at room temperature and operated at a flow rate of 500 nl/min. H₂O / 0.1 % formic acid (FA) was used as solvent A and acetonitrile / 0.1 % FA as solvent B. The mass spectrometer was operated in data-dependent acquisition (DDA). Peptide masses were analyzed in the Orbitrap cell in full ion scan mode, at a resolution of 120,000, a mass range of m/z 350-1550 and an AGC target of $4 \cdot 10^5$. MS/MS were performed in the top speed 3s mode. Peptides were selected for fragmentation by Higher-energy C-trap Dissociation (HCD) with a Normalized Collisional Energy of 27% and a dynamic exclusion of 60 seconds. Fragment masses were measured in an Ion trap in the rapid mode, with an AGC target of $1 \cdot 10^4$. Monocharged peptides and unassigned charge states were excluded from the MS/MS acquisition. The maximum ion accumulation times were set to 100 ms for MS and 35 ms for MS/MS acquisitions respectively.

Raw data were analyzed using the MaxQuant software (Tyanova et al., 2016) version 1.6.3.4 with the embedded Andromeda peptide search engine (Cox et al., 2011). MS/MS spectra were searched against the Uniprot *Saccharomyces cerevisiae* (strain ATCC 204508 / S288c) reference proteome FASTA file (release 2020_06, 6049 entries) modified to include the sequence of rabbit IGHG coupled to the beads in the pulldown (P01870), and to differentiate Nup145N and Nup145C polypeptides. MaxQuant analysis was mostly conducted with default parameters, with the following exceptions: Cys-Cys was set as a fixed modification and oxidation of methionine, protein N-terminal acetylation and deamidation as variables. "Match between runs" and iBAQ were enabled with the default parameters. Peptide and protein identification were performed at both peptide spectrum match and protein false discovery rate of 1%. Protein abundances were estimated using the iBAQ algorithm (Intensity Based Absolute Quantification) on proteins identified with a minimum of two peptides (Schwanhäusser et al., 2011). Perseus version 1.6.2.3 was used to process subsequently the identification and quantification data contained in the final proteinGroups.txt file (Tyanova and Cox, 2018). Proteins matching to the reverse database or common contaminants list were filtered and data were then transformed to a

logarithmic scale by specifying a $\text{Log}_2(x)$ function. A categorical annotation row was added to specify the groups of independent triplicates belonging to each conditions *i.e.* no tag and Nup59, Nup82, Nup116 baits. An additional filtering based on a minimum of three valid values in at least one group was finally applied before imputation of missing values based on a normal distribution (Lazar et al., 2016). The processed data were finally used for statistical analysis *i.e.* evaluation of differentially present proteins between groups using a Student's bilateral t-test and assuming equal variance between groups.

Ribosome footprinting. Ribosome profiling was performed in biological duplicates as described (Panasencko et al., 2019). For the Ribo-Seq samples, all fastq files were adapter stripped using cutadapt (Martin, 2011). Only trimmed reads were retained, with a minimum length of 20 and a quality cutoff of 2 (parameters: `-a 10 CTGTAGGCACCATCAATAGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -- trimmed-only -- minimum-length=20 --quality-cutoff=2`). Histograms were produced of ribosome footprint lengths that were very homogenous with highest reads between 28 and 31 that were kept for the analysis (**Fig. S3J**). Reads were mapped, using default parameters, with HISAT2 (Kim et al., 2015) to R64-1-1, using Ensembl release 84 gtf for transcript definitions. UTR definitions were taken from the Saccharomyces Genome Database and a standard region of 100 bp was used where a gene UTR was not defined. A minimum length of 30 bp was implemented to ensure appropriate mapping around the start and stop codons. For the mapping, only unique alignments to transcripts were retained. A full set of 6692 CDSs were established for R64-1-1 Ensembl release 84 and extended by the same UTR sequences defined above. The filtered reads were then mapped to this transcriptome with bowtie2 (Langmead and Salzberg, 2012), using default parameters. For all downstream analysis, dubious ORFs were filtered to leave 5929 transcripts. The A/P-site position of each read was predicted by riboWaltz (Lauria et al., 2018) and aggregated over all transcripts. Data for plots were produced by counting the number of predicted P-sites covering each codon genome-wide.

RNA extraction and analysis. Input and immunoprecipitated RNAs were cleaned from contaminating DNA and purified with the Nucleospin RNAII kit (Macherey Nagel) according to the manufacturer's instructions. For quantitative PCR (RIP-qPCR), input and immunoprecipitated mRNAs (1 μg) were reverse-transcribed using random hexamers (P(dN)₆, Roche) and Superscript II reverse transcriptase (ThermoFisher Scientific). The cDNAs were quantified with a LightCycler 480 system (Roche) according to the manufacturer's instructions. Each sample was systematically analyzed for control RNAs (*i.e.* *ACT1* mRNA or mitochondrial rRNA). The sequences of the primers used for qPCR in this study are listed in **Table S3**. For microarray analysis (RIP-chip), input and immunoprecipitated mRNAs (1 μg) were reverse-transcribed, labeled and hybridized to Agilent arrays covering the entire yeast ORFeome, as previously reported (Bretes et al., 2014). Comparison of total and immunoprecipitated RNAs was performed twice using independent samples and dye swap.

Protein extraction and Western blot analysis. Total protein extraction from yeast cells was performed by the NaOH-TCA lysis method (Ulrich and Davies, 2009). Whole cell extracts and

immunoprecipitated samples were separated on 10% or 4–12% precast SDS-PAGE gels (ThermoFisher Scientific). Proteins were further detected either by silver staining of the gel or by western-blot following transfer to PVDF membranes. The following validated antibodies were used: monoclonal anti-GFP (Sigma), 1:500; rabbit IgG-HRP polyclonal antibody (to detect protein-A-tagged proteins, Sigma), 1:5,000; anti-Dpm1 (ThermoFisher Scientific), 1:2,000. Specificity of anti-GFP antibodies was confirmed using untagged strains. Images were acquired using a ChemiDoc MP Imaging System (BIORAD).

Cell imaging. Localization of tagged fluorescent proteins was analyzed in live cells using freshly grown cultures ($0.5 \leq OD_{600} \leq 1$). Wide-field fluorescence images of tagged versions of Nup1, Kap60, Nic96, Nup49 and Ulp1 were acquired using a Leica DM6000B microscope with a 100X/1.4 NA (HCX Plan-Apo) oil immersion objective and a CCD camera (CoolSNAP HQ; Photometrics). Esc1-GFP variants were imaged using a wide-field microscopy system based on an inverted microscope (TE2000; Nikon) equipped with a 100X/1.4 NA immersion objective, a CMOS camera (ORCA-Flash C11440; Hamamatsu), a collimated white light-emitting diode for the transmission, and a Spectra X light engine lamp (Lumencor, Inc) to illuminate the samples. Both systems were piloted by the MetaMorph software (Molecular Devices). For all images, z-stacks sections of $0.2 \mu\text{m}$ were acquired using a piezo-electric motor (LVDT; Physik Instrument) mounted underneath the objective lens. Images were scaled equivalently and when indicated, z-projected using ImageJ.

Single molecule fluorescence *in situ* hybridization (smFISH) was performed on fixed cells using Stellaris Custom Probe Sets specific for *NUP1*, *NUP2*, *NSP1* or *GFP* mRNAs (listed in **Table S3**) and RNA FISH buffers, according to the manufacturer's instructions (Biosearch Technologies). *NUP2*, *NSP1* and *GFP* probe sets were labeled with Quasar 570 fluorophores while *NUP1* probe set was labeled with Quasar 670. Images were acquired using a Leica DM6000B microscope as above, using Cy3 and Cy5 filters for detecting Quasar 570 and Quasar 670, respectively. Colocalization with GFP-labeled NPC clusters was checked on individual z-sections.

QUANTIFICATION AND STATISTICAL ANALYSIS

For microarray data analysis, the averaged \log_2 of the immunopurified/input ratios and the standard deviation between the two replicates were calculated for each gene. p-values were defined using the LIMMA package (Ritchie et al., 2015). The mRNAs showing an enrichment in a mock immunoprecipitate (obtained similarly from an untagged strain) associated to a p-value of less than 0.05 were removed from the dataset. The 18 most enriched nuclear mRNAs in the mock immunoprecipitate after filtration were used as a mock dataset (**Fig. 3F**; **Fig. S3G**). Gene Ontology and CDS lengths were obtained from Saccharomyces Genome Database (www.yeastgenome.org). Hek2 target mRNAs were previously identified (Wolf et al., 2010). Ribosome footprinting data (Panasenکو et al., 2019) were previously deposited in the SRA database under accession number PRJNA512900.

The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. For statistics, (n) values correspond to the number of biological replicates (i.e. independent yeast cultures) for immunoprecipitations (RIP, RIP-chip and label-free proteomics), to the total number of mRNA spots counted in at least 2 replicates for FISH experiments, or to the total number of cells counted in at least 2 replicates for live imaging experiments, as indicated in the corresponding figure legends. Individual data points are represented, together with mean values and error bars corresponding to standard deviations. The Mann-Whitney-Wilcoxon test was used to compare RNA binding in RIP experiments (**Fig. 1B-C, 2E-F, 3B-C-E, S3F, S3I**) or CDS length between NPC-bound and unbound mRNAs (**Fig. 3F**). The Fisher exact test was used to compare mRNAs distribution in FISH analyses (**Fig. 2C, 4C, S2H-I, S4A**), Hek2 binding between NPC-bound and unbound mRNAs (**Fig. S3G**), and Nup1 foci formation between control and mutant cells (**Fig. 4E, S4B**). Statistical analyses were performed using Prism (Graphpad) and standard conventions for symbols indicating statistical significance were used: *P<0.05; **P<0.01; ***P<0.001; ns, not significant.

Supplemental table not included in main supplemental PDF :

Table S1. NPC mRNA enrichments in Nup immunoprecipitates, Quantitative proteomics analysis of Nup59, Nup82 and Nup116 immunoprecipitates, and NPC-associated mRNAs scored in RIP-chip experiments. **Related to Figs. 1 and 2.**

References

- Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., et al. (2007). The molecular architecture of the nuclear pore complex. *Nature* 450, 695–701.
- Allegretti, M., Zimmerli, C.E., Rantos, V., Wilfling, F., Ronchi, P., Fung, H.K.H., Lee, C.-W., Hagen, W., Turoňová, B., Karius, K., et al. (2020). In-cell architecture of the nuclear pore and snapshots of its turnover. *Nature* 586, 796-800.
- Beck, M., and Hurt, E. (2017). The nuclear pore complex: understanding its function through structural insight. *Nat Rev Mol Cell Biol* 18, 73–89.
- Blobel, G., and Sabatini, D. (1971). Dissociation of mammalian polyribosomes into subunits by puromycin. *Proc. Natl. Acad. Sci. U. S. A.* 68, 390–394.
- Bonnet, A., and Palancade, B. (2014). Regulation of mRNA trafficking by nuclear pore complexes. *Genes Basel* 5, 767–791.
- Bretes, H., Rouviere, J.O., Leger, T., Oeffinger, M., Devaux, F., Doye, V., and Palancade, B. (2014). Sumoylation of the THO complex regulates the biogenesis of a subset of mRNPs. *Nucleic Acids Res* 42,

5043-58.

Casolari, J.M., Thompson, M.A., Salzman, J., Champion, L.M., Moerner, W.E., and Brown, P.O. (2012). Widespread mRNA association with cytoskeletal motor proteins and identification and dynamics of myosin-associated mRNAs in *S. cerevisiae*. *PLoS One* 7, e31912.

Chadrin, A., Hess, B., San Roman, M., Gatti, X., Lombard, B., Loew, D., Barral, Y., Palancade, B., and Doye, V. (2010). Pom33, a novel transmembrane nucleoporin required for proper nuclear pore complex distribution. *J Cell Biol* 189, 795–811.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 10, 1794–1805.

Dilworth, D.J., Suprpto, A., Padovan, J.C., Chait, B.T., Wozniak, R.W., Rout, M.P., and Aitchison, J.D. (2001). Nup2p dynamically associates with the distal regions of the yeast nuclear pore complex. *J Cell Biol* 153, 1465–1478.

Doye, V., Wepf, R., and Hurt, E.C. (1994). A novel nuclear pore protein Nup133p with distinct roles in poly(A)+ RNA transport and nuclear pore distribution. *EMBO J.* 13, 6062–6075.

Duncan, C.D.S., and Mata, J. (2011). Widespread cotranslational formation of protein complexes. *PLoS Genet.* 7, e1002398.

Fischer, J., Teimer, R., Amlacher, S., Kunze, R., and Hurt, E. (2015). Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. *Nat. Struct. Mol. Biol.* 22, 774–781.

Fleischer, T.C., Weaver, C.M., McAfee, K.J., Jennings, J.L., and Link, A.J. (2006). Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev.* 20, 1294–1307.

Halbach, A., Zhang, H., Wengi, A., Jablonska, Z., Gruber, I.M., Halbeisen, R.E., Dehe, P.M., Kemmeren, P., Holstege, F., Geli, V., et al. (2009). Cotranslational assembly of the yeast SET1C histone methyltransferase complex. *EMBO J* 28, 2959–2970.

Hampoelz, B., Schwarz, A., Ronchi, P., Bragulat-Teixidor, H., Tischer, C., Gaspar, I., Ephrussi, A., Schwab, Y., and Beck, M. (2019). Nuclear Pores Assemble from Nucleoporin Condensates During Oogenesis. *Cell* 179, 671-686.e17.

Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Mol. Cell* 31, 925–932.

Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D., and Brown, P.O. (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol* 6, e255.

Jäkel, S., Mingot, J.-M., Schwarzmaier, P., Hartmann, E., and Görlich, D. (2002). Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO J.* 21, 377–386.

Kamenova, I., Mukherjee, P., Conic, S., Mueller, F., El-Saafin, F., Bardot, P., Garnier, J.-M., Dembele, D., Capponi, S., Timmers, H.T.M., et al. (2019). Co-translational assembly of mammalian nuclear multisubunit complexes. *Nat. Commun.* 10, 1740.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* *12*, 357–360.

Kim, S.J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B.D., Hogan, J.A., Upla, P., et al. (2018). Integrative structure and functional anatomy of a nuclear pore complex. *Nature* *555*, 475–482.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* *9*, 357–359.

Lauria, F., Tebaldi, T., Bernabò, P., Groen, E.J.N., Gillingwater, T.H., and Viero, G. (2018). riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput. Biol.* *14*, e1006169.

Lazar, C., Gatto, L., Ferro, M., Bruley, C., and Burger, T. (2016). Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. *J. Proteome Res.* *15*, 1116–1125.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N., et al. (2016). Architecture of the symmetric core of the nuclear pore. *Science* *352*, aaf1015.

Long, R.M., Singer, R.H., Meng, X., Gonzalez, I., Nasmyth, K., and Jansen, R.P. (1997). Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* *277*, 383–387.

Martin, M. (2011). Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet J.* *17*, 10–12.

Mészáros, N., Cibulka, J., Mendiburo, M.J., Romanauska, A., Schneider, M., and Köhler, A. (2015). Nuclear pore basket proteins are tethered to the nuclear envelope and can regulate membrane curvature. *Dev. Cell* *33*, 285–298.

Natan, E., Wells, J.N., Teichmann, S.A., and Marsh, J.A. (2017). Regulation, evolution and consequences of cotranslational protein complex assembly. *Curr Opin Struct Biol* *42*, 90–97.

Onischenko, E., Tang, J.H., Andersen, K.R., Knockenhauer, K.E., Vallotton, P., Derrer, C.P., Kralt, A., Mugler, C.F., Chan, L.Y., Schwartz, T.U., et al. (2017). Natively Unfolded FG Repeats Stabilize the Structure of the Nuclear Pore Complex. *Cell* *171*, 904-917.e19.

Padavannil, A., Sarkar, P., Kim, S.J., Cagatay, T., Jiou, J., Brautigam, C.A., Tomchick, D.R., Sali, A., D’Arcy, S., and Chook, Y.M. (2019). Importin-9 wraps around the H2A-H2B core to act as nuclear importer and histone chaperone. *ELife* *8*, e43630.

Panasenko, O.O., Somasekharan, S.P., Villanyi, Z., Zagatti, M., Bezrukov, F., Rashpa, R., Cornut, J., Iqbal, J., Longis, M., Carl, S.H., et al. (2019). Co-translational assembly of proteasome subunits in NOT1-containing assembliesomes. *Nat. Struct. Mol. Biol.* *26*, 110–120.

Panse, V.G., Kuster, B., Gerstberger, T., and Hurt, E. (2003). Unconventional tethering of Ulp1 to the transport channel of the nuclear pore complex by karyopherins. *Nat Cell Biol* *5*, 21–27.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* *43*, e47.

Rout, M.P., Aitchison, J.D., Suprpto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. (2000). The yeast nuclear

pore complex: composition, architecture, and transport mechanism. *J Cell Biol* *148*, 635–651.

Rouviere, J.O., Bulfoni, M., Tuck, A., Cosson, B., Devaux, F., and Palancade, B. (2018). A SUMO-dependent feedback loop senses and controls the biogenesis of nuclear pore subunits. *Nat. Commun.* *9*, 1665.

Ryan, K.J., Zhou, Y., and Wente, S.R. (2007). The karyopherin Kap95 regulates nuclear pore complex assembly into intact nuclear envelopes in vivo. *Mol. Biol. Cell* *18*, 886–898.

Santos, D.A., Shi, L., Tu, B.P., and Weissman, J.S. (2019). Cycloheximide can distort measurements of mRNA levels and translation efficiency. *Nucleic Acids Res.* *47*, 4974–4985.

Schlenstedt, G., Smirnova, E., Deane, R., Solsbacher, J., Kutay, U., Görlich, D., Ponstingl, H., and Bischoff, F.R. (1997). Yrb4p, a yeast ran-GTP-binding protein involved in import of ribosomal protein L25 into the nucleus. *EMBO J.* *16*, 6237–6249.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* *473*, 337–342.

Schwarz, A., and Beck, M. (2019). The Benefits of Cotranslational Assembly: A Structural Perspective. *Trends Cell Biol.* *29*, 791–803.

Shiber, A., Döring, K., Friedrich, U., Klann, K., Merker, D., Zedan, M., Tippmann, F., Kramer, G., and Bukau, B. (2018). Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* *561*, 268–272.

Shieh, Y.-W., Minguéz, P., Bork, P., Auburger, J.J., Guilbride, D.L., Kramer, G., and Bukau, B. (2015). Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science* *350*, 678–680.

Singer-Krüger, B., and Jansen, R.-P. (2014). Here, there, everywhere. mRNA localization in budding yeast. *RNA Biol.* *11*, 1031–1039.

Tyanova, S., and Cox, J. (2018). Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. *Methods Mol. Biol.* *1711*, 133–148.

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* *11*, 2301–2319.

Ulrich, H.D., and Davies, A.A. (2009). In vivo detection and characterization of sumoylation targets in *Saccharomyces cerevisiae*. *Methods Mol Biol* *497*, 81–103.

Walther, T.C., Askjaer, P., Gentzel, M., Habermann, A., Griffiths, G., Wilm, M., Mattaj, I.W., and Hetzer, M. (2003). RanGTP mediates nuclear pore complex assembly. *Nature* *424*, 689–694.

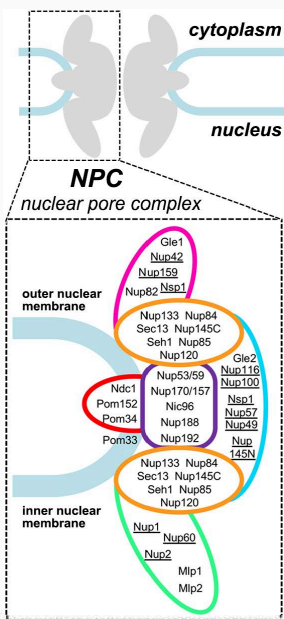
Weber, M., and Antonin, W. (2016). Perforating the nuclear boundary - how nuclear pore complexes assemble. *J Cell Sci* *129*, 4439–4447.

Wolf, J.J., Dowell, R.D., Mahony, S., Rabani, M., Gifford, D.K., and Fink, G.R. (2010). Feed-forward regulation of a cell fate determinant by an RNA-binding protein generates asymmetry in yeast. *Genetics* *185*, 513–522.

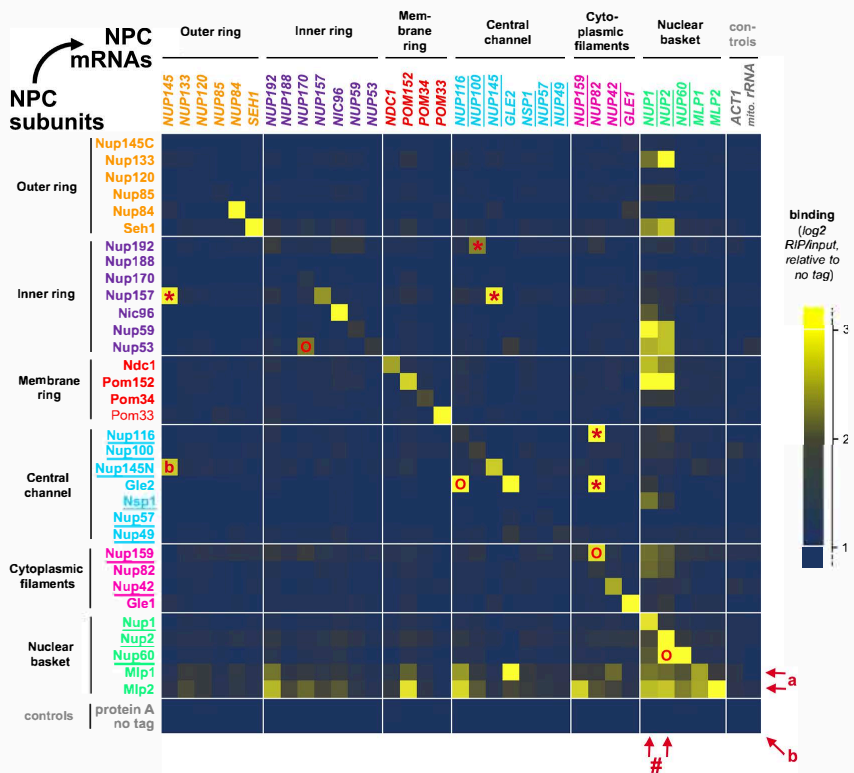
Yoshida, K., Seo, H.-S., Debler, E.W., Blobel, G., and Hoelz, A. (2011). Structural and functional analysis of an essential nucleoporin heterotrimer on the cytoplasmic face of the nuclear pore complex. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 16571–16576.

Figure 1

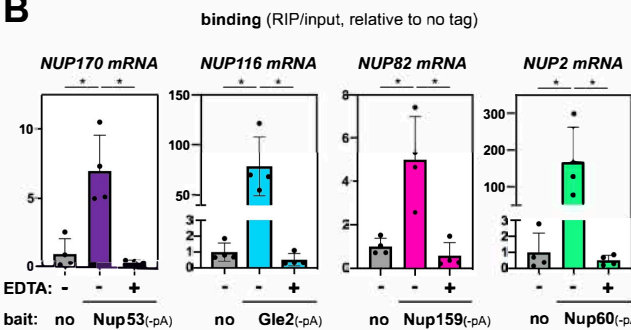
A



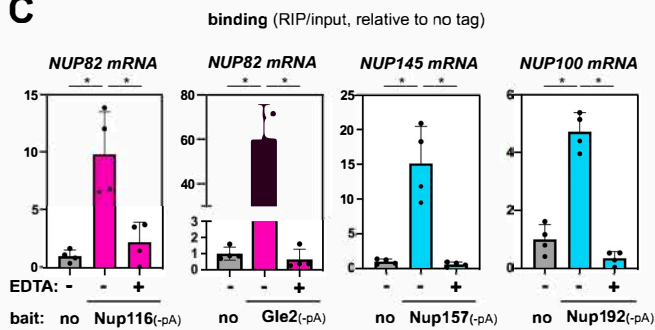
- intra-module co-translational interaction
- * inter-module co-translational interaction
- # mRNAs at NPCs
- a mRNA docking at Mlp1-Mlp2 (export)
- b interaction with own mRNA (translation)



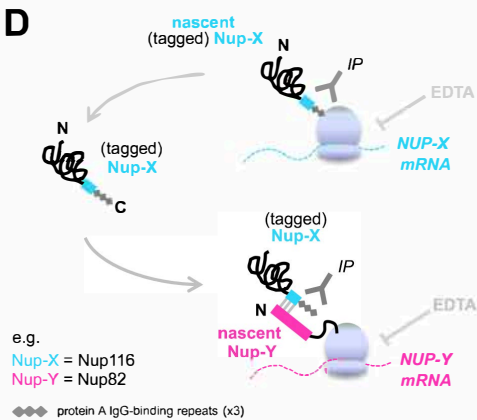
B



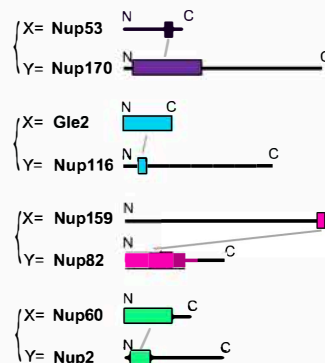
C



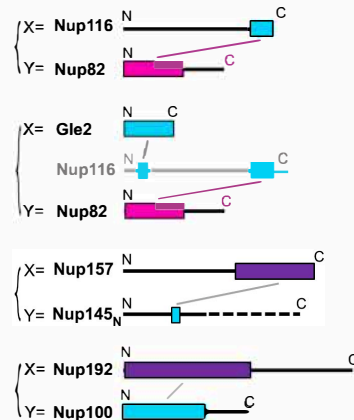
D



E

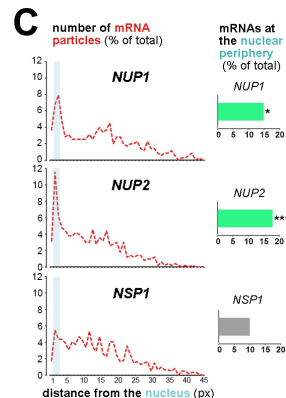
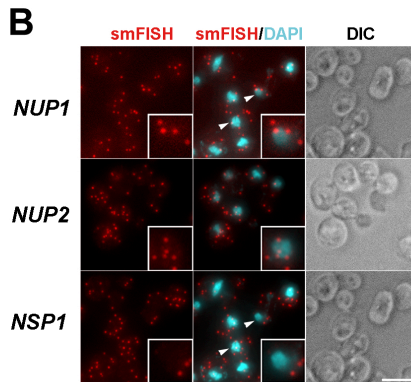
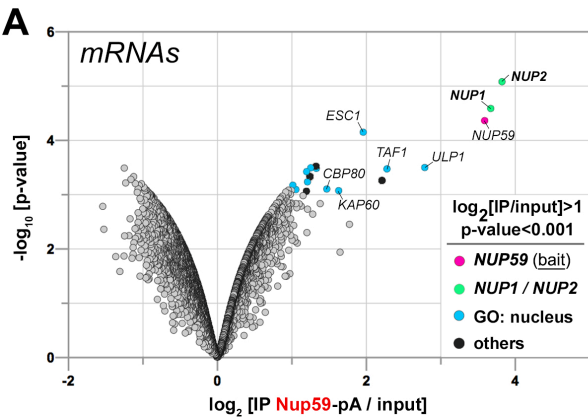


F

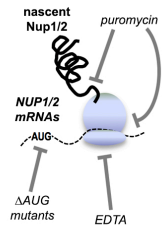


100 aa —

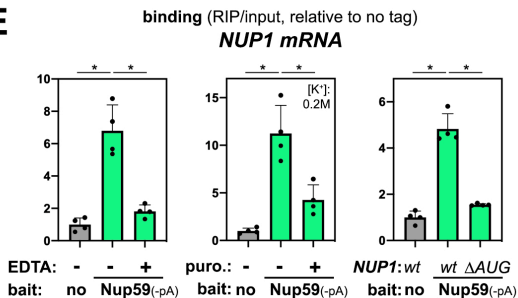
Figure 2



D



E



F

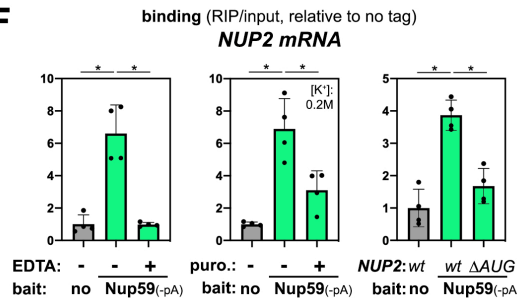
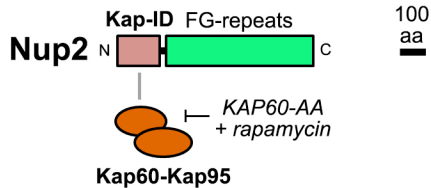
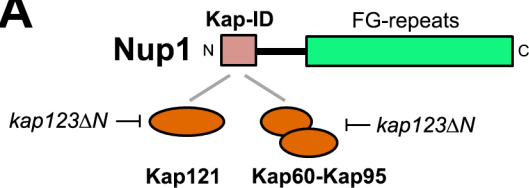
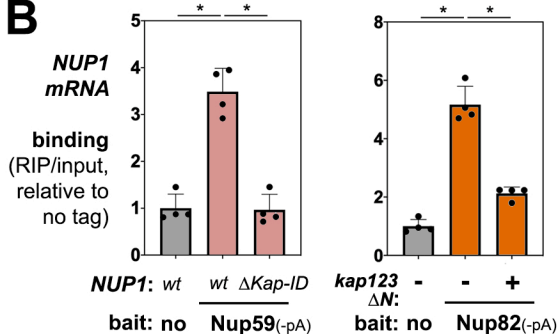


Figure 3

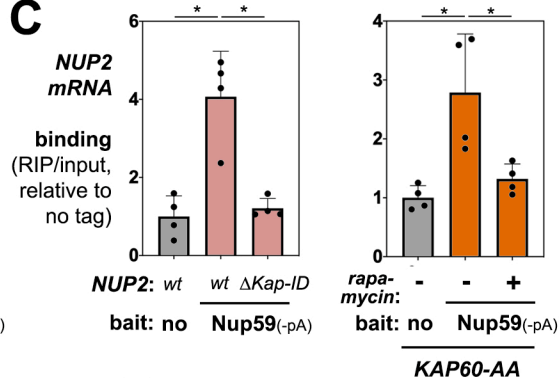
A



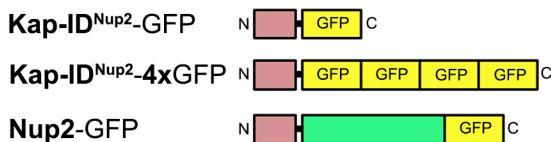
B



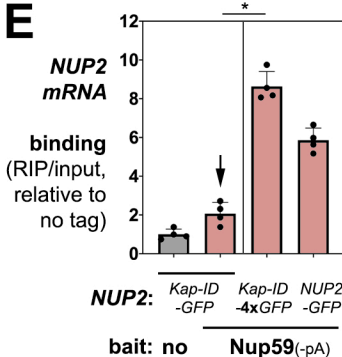
C



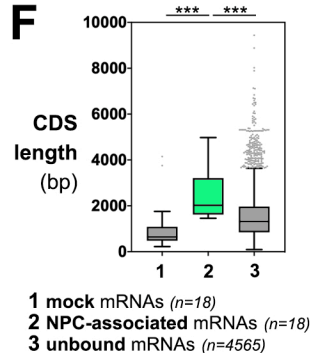
D



E



F



G

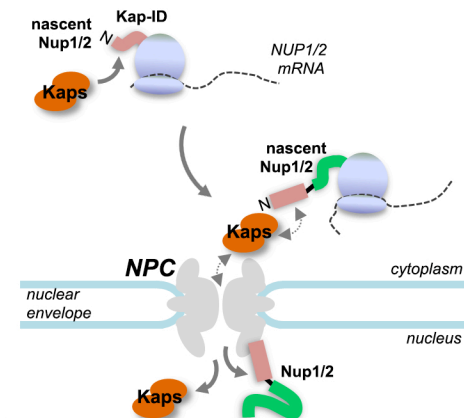
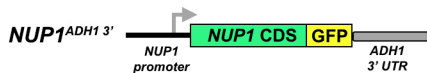
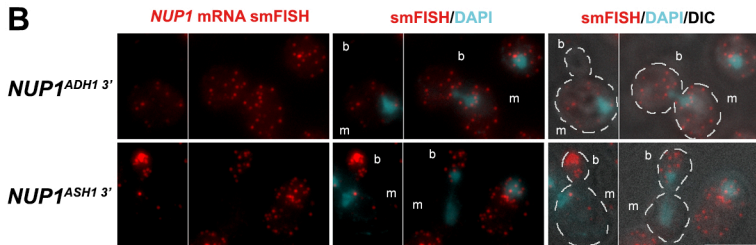


Figure 4

A

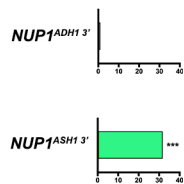


B

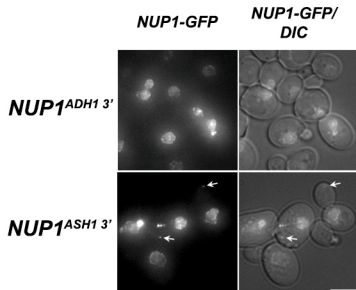


C

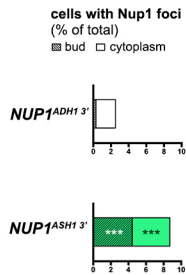
cells with *NUP1* mRNAs mislocalized to the bud (% of total)



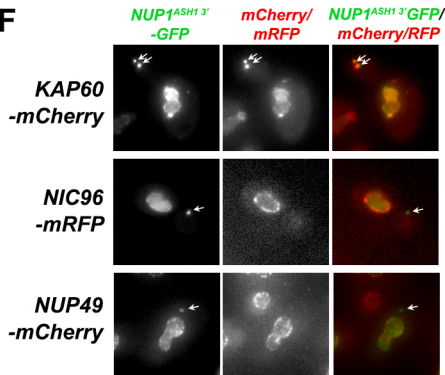
D



E



F



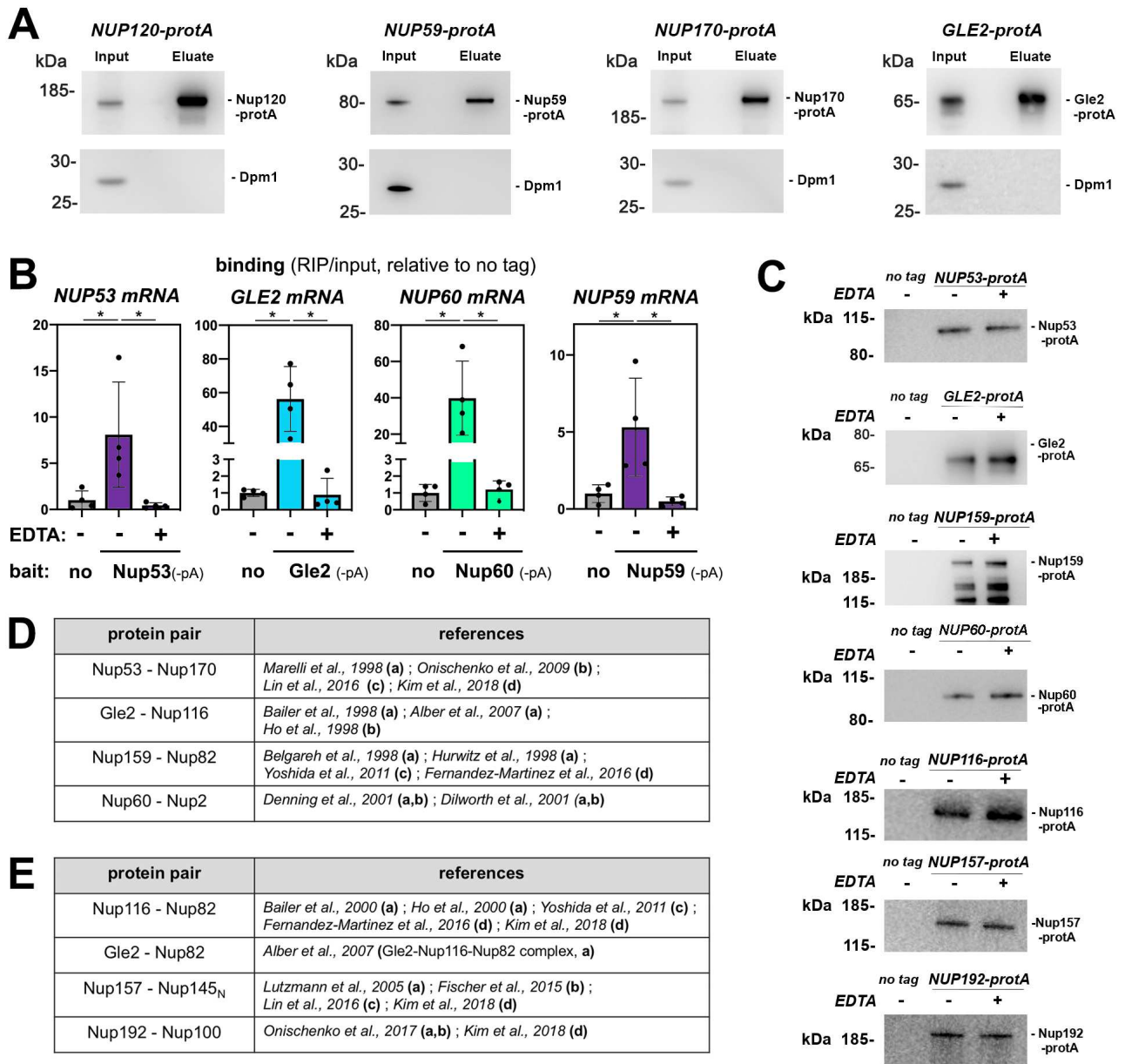
SUPPLEMENTAL INFORMATION

Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis

Ophélie Lautier, Arianna Penzo, Jérôme O. Rouvière, Guillaume Chevreux, Louis Collet, Isabelle Loïodice, Angela Taddei, Frédéric Devaux, Martine A. Collart & Benoit Palancade

includes:

- **Figures S1-4.**
- **Table S2** (Strains used in this study).
- **Table S3** (qPCR primers and FISH probes used in this study).
- **Table S4** (Plasmids used in this study).
- **Supplemental references.**



(a) complex identification - (b) biochemical reconstitution - (c) structural analysis - (d) *in vivo* crosslink

Figure S1, related to Fig. 1. Control experiments related to the co-translational interaction screen.

A, For each of the protA-tagged strains, soluble extracts (“input”) and immunopurified proteins (“eluate”, 10X equivalent) were analyzed by SDS-PAGE and western-blotting using rabbit IgG-HRP antibodies (to detect protA-tagged Nups, *top panel*) or anti-Dpm1 antibodies (to detects an unrelated, control ER protein). The position of molecular weights is represented (kDa). Note that the bait is systematically enriched over the control protein. **B**, The association between the indicated pA-tagged Nups (baits) and their own mRNAs was analyzed by RIP followed by RT-qPCR and represented as the ratios between immunopurified and input RNAs (mean±SD, n=4, relative to untagged). When indicated, the purification was performed in the presence of 40 mM EDTA. *, p<0.05 (Mann-Whitney-Wilcoxon test). **C**, The protein eluates obtained upon immunopurification of the indicated pA-tagged Nups (as in **Fig. 1B-C**) were analyzed by SDS-PAGE and western-blotting using rabbit IgG-HRP antibodies. When indicated, the experiment was performed in the presence of 40 mM EDTA. The position of molecular weights is represented (kDa). Note that equal purification of bait proteins is achieved in both untreated and EDTA-treated samples. **D-E**, The physical interactions and the corresponding domains in the pairs of proteins found in the co-translational interaction screen (depicted in **Fig. 1B-F**) were previously described in the indicated studies (Alber et al., 2007; Bailer et al., 1998, 2000; Belgareh et al., 1998; Denning et al., 2001; Dilworth et al., 2001; Fernandez-Martinez et al., 2016; Fischer et al., 2015; Ho et al., 1998, 2000; Hurwitz et al., 1998; Kim et al., 2018; Lin et al., 2016; Lutzmann et al., 2005; Marelli et al., 1998; Onischenko et al., 2009, 2017; Yoshida et al., 2011). The type of experimental evidences is indicated (a-d). Note that the interaction domains within the paralogous proteins Nup157 and Nup170 were mapped within their unique orthologue in the fungus *Chaetomium thermophilum* (Lin et al., 2016). Structural analysis of Gle2 binding to Gle2-binding sequence (GLEBS) was achieved for the human Gle2^{Rae1}-Nup98 complex (Ren et al., 2010).

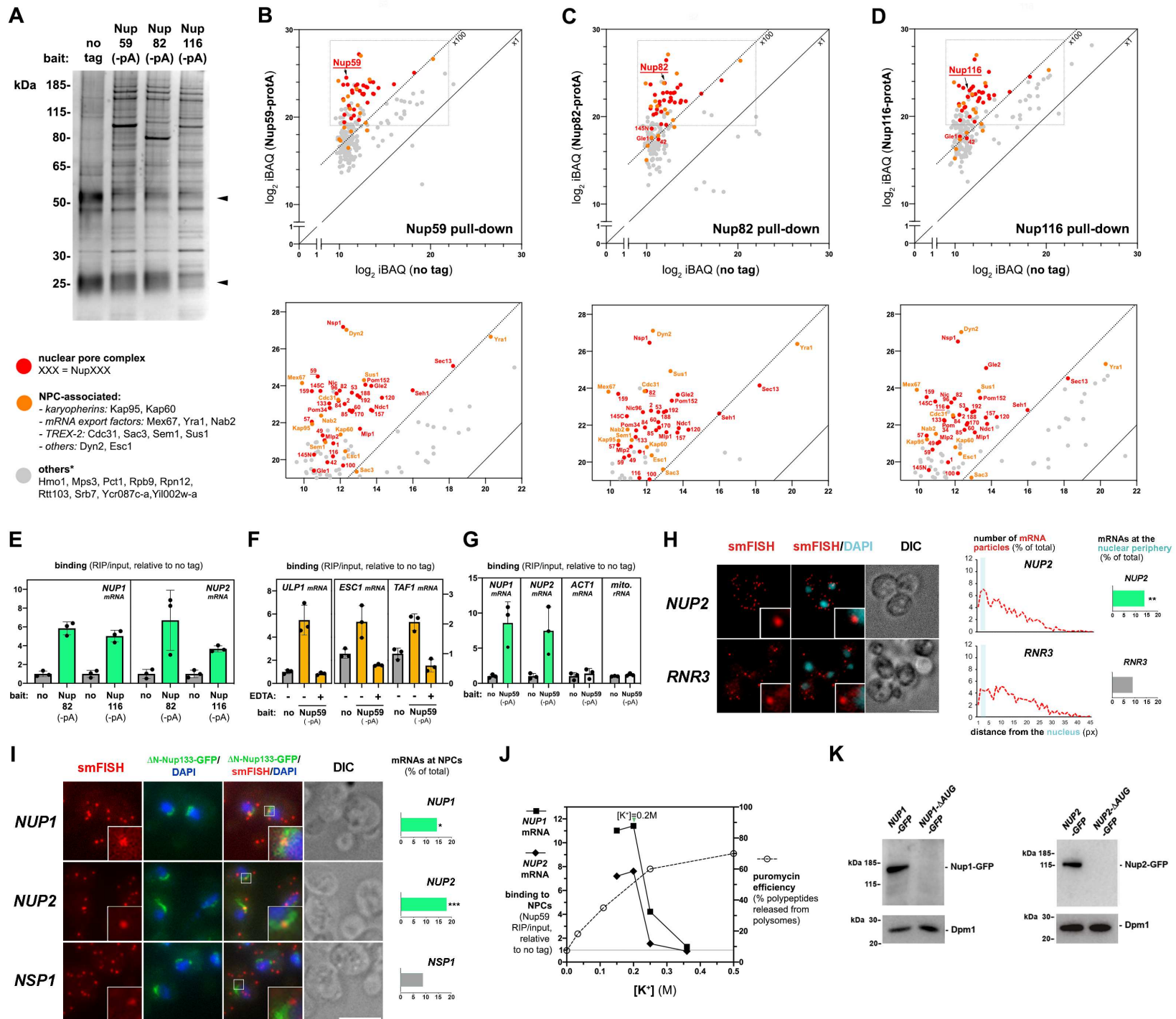


Figure S2. (see legend on next page)

Figure S2, related to Fig. 2. Identification and characterization of NPC-associated mRNAs

A, Immunoprecipitates from control (no tag), *NUP59-protA*, *NUP82-protA* and *NUP116-protA* cells were analyzed by SDS-PAGE followed by silver staining. Arrowheads indicate the heavy and light chains of the IgG used in the pull-down. The position of molecular weights is indicated (kDa). Note the specific, similar banded pattern observed in the three Nup immunoprecipitates as compared to the control sample. **B-D**, Protein abundances (mean log₂ iBAQ ; n=3) were determined by label free mass spectrometry for control (no tag) and Nup59-pA (**B**), Nup82-pA (**C**) or Nup116-pA (**D**) immunoprecipitates and plotted for each polypeptide significantly enriched in one of the samples (p<0.01; see also **Table S1** for the complete dataset). NPC subunits and NPC-associated proteins are highlighted in red and orange, respectively. The diagonals indicate 1X and 100X enrichments in the tagged sample, and the bottom panels represent an enlargement of the areas of interest in the top panels. *, other proteins commonly enriched by more than 100X in the three pull-downs and appearing in the bottom panels are listed. **E-G**, The association between the indicated mRNAs and NPCs was analyzed by RIP followed by RT-qPCR and represented as the ratios between immunopurified and input RNAs (mean±SD, n=3, relative to untagged). Protein A tagged versions of Nup82 (**E**), Nup116 (**E**) or Nup59 (**F-G**) were used as baits in the immunopurification procedure. When indicated, the purification was performed in the presence of 40 mM EDTA. Note that NPC isolates are enriched in *NUP1* and *NUP2* mRNAs, but not in highly-expressed *ACT1* mRNAs or mitochondrial (mito.) rRNAs. **H**, Single-molecule FISH was performed on *RNR3-GFP* cells using sets of probes specific for *NUP2* or *GFP* mRNAs. *Left panels*, smFISH images, as well as merged images with the DAPI channel (nuclear staining) are shown. Z-projections are displayed, except for Differential Interference Contrast (DIC) images. Scale bar, 5 μm. *Middle panels*, The distribution of mRNA particles in the smFISH experiments (% of total) is represented according to their distance to the nuclear staining (pixels, px). *Right panels*, Plotted is the fraction of particles localized in the perinuclear area (1px < distance < 3px=200 nm, highlighted in blue). The number of counted mRNA particles and cells (n₁/n₂) is as follows: *NUP2*: 637/101; *RNR3-GFP*: 655/66 (from two independent experiments for *NUP2*, three for *GFP*). **, p<0.01 (Fisher exact test). **I**, smFISH was performed on GFP-Δ*N-nup133* cells using sets of probes specific for *NUP1*, *NUP2* or *NSP1* mRNAs. SmFISH images, as well as merged images from the GFP and DAPI (nuclear staining) channels are shown. Z-projections are displayed, except for Differential Interference Contrast (DIC) images. The fraction of mRNA particles colocalized with GFP-tagged NPC clusters is represented (*right panels*). The total number of counted mRNA particles and cells (n₁/n₂) from two independent experiments is as follows: *NUP1*, 318/52; *NUP2*, 583/69; *NSP1*, 626/52. Scale bar, 5 μm. *, p<0.05; ***, p<0.001 (Fisher exact test). **J**, Optimization of the RNA immunoprecipitation assay in conditions of *in vitro* puromycin treatment. Since puromycin-mediated release of nascent polypeptides is only effective at high [K⁺] concentration (data from Blobel and Sabatini, 1971; right Y-axis), *NUP1* and *NUP2* binding to NPCs (Nup59-pA) was analyzed by RIP at increasing [K⁺] concentrations (in the absence of puromycin) and represented as the ratios between immunopurified and input RNAs (relative to untagged and normalized to *ACT1* mRNA, left Y-axis). The maximal [K⁺] concentration preserving *NUP1* and *NUP2* mRNA interactions with NPCs (= 0.2 M) was chosen to perform the *in vitro* puromycin dissociation experiments featured in **Fig. 2D-F**. **K**, Whole cell extracts from *wt* cells expressing GFP-tagged versions of *NUP1* and *NUP2*, either *wt* or Δ*AUG*, were analyzed by SDS-PAGE and western blotting using anti-GFP and anti-Dpm1 antibodies. Dpm1 was used as a loading control. The position of molecular weights is indicated (kDa). Note that Nup1 and Nup2 proteins are not detected upon deletion of their respective start codons. Transgene expression was achieved in the presence of the endogenous *wt* proteins.

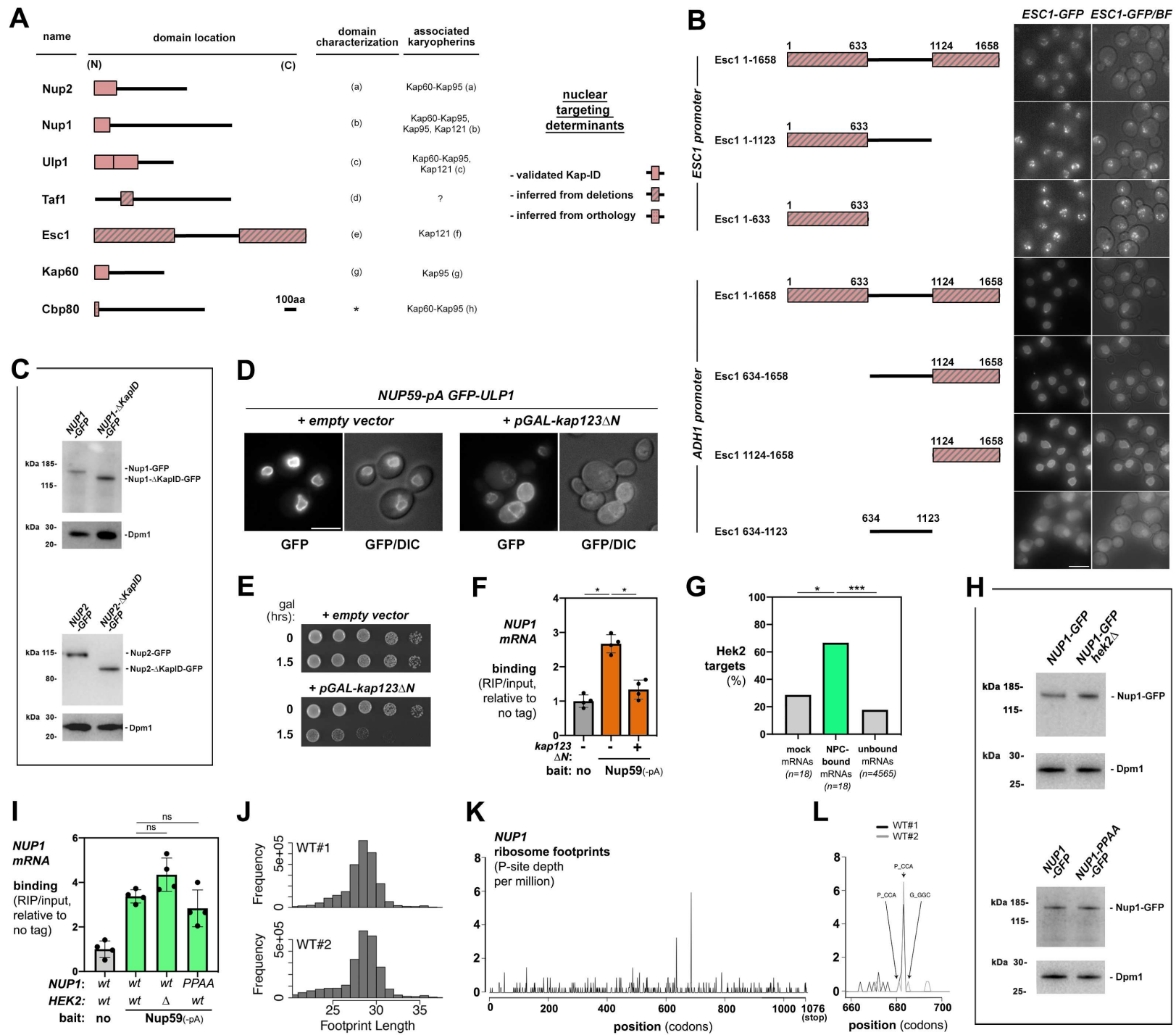


Figure S3. (see legend on next page)

Figure S3, related to Fig. 3. Mechanisms of targeting of NPC-bound mRNAs

A, Domain organization of proteins encoded by the most enriched NPC-associated mRNAs (\log_2 IP/input > 1.5; $p < 0.001$). Represented are sequences required for nuclear targeting (*pink boxes*), including validated karyopherin interaction domains (Kap-ID; *plain boxes*) and other nuclear localization determinants inferred from deletion studies (*shaded boxes*) or orthology (*dotted box*, *). When identified, the interacting Kaps are indicated. (a) (Solsbacher et al., 2000); (b) (Mészáros et al., 2015); (c) (Panse et al., 2003); (d) (Irvin and Pugh, 2006); (e) this study, see panel **B**; (f) (Niepel et al., 2013); (g) (Enekel et al., 1996); (h) (Oeffinger et al., 2007). **B**, Live cell imaging of GFP-tagged Esc1 truncation mutants. A scheme of the deletions is represented, together with z-projections for the GFP channel, as such, or merged with the bright-field (BF) images. N-terminal truncations were expressed from the *ADHI* promoter. Scale bar, 5 μ m. Note that Esc1 N-terminal region (aa 1-633) is sufficient for nuclear targeting, even if another autonomous nuclear localization determinant is present in the C-terminus of the protein. **C**, Whole cell extracts from *wt* cells expressing GFP-tagged versions of *NUP1* and *NUP2*, either *wt* or Δ *Kap-ID*, were analyzed by SDS-PAGE and western blotting using anti-GFP and anti-Dpm1 antibodies. Dpm1 was used as a loading control. The position of molecular weights is indicated (kDa). In view of Nup1 Kap-ID essentiality (Mészáros et al., 2015), transgene expression was achieved in the presence of the endogenous *wt* proteins. Note that these N-terminal truncations do not affect the total levels of Nup1 and Nup2 (-GFP) proteins. **D**, Live imaging of *NUP59-protA* cells expressing *GFP-ULP1* and either an empty vector, or the *GAL-kap123 Δ N* construct, following 1.5 hrs galactose induction. Single plane images are shown for the GFP and DIC channels. Scale bar, 5 μ m. **E**, Following the indicated time of galactose induction, the same cells as in **D** were spotted as serial dilutions on glucose-containing SC medium. Note that the cells having experienced 1.5 hrs of galactose induction still support colony formation, indicating that the mutant did not cause an irreversible and complete growth arrest. **F**, The association between *NUP1* mRNAs and NPCs (Nup59-pA) was analyzed by RIP followed by RT-qPCR and represented as the ratios between purified and input RNAs (mean \pm SD, n=4, relative to untagged and normalized to *ACT1* mRNA). When indicated, the experiment was performed upon galactose-induced over-expression of *kap123 Δ N* for 1.5 hrs. *, $p < 0.05$ (Mann-Whitney-Wilcoxon test). **G**, The fraction (%) of Hek2 target mRNAs was represented for mock mRNAs (most enriched mRNAs in a mock immunoprecipitate, see STAR Methods), NPCs-enriched mRNAs (from **Fig. 2A**) or unbound mRNAs (\log_2 IP/input < 1 or $p > 0.001$ in **Fig. 2A**). Hek2 target mRNAs were previously identified by Cross-Linking Immunoprecipitation (Wolf et al., 2010). *, $p < 0.05$; ***, $p < 0.001$ (Fisher exact test). **H**, Whole cell extracts from *wt* or *hek2 Δ* cells expressing GFP-tagged versions of *NUP1*, either *wt* or P₆₈₂A, P₆₈₃A (PPAA), were analyzed by SDS-PAGE and western blotting using anti-GFP and anti-Dpm1 antibodies. Dpm1 was used as a loading control. The position of molecular weights is represented (kDa). Note that the indicated mutations do not impact the total levels of Nup1-GFP proteins. **I**, The association between *NUP1* mRNAs and NPCs (Nup59-pA) was analyzed by RIP followed by RT-qPCR in *wt* or *hek2 Δ* mutant cells expressing GFP-tagged versions of *NUP1*, either *wt* or P₆₈₂A, P₆₈₃A (PPAA), and represented as the ratios between immunopurified and input RNAs (mean \pm SD, n=4, relative to untagged and normalized to *ACT1* mRNA). Transgene-derived mRNAs were detected using *GFP*-specific primer pairs. ns, not significant (Mann-Whitney-Wilcoxon test). **J-L**, Ribosome footprinting analysis of *NUP1* mRNA in *wt* cells: histograms of ribosome footprint length in two independent biological replicates (**J**) and distribution of ribosome footprints along *NUP1* CDS (**K**). Note the presence of a strong pause site at the P₆₈₂P₆₈₃ dicodon within *NUP1* CDS (enlarged in **L** for the two replicates).

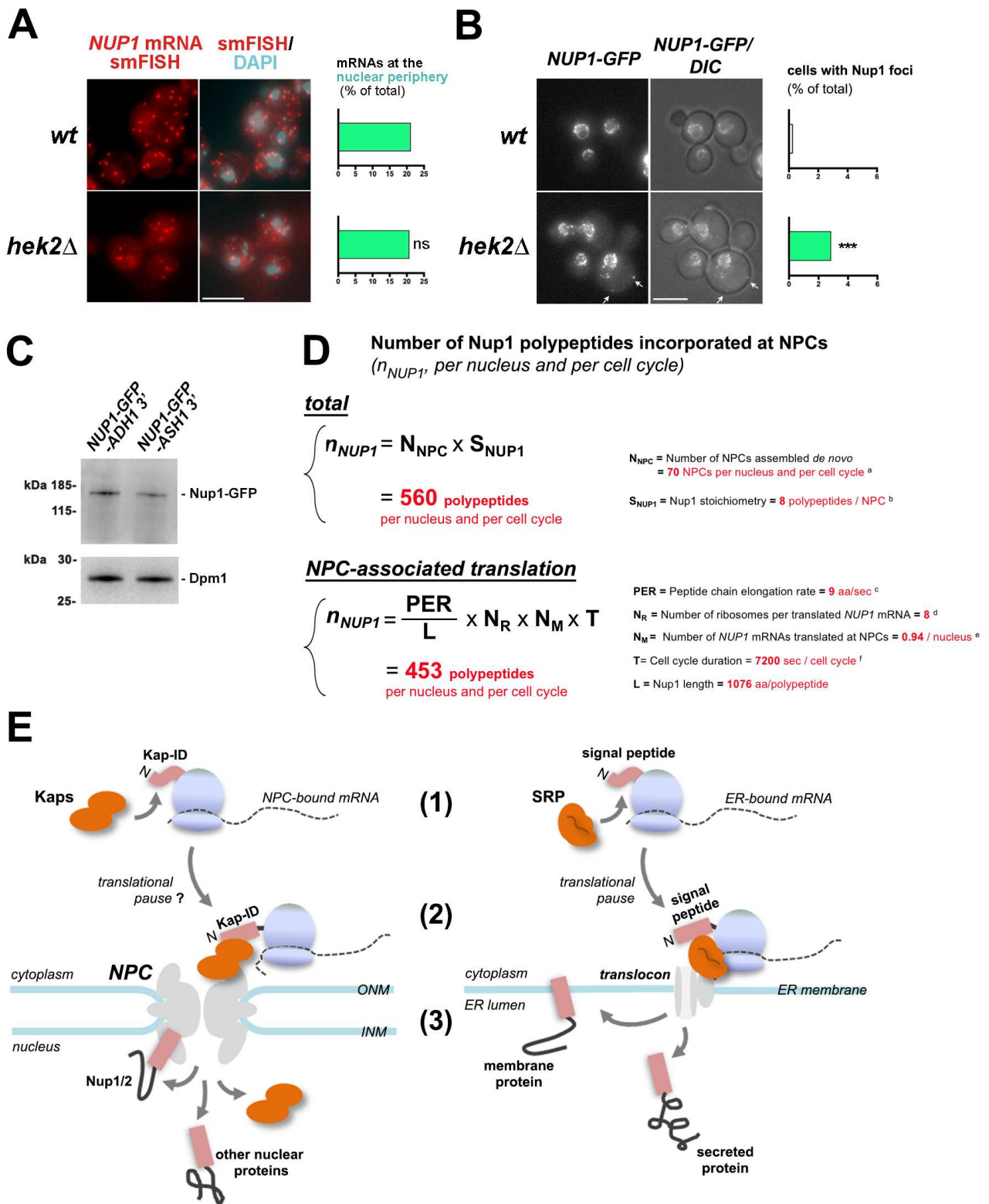


Figure S4. (see legend on next page)

Figure S4, related to Fig. 4. Co-translational interactions and nucleoporin homeostasis.

A, *Left panel*, smFISH was performed on *wt* and *hek2Δ* cells using sets of probes specific for *NUP1* mRNAs. SmFISH images, as well as merged images with the DAPI channel (nuclear staining) are shown. Z-projections are displayed. Scale bar, 5 μm. *Right panel*, Plotted is the fraction of particles localized in the perinuclear area (determined as in **Fig. 2C**). The total number of counted mRNA particles and cells (n_1/n_2) from two independent experiments is as follows: *wt*: 330/63; *hek2Δ*: 358/82. ns, not significant (Fisher exact test). **B**, *Left panel*, live imaging of *wt* or *hek2Δ* mutant cells expressing an integrated, GFP-tagged version of Nup1. Z-projections and single plane images are shown for the GFP and DIC channels, respectively, as well as arrows indicating Nup1-GFP foci. Scale bar, 5 μm. *Right panel*, the fraction of cells exhibiting cytoplasmic foci of Nup1-GFP is represented. The total number of counted cells from three independent experiments is as follows: *wt*, 737; *hek2Δ*, 843. ***, $p < 0.001$ (Fisher exact test). **C**, Whole cell extracts from *wt* cells carrying *NUP1-GFP^{ADHI 3'}* or *NUP1-GFP^{ASHI 3'}* constructs were analyzed by SDS-PAGE and western blotting using anti-GFP and anti-Dpm1 antibodies. Dpm1 was used as a loading control. The position of molecular weights is indicated (kDa). **D**, NPC-associated translation can supply most Nup1 polypeptides required for interphase NPC biogenesis. The total number of Nup1 proteins incorporated (per nucleus and per cell cycle) during interphase NPC biogenesis was calculated as follows (*top panel*): (a) the number of NPCs assembled *de novo* during interphase (N_{NPC}) was previously reported based on three-dimensional reconstruction of electron microscopy images of the yeast nucleus (Winey et al., 1997), (b) the stoichiometry of Nup1 (S_{NUP1}) within native NPCs was determined by quantitative proteomics of isolated yeast NPCs (Kim et al., 2018). The number of Nup1 proteins supplied by NPC-associated translation (*bottom panel*) was estimated based on (c) the peptide chain elongation rate (PER) previously determined for yeast cells grown upon optimal conditions (Boehlke and Friesen, 1975), (d) the average number of ribosomes per translated *NUP1* mRNAs (N_R), according to our earlier polysome profiling data (Rouviere et al., 2018), (e) the average number of *NUP1* mRNAs present at NPCs (N_M , this study, **Fig. 2C**), (f) the duration of the yeast cell cycle (120 min) and the length of the Nup1 polypeptide. Note that while these calculations only provide estimates, they suggest that the numbers of Nup1 polypeptides required for interphase NPC biogenesis or produced by NPC-associated translation are in the same range. **E**, Comparison between translational processes occurring at nuclear pores and at the endoplasmic reticulum (ER). *Left panel*, during early translation of *NUP1*, *NUP2* or other NPC-bound mRNAs, recognition of the nascent karyopherin interaction domain (Kap-ID) by soluble karyopherins (Kaps, (1)) would trigger the association of translating ribosomes with NPCs by virtue of interactions between Kaps and cytoplasmically oriented FG-Nups (2). Fully-translated polypeptides would be further translocated through NPCs as for classical nuclear import (3). Translational pausing could contribute to synchronize translation and NPC targeting, although not essential in our experimental conditions. ONM, outer nuclear membrane, INM, inner nuclear membrane. *Right panel*, during early translation of ER-bound mRNAs, recognition of the signal peptide by the signal recognition particle (SRP, (1)) triggers translational pausing and delivery to the Sec61 translocon complex (2). The polypeptide is then translocated co-translationally (3), allowing its targeting to the ER membrane (for membrane proteins) or to the ER lumen (for ER-resident or secreted proteins; reviewed in Aviram and Schuldiner, 2017). Note that despite this functional convergence, both translocation machineries show independent evolutionary origins, the Sec61 complex only having a prokaryotic counterpart (Devos et al., 2004; Mans et al., 2004; Mandon et al., 2009).

Table S2 - Strains used in this study.

Lab code	Strain name	Genotype	Source	Usage (Figures)
YBP2149	<i>wt (DF5)</i>	<i>MATalpha ura3 his3 trp1 leu2 lys2</i>	<i>(Rout et al., 2000) (a)</i>	
YBP1076	<i>wt (W303)</i>	<i>MATalpha ade2 ura3 his3 trp1 leu2 can1</i>	<i>(Rout et al., 2000) (a)</i>	
YBP539	<i>wt (BY4742)</i>	<i>MATalpha ura3 his3 leu2 lys2</i>	<i>Euroscarf</i>	
YBP2213	<i>wt (BY4741)</i>	<i>MATa ura3 his3 leu2 met15</i>	<i>Euroscarf</i>	2B-C, 4B-E and S4C (transformed with <i>NUP1-GFP-ADH1^{3'UTR}</i> or <i>NUP1-GFP-ASH1^{3'UTR}</i> expression vectors), 4F (transformed with <i>NUP1-GFP-ASH1^{3'UTR}</i> and <i>NUP49-mCherry</i> expression vectors)
YBP937	<i>wt (BY4743)</i>	<i>MATalpha/ura3/ura3 his3/his3 leu2/leu2 lys2/LYS2 met15/MET15</i>	<i>Euroscarf</i>	
YBP2118	<i>MLP1-protA</i>	<i>(W303) MLP1-protA::HIS3-URA3</i>	<i>(Niepel et al., 2005) (a)</i>	1A
YBP2119	<i>MLP2-protA</i>	<i>(W303) MLP2-protA::HIS3-URA3</i>	<i>(Niepel et al., 2005) (a)</i>	1A
YBP2120	<i>NUP192-protA</i>	<i>(DF5) NUP192-protA::HIS3-URA3</i>	<i>(Rout et al., 2000) (a)</i>	1A, 1C, S1C
YBP2121	<i>NUP188-protA</i>	<i>(DF5) NUP188-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2122	<i>NUP170-protA</i>	<i>(DF5) NUP170-protA::HIS3-URA3</i>	<i>(Rout et al., 2000) (a)</i>	1A, S1A
YBP2123	<i>NUP159-protA</i>	<i>(DF5) NUP159-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A, 1B, S1C
YBP2124	<i>NUP157-protA</i>	<i>(DF5) NUP157-protA::HIS3-URA3</i>	<i>(Rout et al., 2000) (a)</i>	1A, 1C, S1C
YBP2125	<i>POM152-protA</i>	<i>(DF5) POM152-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2126	<i>NUP145-protA</i>	<i>(DF5) NUP145-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2127	<i>NUP133-protA</i>	<i>(DF5) NUP133-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2128	<i>NUP120-protA</i>	<i>(DF5) NUP120-protA::HIS3-URA3</i>	<i>(Rout et al., 2000) (a)</i>	1A, S1A
YBP2129	<i>NUP116-protA</i>	<i>(DF5) NUP116-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A, 1C, S1C, S2A, S2D, S2E
YBP2130	<i>NUP1-protA</i>	<i>(DF5) NUP1-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2131	<i>NUP100-protA</i>	<i>(DF5) NUP100-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2132	<i>NIC96-protA</i>	<i>(DF5) NIC96-protA::HIS3-URA3</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2133	<i>NSP1-protA</i>	<i>(DF5) NSP1-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2134	<i>NUP85-protA</i>	<i>(DF5) NUP85-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2135	<i>NUP84-protA</i>	<i>(DF5) NUP84-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2136	<i>NUP82-protA</i>	<i>(W303) NUP82-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A, 3B (transformed with the <i>kap123ΔN</i> plasmid), S2A, S2C, S2E
YBP2137	<i>NUP2-protA</i>	<i>(DF5) NUP2-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2138	<i>NDC1-protA</i>	<i>(DF5) NDC1-protA::HIS3-URA3</i>	<i>(Rout et al., 2000) (a)</i>	1A
YBP2139	<i>GLE1-protA</i>	<i>(DF5) GLE1-protA::HIS5</i>	<i>(Rout et al., 2000) (a)</i>	1A

Lab code	Strain name	Genotype	Source	Usage (Figures)
YBP2140	<i>NUP60-protA</i>	(DF5) <i>NUP60-protA::HIS3-URA3</i>	(Rout et al., 2000) (a)	1A, 1B, S1B-C
YBP2141	<i>NUP59-protA</i>	(DF5) <i>NUP59-protA::HIS3-URA3</i>	(Rout et al., 2000) (a)	1A, 2A, 2E-F (EDTA and puromycin panels), S1A-B, S2A-B, S2F-G, S2J
YBP2142	<i>NUP57-protA</i>	(DF5) <i>NUP57-protA::HIS5</i>	(Rout et al., 2000) (a)	1A
YBP2143	<i>NUP53-protA</i>	(DF5) <i>NUP53-protA::HIS3-URA3</i>	(Rout et al., 2000) (a)	1A, 1B, S1B-C
YBP2144	<i>NUP49-protA</i>	(DF5) <i>NUP49-protA::HIS5</i>	(Rout et al., 2000) (a)	1A
YBP2145	<i>NUP42-protA</i>	(DF5) <i>NUP42-protA::HIS5</i>	(Rout et al., 2000) (a)	1A
YBP2146	<i>GLE2-protA</i>	(DF5) <i>GLE2-protA::HIS3-URA3</i>	(Rout et al., 2000) (a)	1A-C, S1A-C
YBP2147	<i>SEH1-protA</i>	(DF5) <i>SEH1-protA::HIS5</i>	(Rout et al., 2000) (a)	1A
YBP1347	<i>POM34-protA</i>	(DF5) <i>POM34-protA::HIS5</i>	(Rout et al., 2000) (a)	1A
YBP1025	<i>POM33-protA</i>	(DF5) <i>POM33-protA::HIS5</i>	(Chadrin et al., 2010)	1A
YBP2300	<i>NUP145N-protA</i>	(BY4742) <i>NUP145N (ΔC)-protA::HIS5</i>	this study (b)	1A
YBP1519	<i>protA</i>	(W303) <i>ZPR1^{prom}-protA::HIS5</i>	(Oeffinger et al., 2007) (a)	1A
YBP2299	<i>RNR3-GFP</i>	(BY4741) <i>RNR3-GFP::HIS5</i>	(Huh et al., 2003)	S2H
YBP517	<i>GFP-Nup133ΔN</i>	(S288c) <i>ura3 trp1 leu2 lys2 met14 arg8 Mfa1^{prom}-LEU2 GFP-Nup133ΔN</i>	this study (c)	S2I
YBP2215	<i>NUP59-protA</i>	(BY4741) <i>NUP59-pA::HIS5</i>	this study (d)	2E-F and S2K (transformed with <i>wt</i> or Δ AUG <i>NUP1/NUP2-GFP</i> expression vectors), 3B-C and S3C (transformed with <i>wt</i> or Δ Kap-ID <i>NUP1/NUP2-GFP</i> expression vectors), S3F (transformed with the <i>kap123ΔN</i> plasmid), S3H-I (transformed with <i>wt/PPAA NUP1-GFP</i> expression vectors)
YBP2268	<i>ESC1-GFP</i>	(W303) <i>ESC1-GFP::TRP1</i>	this study (e)	S3B
YBP2269	<i>ESC1(1-1123)-GFP</i>	(W303) <i>ESC1(1-1123)-GFP::TRP1</i>	this study (e)	S3B
YBP2270	<i>ESC1(1-633)-GFP</i>	(W303) <i>ESC1(1-633)-GFP::TRP1</i>	this study (e)	S3B
YBP2271	<i>ADH1^{prom}-ESC1-GFP</i>	(W303) <i>esc1::KanMX::ADH1^{prom}-ESC1-GFP::TRP1</i>	this study (f)	S3B
YBP2272	<i>ADH1^{prom}-ESC1(Δ1-1123)-GFP</i>	(W303) <i>esc1::KanMX::ADH1^{prom}-ESC1(Δ1-1123)-GFP::TRP1</i>	this study (f)	S3B
YBP2273	<i>ADH1^{prom}-ESC1(Δ1-633)-GFP</i>	(W303) <i>esc1::KanMX::ADH1^{prom}-ESC1(Δ1-633)-GFP::TRP1</i>	this study (f)	S3B
YBP2274	<i>ADH1^{prom}-ESC1(634-1123)-GFP</i>	(W303) <i>esc1::KanMX::ADH1^{prom}-ESC1(634-1123)-GFP::TRP1</i>	this study (f)	S3B
YBP2220	<i>KAP60-AnchorAway</i>	(W303) <i>tor1-1 fpr1::NatMX RPL13A-2xFKB12::TRP1 SRP1(KAP60)-FRB::kanMX6</i>	(Haruki et al., 2008) (g)	3C
YBP2225	<i>KAP60-AnchorAway NUP59-protA</i>	(W303) <i>tor1-1 fpr1::NatMX RPL13A-2xFKB12::TRP1 SRP1(KAP60)-FRB::kanMX6 NUP59-protA::HIS5</i>	this study (d)	3C
YBP2275	<i>NUP2-Kap-ID-GFP</i>	(DF5) <i>NUP2(Δ176-720)-GFP::KanMX</i>	this study (h)	3E
YBP2276	<i>NUP2-Kap-ID-GFP NUP59-protA</i>	(DF5) <i>NUP2(Δ176-720)-GFP::KanMX NUP59-protA::HIS3-URA3</i>	this study (h)	3E
YBP2261	<i>NUP2-Kap-ID-4xGFP NUP59-protA</i>	(DF5) <i>NUP2(Δ176-720)-4xGFP::KanMX NUP59-protA::HIS3-URA3</i>	this study (h)	3E
YBP2205	<i>NUP2-GFP NUP59-protA</i>	(DF5) <i>NUP2-GFP::TRP1 NUP59-protA::HIS3-URA3</i>	this study (i)	3E
YBP2254	<i>hek2Δ NUP59-protA</i>	(BY4741) <i>hek2::KanMX NUP59-protA::HIS5</i>	this study (j)	S3H-I (transformed with <i>wt NUP1-GFP</i> expression vector)

Lab code	Strain name	Genotype	Source	Usage (Figures)
YBP2234	<i>NUP1-GFP W303</i>	(W303) <i>tor1-1 fpr1::NatMX PMA1-2xFKBP12::TRP1 NUP1-GFP-12xMS2loops</i>	<i>this study (k)</i>	S4A-B
YBP2245	<i>NUP1-GFP hek2Δ W303</i>	(W303) <i>tor1-1 fpr1::NatMX PMA1-2xFKBP12::TRP1 NUP1-GFP-12xMS2loops hek2::KanMX</i>	<i>this study (j)</i>	S4A-B
YBP2259	<i>KAP60-mCherry</i>	(BY4742) <i>SRP1(KAP60)-mCherry::KanMX</i>	<i>this study (l)</i>	4F (transformed with <i>NUP1-GFP-ASH1^{3'UTR}</i> expression vector)
YBP2277	<i>NIC96-mRFP</i>	(BY4742) <i>NIC96-mRFP::KanMX</i>	(Huh et al., 2003) (m)	4F (transformed with <i>NUP1-GFP-ASH1^{3'UTR}</i> expression vector)

(a) A gift from M. Rout (*Rockefeller University, New York*). The genomic copy of each nucleoporin-coding gene is tagged through the C-terminal integration of a DNA fragment encoding the IgG-binding domains of *S. aureus* protein A. Expression of the tagged nucleoporin is driven by its endogenous promoter. Expression of the protA tag alone is driven by the *ZPR1* promoter.

(b) C-ter deletion of *NUP145* was obtained by integration of a *protA-HIS5* cassette amplified from pBXA in a diploid *wt* strain (BY4743). The *NUP145N(Δ606-1317)-protA* haploid strain was further recovered following sporulation.

(c) N-ter deletion and GFP-tagging of *NUP133* was achieved by integration of the *nup133::GFP-nup133ΔN* allele (Belgareh and Doye, 1997) into the *nup133::URA3* locus of YV460 (Loeillet et al., 2005). Expression of the GFP-tagged version of the ΔN-nup133 truncation is driven by the endogenous *NUP133* promoter.

(d) *NUP59* was C-terminally tagged by homologous recombination using a *protA-HIS5* cassette amplified from pBXA. Expression of Nup59-protA is driven by its endogenous promoter.

(e) C-ter deletions of *ESC1* were obtained by C-terminal integration of a *GFP::TRP1* cassette amplified from pYM26. Expression of Esc1-GFP truncations are driven by the endogenous *ESC1* promoter.

(f) N-ter deletions of *ESC1* were obtained by integration of *KanMX::ADHI^{prom}* cassettes amplified from pYM-N6 in *ESC1-GFP* or *ESC1(1-1123)-GFP* cells. Expression of Esc1-GFP truncations are driven by the *ADHI* promoter.

(g) Obtained from Euroscarf.

(h) C-ter deletions of *NUP2* were obtained by C-terminal integration of a *GFP::KanMX* cassette (amplified from pYM12) or a *4xGFP::KanMX* cassette (amplified from pSM1023) in *wt (DF5)* or *NUP59-protA* cells. Expression of the Nup2 N-terminal Kap-Interacting Domain is driven by the endogenous genomic *NUP2* promoter.

(i) *NUP2* was C-terminally tagged with GFP in *NUP59-protA* cells by homologous recombination using a *GFP::TRP1* cassette amplified from pFA6a-GFP::TRP1. Expression of Nup2-GFP is driven by its endogenous promoter.

(j) The complete *HEK2* CDS was deleted by homologous recombination using a *KanMX* cassette in *NUP59-protA* (YBP2215) or *NUP1-GFP* (YBP2234) cells.

(k) *NUP1* was C-terminally tagged with GFP by homologous recombination with a *GFP-HIS5-MS2loops* cassette amplified through fusion PCR using *pFA6a-GFP::KanMX* and *pLOXHIS5MS2L* as templates. The *lox-HIS5-lox* marker was removed by Cre-mediated pop-out. Nup1-GFP expression is driven by its endogenous promoter and terminator.

(l) *KAP60 (SRP1)* was C-terminally tagged with mCherry by homologous recombination using a cassette amplified from *pYM-mCherryFP::KanMX*.

(m) A gift from W.K. Huh.

Table S3 – Oligonucleotides used in this study.**A.** qPCR primers used in this study.

Primer name	Sequence (5'-> 3')
<i>NUP145N-F (Fig. 1A)</i>	CTCTAGAGAATGTTTATCCT
<i>NUP145N-R (Fig. 1A)</i>	TGAAATTCAGCAAACCTCGTCGT
<i>NUP145-F (Fig. 1C)</i>	GGGAATTGAAGCTTCCCCTA
<i>NUP145-R (Fig. 1C)</i>	TGGCCCGTCAAATTTCTTAG
<i>NUP133-F</i>	CGCCAGGTGCATACTAACT
<i>NUP133-R</i>	AATGATAAGCCCTCCGGTTT
<i>NUP120-F</i>	GCATTCACGAGGCTTCTTC
<i>NUP120-R</i>	GGGAATGATTATACGGTTGGAA
<i>NUP85-F</i>	ACCATGCGTCGACATCATT
<i>NUP85-R</i>	TAGCGGTAAATGCTGCTGTG
<i>NUP84-F</i>	AGAGGACCCCAAGTAAGGA
<i>NUP84-R</i>	CGCTGTGTGGTTCTTCTCAA
<i>SEH1-F</i>	GCAGTCTCTGCATTGGAACA
<i>SEH1-R</i>	TTATGACCGGAAAGTTTGC
<i>NUP192-F</i>	ACGGGCTTATGCTCTTTGA
<i>NUP192-R</i>	ATTTTCGCAATCCACCAAAG
<i>NUP188-F</i>	CACAACATTTGGAGCAATGG
<i>NUP188-R</i>	GGCACGTCTCAGTAAAACC
<i>NUP170-F</i>	TGTGGATCATTCTGCTCTGC
<i>NUP170-R</i>	CGCAAGCCAATTTCTTTAGC
<i>NUP157-F</i>	AGGCTTGCTCTACGGCACT
<i>NUP157-R</i>	CCAGGTATACCGCAGAGAA
<i>NIC96-F</i>	GTTGATCGGAAGATGCGACT
<i>NIC96-R</i>	ATGAGATGCATCCACAACCA
<i>NUP59-F</i>	CACCACAGACAACCCAGATG
<i>NUP59-R</i>	AATTGCAAGTGTGCTGCTG
<i>NUP53-F</i>	AAGCCTACAGCCACACCTTC
<i>NUP53-R</i>	GTATTCGGCGTTGTTTTCGT
<i>NDC1-F</i>	CCTGCTTAGTGGGCTTTCTG
<i>NDC1-R</i>	GATCCAGTGAAGTGGCCCTA
<i>POM152-F</i>	TTGTCCAGGTGAAATTGTGG
<i>POM152-R</i>	TCCCACACACTGGTCCAATA
<i>POM34-F</i>	ACAGTTTTGCAACCGGCTAT
<i>POM34-R</i>	ACCAACTGAACGTGGTGTGA
<i>POM33-F</i>	GCGTTGCTGGACTCCTAAAG
<i>POM33-R</i>	TGGAATACAGCAAACGCGAA
<i>NUP116-F</i>	CCTTTGTCAGGTGAATCGT
<i>NUP116-R</i>	TTGCGTTAGCGTTTGATTG
<i>NUP100-F</i>	GGGATCTTGTCACCTTTGGA
<i>NUP100-R</i>	ATTAATGCCTTCGCCCTTTT

Primer name	Sequence (5'-> 3')
<i>GLE2-F</i>	TGAAGCCGATGGATATGCGA
<i>GLE2-R</i>	GGATTGGTTTGACGATGGCA
<i>NSP1-F</i>	CCCTTTCATTTGGTTCAGGA
<i>NSP1-R</i>	GCTGGTTTTGCTGGTTCATT
<i>NUP57-F</i>	CGGCAATAGCACTCAAAACA
<i>NUP57-R</i>	CCAAATAGGCCTCCCGTAGT
<i>NUP49-F</i>	TGGTAATCAGCAGCAGCAAC
<i>NUP49-R</i>	AATCCGTGTCATTGGGGTAA
<i>NUP159-F</i>	TCACCTTCCCATCTTCTGG
<i>NUP159-R</i>	GCATGGTCTTTGGTTTCGT
<i>NUP82-F</i>	TGTCAATGTCGTGGGATGTT
<i>NUP82-R</i>	TTCCCAACTGTTCAAGTAACG
<i>NUP42-F</i>	ACTTTCGGGCAGCTACTAA
<i>NUP42-R</i>	AAATGTGCCAAATGGGGATA
<i>GLE1-F</i>	CCGTACAGTTTCTTCAGGCG
<i>GLE1-R</i>	CTTAATCTCGCAGCACCCAC
<i>NUP1-F</i>	CTCTGAGGGAAGTGCGAAAC
<i>NUP1-R</i>	CGAAAACGAGGGTTTAGCTG
<i>NUP2-F</i>	CGCAAGATGCAACCAAAGTA
<i>NUP2-R</i>	AAGCCACTTCGTCTTCTCA
<i>NUP60-F</i>	CTAGCCTGCCCTACCATAC
<i>NUP60-R</i>	TGAGCAGATGTGTGGCTGT
<i>MLP1-F</i>	TTGGCTAGAAAAGGAGCTACG
<i>MLP1-R</i>	TCTTCCATTTGAAAATCATTCC
<i>MLP2-F</i>	GCACGTAGAGAACCTCGAAGA
<i>MLP2-R</i>	CTTCTCCACCGACTGGGATA
<i>ACT1-F</i>	ACGTTACCAATTGAAACACG
<i>ACT1-R</i>	AGAACAGGGTGTCTTCTGG
<i>mito_21S_rRNA-F</i>	GTGTGAACTCTGTCCATGC
<i>mito_21S_rRNA-R</i>	TCATGCGGGACCTCAATTAT
<i>ULP1-F</i>	ACCCCTAATACAGTGGCGTT
<i>ULP1-R</i>	TTCTCTTCATCCACCTCCGG
<i>ESC1-F</i>	TGGAAGCTGGAGTGACGAAA
<i>ESC1-R</i>	CATCCAAATCGCTCACCTGT
<i>TAF1-F</i>	GCAGCAAGGCAAAATTCGAA
<i>TAF1-R</i>	GCTTCCCGTTCAGTTACACC
<i>GFP-F</i>	AGTGGAGAGGGTGAAGGTGA
<i>GFP-R</i>	GTTGGCCATGGAACAGGTAG
<i>NUP2-KapID-F</i>	CCGATGCGCAAATACAGAGA
<i>NUP2-KapID-R</i>	ATCACAGCAGATGACGCAAC

B. FISH probes used in this study.

Probe #	Sequence (5'→ 3')	Probe position
NUP1-1	TCGATGGCCGTTTTTTATTG	90
NUP1-2	TGATTCGCGTATGAGAGGTT	130
NUP1-3	ACATGCAGTGTGTCTCAAC	166
NUP1-4	AGTGTTCGCGTTCATATA	266
NUP1-5	AAATTGGCAAAAGCGGCGGT	291
NUP1-6	TTCTCCCTTAATAACCTCAG	325
NUP1-7	CAAGCTCACGCATATTTCTC	354
NUP1-8	CCCAATATAACCGACGATGT	406
NUP1-9	CCTTTTTTGTATAGCGTCA	657
NUP1-10	GCCAGATTGCCATTTGAAAT	751
NUP1-11	GGGAACAAAGCTCTCCTGAA	899
NUP1-12	TCCTTTTTGGGCTCAATATT	955
NUP1-13	CCTACGGTGGGCAAAACAAT	985
NUP1-14	GTTGCCCTAGGAGAAGTTTT	1042
NUP1-15	GTTTTCGTTGACTTATCGGT	1108
NUP1-16	AGTTTTATTAGCCTTTGGC	1160
NUP1-17	GGAAGGGACAGTATTGTGCA	1187
NUP1-18	GGTATCTGATTACCACCAA	1223
NUP1-19	TTCATCATCACCTTCATCAC	1370
NUP1-20	CAGGTAACGCCTTTTTCTT	1401
NUP1-21	GCAGAGGAGTCAATATTGCT	1603
NUP1-22	GTAGGCTTCTGACTATTGT	1690
NUP1-23	TCCCTCAGAGATCTTATTTT	1736
NUP1-24	CCTCCTCTGACTTAGAGAAA	1776
NUP1-25	GCTTGAACATCAACAGGCTT	1855
NUP1-26	GGCTTGAGAGTCTTATCATC	1882
NUP1-27	CTTTTTGAGCAGGTTCAAGTA	1914
NUP1-28	ATCTGTCTCATTAGCAGGAG	2081
NUP1-29	GCCAAAGGTGAAAGATGGCG	2114
NUP1-30	TTTTGTGCTAGTGGTTGTTG	2153
NUP1-31	CGGGAGCCCCAAAATAAAA	2178
NUP1-32	TTACCCAAGACAGGAATCGG	2338
NUP1-33	GTGTTAGCAGTACCGAATGA	2401
NUP1-34	AAGTAGTAGTTGTGCCGTTA	2481
NUP1-35	CTATATTGGTCTCAGAGGTA	2505
NUP1-36	CGATACTTTGATCCGGTTTT	2544
NUP1-37	CCTGAGCTCGAAAAGCCAAA	2596
NUP1-38	ATTAGAAGCTGCACCAGTTG	2624
NUP1-39	CCAGCATTAGCATTAGTTGA	2710
NUP1-40	CACCGTATTAGCATTAGTGT	2846
NUP1-41	TTTGCGATTGATGTGGTTGG	2883
NUP1-42	GGAAGTGTGATGGCGTGAA	2929
NUP1-43	ATTAGTAATTCGCCATTCA	2969
NUP1-44	CCAAATATATCACTTGGCCT	3007
NUP1-45	AAAGAAGTTGTCCGCACACC	3094
NUP1-46	CATGCTCATGCCATTATTTG	3155
NUP1-47	ATCTTTCTGTTCGCCATAAC	3184
NUP1-48	TTTAGAGTGCCCATCTTGT	3206

Probe #	Sequence (5'→ 3')	Probe position
NUP2-1	CGTACGTTTCTCTCTGTATT	27
NUP2-2	TCATCGTCAGACTCGTTAGA	49
NUP2-3	CATCACAGCAGATGACGCAA	89
NUP2-4	TGGCATGGCAATTTTCTTC	113
NUP2-5	ACCAAAAGGTTTGAACGCCA	143
NUP2-6	CTTGGTTTCATCCGATTTTG	167
NUP2-7	GTAGGGCTATTATCAACCTG	238
NUP2-8	GCTTTTGTCTGGAATTGCT	265
NUP2-9	CATCAACCTTAGCCTTGAAC	297
NUP2-10	TAACGGCTTGCCTAGAATA	320
NUP2-11	GATAGATTCACGGGAGCTT	395
NUP2-12	GACATCTTCTACTTTGGCAG	452
NUP2-13	TCACTTCATCTTCTTGGCG	592
NUP2-14	CCCTTGATTTCTATATCGTT	619
NUP2-15	CATTTTATCGGTGCTTGGGA	690
NUP2-16	TAGCATTAGTCTCGGTTTTT	714
NUP2-17	AGTAGTTGAAGTGCCGAAG	749
NUP2-18	CTTGGTAGCTTCTGTCAATG	797
NUP2-19	TTACTGTGTGTGCCACATT	820
NUP2-20	TTTTGAGAATCCGCAGCATG	871
NUP2-21	CAGCTGCTTGACCAATAACA	906
NUP2-22	ATGAGCTTTTTCTAGCGAT	933
NUP2-23	GAGGTTGAGTTTTCGTCATT	985
NUP2-24	GATGCATCAGGTGTATTCTT	1075
NUP2-25	AAGAGTTTGGGACCCAAAA	1110
NUP2-26	TACCGTTTTGAAGTTTCGT	1136
NUP2-27	ACGTTGTGTGTGCATCTTC	1201
NUP2-28	TGAAGGCAGGCTTAGAGGAA	1230
NUP2-29	TTGAGTCCTTTTACTTTCC	1284
NUP2-30	CTCCGTTTGATATGCCAAA	1318
NUP2-31	TCAACCGCAGAGGGTAAAGA	1363
NUP2-32	TTGTGTGCTTCTTCTTGTG	1393
NUP2-33	TTAGTATCCGCGGTTTAG	1448
NUP2-34	CTCTTTATTGTCAGCGAGTG	1496
NUP2-35	GGAGTATTATTAGGCTGGGA	1546
NUP2-36	TTTGCTGTGTTTTTCCGAA	1576
NUP2-37	GTATAGAGGGAGCAGGAGAT	1617
NUP2-38	CCTTGCTATCATTGTAGTT	1689
NUP2-39	GTTGCTTCTGTGTGTGATTC	1711
NUP2-40	TCTACTTTGGTTGCATCTTG	1744
NUP2-41	CTTTGATTCTTCTGGGGTAG	1766
NUP2-42	GTACGATTGGTTTCAGCAT	1859
NUP2-43	TCAAAGCTTCATTTCGCCT	1893
NUP2-44	CACTTTGAAGGATCGTCCT	1919
NUP2-45	ACCGTCAGACCTACAAAGTA	1943
NUP2-46	AGGAGTCTACAACAGTTGCA	1983
NUP2-47	TTATCATTTCCGGGAGCTAA	2017
NUP2-48	AGTTACAAGTTCCCATCAG	2063

Probe #	Sequence (5'→ 3')	Probe position
<i>GFP-1</i>	TGTTAATTAACCCGGGGATC	2
<i>GFP-2</i>	CCAGTGAAAAGTTCTTCTCC	27
<i>GFP-3</i>	CCCATTAACATCACCATCTA	70
<i>GFP-4</i>	CCTCTCCACTGACAGAAAAT	95
<i>GFP-5</i>	GTAAGTTTCCGTATGTTGC	126
<i>GFP-6</i>	GTAGTTTCCAGTAGTGCAA	158
<i>GFP-7</i>	ACAAGTGTGGCCATGGAAC	180
<i>GFP-8</i>	GCATTGAACACCATAAGTGA	208
<i>GFP-9</i>	TCATGCCGTTTCATATGATC	243
<i>GFP-10</i>	GGGCATGGCACTCTTAAAA	265
<i>GFP-11</i>	TTCTTCTGTACATAACCT	287
<i>GFP-12</i>	GTTCCCGTCATCTTTGAAAA	313
<i>GFP-13</i>	TGACTTCAGCAGCGTGTCTTG	335
<i>GFP-14</i>	TAACAAGGGTATCACCTCA	359
<i>GFP-15</i>	ATACCTTTAACTCGATTCT	381
<i>GFP-16</i>	GTGTCCAAGAATGTTCCAT	415
<i>GFP-17</i>	GTGAGTTATAGTTGTATTCC	440
<i>GFP-18</i>	GTCTGCCATGATGTATACAT	463
<i>GFP-19</i>	CTTTGATTCCATTCTTTTGT	485
<i>GFP-20</i>	CCATCTCAATGTTGTGTCT	519
<i>GFP-21</i>	ATGGTCTGCTAGTTGAACGC	541
<i>GFP-22</i>	CGCCAATTGGAGTATTTTGT	566
<i>GFP-23</i>	GTCTGGTAAAAGGACAGGGC	589
<i>GFP-24</i>	AAGGGCAGATTGTGTGGACA	619
<i>GFP-25</i>	TCTTTTTCGTTGGGATCTTTC	641
<i>GFP-26</i>	TCAAGAAGGACCATGTGGTC	663
<i>GFP-27</i>	AATCCCAGCAGCTGTTACAA	685
<i>GFP-28</i>	TATAGTTCATCCATGCCATG	708

Probe #	Sequence (5'→ 3')	Probe position
<i>NSP1-1</i>	CTGATTGGTGGTGTAGAGT	62
<i>NSP1-2</i>	TTGATTGACCTGTCCGAAG	105
<i>NSP1-3</i>	GGCGCAGAATTGTTGAAACC	130
<i>NSP1-4</i>	ATTGCTACCAAATGCAGGTG	182
<i>NSP1-5</i>	ACCGAATGCAGTATTACCAG	206
<i>NSP1-6</i>	AGTCGTAGAATTGTTGCTGC	254
<i>NSP1-7</i>	TCGTTTGTGAGCACTGGAA	318
<i>NSP1-8</i>	ACCAAATGTATTTCCTCCAG	353
<i>NSP1-9</i>	AACGCAGGTTTTGTCGTATT	406
<i>NSP1-10</i>	CGAGCTAGGAGTTGTGTTAT	452
<i>NSP1-11</i>	CCAAAGGAAAAGGCTGGCTT	526
<i>NSP1-12</i>	CTTGTCTGGTCTGTTTTTTT	566
<i>NSP1-13</i>	TCTGTCTATTACCACACT	607
<i>NSP1-14</i>	AAAATCCTGTAGTGGGAGCT	633
<i>NSP1-15</i>	GTTTTGCTGGTTCATTAGTT	771
<i>NSP1-16</i>	TTATTGTCGGATGTTGCAGT	808
<i>NSP1-17</i>	AACTTGGGGTGTGTTAGTT	831
<i>NSP1-18</i>	TCATCCGATTTAGCACCAAA	856
<i>NSP1-19</i>	GAGAAGGCAGGCTTAGAAGT	892
<i>NSP1-20</i>	ATCATCTTCTTTTCTTCTG	926
<i>NSP1-21</i>	CATCTTGCTTGCTTCAATTT	984
<i>NSP1-22</i>	GGCTTGGCACCAAAGGAAAA	1021
<i>NSP1-23</i>	ATCCTTTTCTCATCAGACT	1094
<i>NSP1-24</i>	GCCAAAAGAGAAGGCAGGTT	1127
<i>NSP1-25</i>	CTAGCTTATTTTCATCCGG	1153
<i>NSP1-26</i>	GAGAAGGCAGGCTTAGAAGT	892
<i>NSP1-27</i>	ATCATCTTCTTTTCTTCTG	926
<i>NSP1-28</i>	CATCTTGCTTGCTTCAATTT	984
<i>NSP1-29</i>	GGCTTGGCACCAAAGGAAAA	1021
<i>NSP1-30</i>	GGCTTGAAGTTTCATTGTT	1339
<i>NSP1-31</i>	TTTTTCTCATCGGACTTTGC	1375
<i>NSP1-32</i>	TCCTTTTTTTCATCCGATTT	1435
<i>NSP1-33</i>	CTGTCCTTTTTTCATTCGA	1495
<i>NSP1-34</i>	GAAGGCAGGCTTCGAAGAAC	1517
<i>NSP1-35</i>	TCATTCTTCTTTTCATCGGG	1552
<i>NSP1-36</i>	CGAATGAGAAGGCAGGCTTG	1581
<i>NSP1-37</i>	TTTCTTTTCGTTAGCCTTTG	1604
<i>NSP1-38</i>	CCTGTAGGCTTAGAACCAAA	1654
<i>NSP1-39</i>	TTTCTGTTCTTCGGTTTAG	1718
<i>NSP1-40</i>	ATTTGACATCTGCGTTGAC	1821
<i>NSP1-41</i>	ACTGGCTTCAATTCCTACTGG	1870
<i>NSP1-42</i>	ATCGTCCAGTGTTTTATTGT	1898
<i>NSP1-43</i>	TTGGTTAGTCCATTTTCGTTA	1922
<i>NSP1-44</i>	TGACTAATTTGTTCACTCC	2014
<i>NSP1-45</i>	GCCTCTCGATATATTGTAGA	2088
<i>NSP1-46</i>	AGCACCTGAACCTGTAGACA	2177
<i>NSP1-47</i>	AAGTCTGGGCTGTCTTATAG	2232
<i>NSP1-48</i>	CTAGTGACCTTAAAGCGTCG	2403

Table S4 - Plasmids used in this study.

Lab code	Plasmid name	Main features / Usage	Source
pBP179	pUN100-GFP- Δ N-nup133	<i>CEN/URA3/NUP133^{prom}-GFP-ΔN-nup133-NUP133^{3'UTR}</i>	(Belgareh and Doye, 1997)
pBP1025	pBXA	for protein A tagging (<i>HIS5</i>)	(Rout et al., 2000)
pBP2136	pYM26	for GFP tagging (<i>TRP1</i>)	(Janke et al., 2004)
pBP2135	pYM12	for GFP tagging (<i>KanMX</i>)	(Janke et al., 2004)
pBP2137	pYM-N6	For N-ter truncations (<i>KanMX::ADH1^{prom}</i>)	(Janke et al., 2004)
pBP1166	pSM1023	for 4xGFP tagging (<i>KanMX</i>)	(Maekawa et al., 2003)
pBP2098	pFA6a-GFP:: <i>TRP1</i>	for GFP tagging (<i>TRP1</i>)	(Longtine et al., 1998)
pBP414	pFA6a-KanMX	for deletion (<i>KanMX</i>)	(Longtine et al., 1998)
pBP1017	pYM-mCherryFP:: <i>KanMX</i>	for mCherryFP tagging (<i>KanMX</i>)	A gift from S. Léon.
pBP314	pFA6a-GFP:: <i>KanMX</i>	for GFP tagging (<i>KanMX</i>)	(Longtine et al., 1998)
pBP1696	pLOXHIS5MS2L	for MS2 loop tagging (<i>loxP-HIS5-loxP</i>)	(Haim-Vilimovsky and Gerst, 2009)
pBP2118	pRS316-NUP1-GFP	<i>CEN/URA3/NUP1^{prom}-NUP1-GFP-ADH1^{3'UTR}</i>	this study (a)
pBP2133	pRS316-nup1- Δ AUG-GFP	<i>CEN/URA3/NUP1^{prom}-nup1-ΔAUG-GFP-ADH1^{3'UTR}</i>	this study (b)
pBP2120	pRS316-NUP2-GFP	<i>CEN/URA3/NUP2^{prom}-NUP2-GFP-ADH1^{3'UTR}</i>	this study (a)
pBP2134	pRS316-nup2- Δ AUG-GFP	<i>CEN/URA3/NUP2^{prom}-nup2-ΔAUG-GFP-ADH1^{3'UTR}</i>	this study (b)
pBP2119	pRS316-nup1- Δ Kap-ID-GFP	<i>CEN/URA3/NUP1^{prom}-nup1-ΔKap-ID-ADH1^{3'UTR}</i>	this study (c)
pBP2121	pRS316-nup2- Δ Kap-ID-GFP	<i>CEN/URA3/NUP2^{prom}-nup2-ΔKap-ID-GFP-ADH1^{3'UTR}</i>	this study (c)
pBP795	YCp-GAL- <i>yrb4</i> Δ N (<i>kap123</i> Δ N)	<i>CEN/URA3/GAL1-10^{prom}-ΔN-kap123</i>	(Panse et al., 2003)
pBP731	pRS315-GFP-ULP1	<i>CEN/LEU2/NOP1^{prom}-GFP-ULP1</i>	(Panse et al., 2003)
pBP2130	pRS316-nup1-P682A,P683A-GFP	<i>CEN/URA3/NUP1^{prom}-nup1-P682A,P683A-ADH1^{3'UTR}</i>	this study (d)
pBP2131	pRS316-NUP1-GFP- <i>ASH1</i> ^{3'UTR}	<i>CEN/URA3/NUP1^{prom}-NUP1-GFP-ASH1^{3'UTR}</i>	this study (e)
pBP1028	pUN100-NUP49-mCherry	<i>CEN/LEU2/NUP49-mCherry</i>	(Chadrin et al., 2010)

- (a) The expression cassette encompasses *NUP1* (or *NUP2*) promoter (500bp upstream the start codon), *NUP1* (or *NUP2*) complete CDS, the GFP coding sequence (in frame with *NUP1/NUP2* CDS) and the *ADH1* 3'UTR.
- (b) For *NUP1*, the first AUG (position +1) and a secondary AUG codon, in frame with the start codon (position +25) were both removed by PCR-based techniques. For *NUP2*, only the start codon (position +1) was removed.
- (c) The sequences encoding Nup1 KapID (aa2-122) (Mészáros et al., 2015) and Nup2-KapID (aa2-175) (Solsbacher et al., 2000) were removed by PCR-based techniques.
- (d) P682A and P683A mutations were introduced in *pRS316-NUP1-GFP* by PCR-based techniques.
- (e) The *ADH1* 3'UTR flanking the *NUP1-GFP* transgene in *pRS316-NUP1-GFP* was replaced by the 3'UTR from *ASH1* (position +1750-2018 with respect to *ASH1* start codon), a motif sufficient to drive mRNA localization to the bud (Long et al., 1997).

Supplemental references

- Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., et al. (2007). The molecular architecture of the nuclear pore complex. *Nature* 450, 695–701.
- Aviram, N., and Schuldiner, M. (2017). Targeting and translocation of proteins to the endoplasmic reticulum at a glance. *J. Cell. Sci.* 130, 4079–4085.
- Bailer, S.M., Siniosoglou, S., Podtelejnikov, A., Hellwig, A., Mann, M., and Hurt, E. (1998). Nup116p and nup100p are interchangeable through a conserved motif which constitutes a docking site for the mRNA transport factor gle2p. *EMBO J* 17, 1107–1119.
- Bailer, S.M., Balduf, C., Katahira, J., Podtelejnikov, A., Rollenhagen, C., Mann, M., Pante, N., and Hurt, E. (2000). Nup116p associates with the Nup82p-Nsp1p-Nup159p nucleoporin complex. *J. Biol. Chem.* 275, 23540–23548.
- Belgareh, N., and Doye, V. (1997). Dynamics of nuclear pore distribution in nucleoporin mutant yeast cells. *J Cell Biol* 136, 747–759.
- Belgareh, N., Snay-Hodge, C., Pasteau, F., Dagher, S., Cole, C.N., and Doye, V. (1998). Functional characterization of a Nup159p-containing nuclear pore subcomplex. *Mol. Biol. Cell* 9, 3475–3492.
- Blobel, G., and Sabatini, D. (1971). Dissociation of mammalian polyribosomes into subunits by puromycin. *Proc. Natl. Acad. Sci. U.S.A.* 68, 390–394.
- Boehlke, K.W., and Friesen, J.D. (1975). Cellular content of ribonucleic acid and protein in *Saccharomyces cerevisiae* as a function of exponential growth rate: calculation of the apparent peptide chain elongation rate. *J. Bacteriol.* 121, 429–433.
- Chadrin, A., Hess, B., San Roman, M., Gatti, X., Lombard, B., Loew, D., Barral, Y., Palancade, B., and Doye, V. (2010). Pom33, a novel transmembrane nucleoporin required for proper nuclear pore complex distribution. *J Cell Biol* 189, 795–811.
- Denning, D., Mykytka, B., Allen, N.P., Huang, L., Al Burlingame, null, and Rexach, M. (2001). The nucleoporin Nup60p functions as a Gsp1p-GTP-sensitive tether for Nup2p at the nuclear pore complex. *J. Cell Biol.* 154, 937–950.
- Devos, D., Dokudovskaya, S., Alber, F., Williams, R., Chait, B.T., Sali, A., and Rout, M.P. (2004). Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol* 2, e380.
- Dilworth, D.J., Suprpto, A., Padovan, J.C., Chait, B.T., Wozniak, R.W., Rout, M.P., and Aitchison, J.D. (2001). Nup2p dynamically associates with the distal regions of the yeast nuclear pore complex. *J Cell Biol* 153, 1465–1478.
- Enenkel, C., Schülke, N., and Blobel, G. (1996). Expression in yeast of binding regions of karyopherins alpha and beta inhibits nuclear import and cell growth. *Proc. Natl. Acad. Sci. U.S.A.* 93, 12986–12991.
- Fernandez-Martinez, J., Kim, S.J., Shi, Y., Upla, P., Pellarin, R., Gagnon, M., Chemmama, I.E., Wang, J., Nudelman, I., Zhang, W., et al. (2016). Structure and Function of the Nuclear Pore Complex Cytoplasmic mRNA Export Platform. *Cell* 167, 1215–1228.e25.
- Fischer, J., Teimer, R., Amlacher, S., Kunze, R., and Hurt, E. (2015). Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. *Nat. Struct. Mol. Biol.* 22, 774–781.

Haim-Vilmovsky, L., and Gerst, J.E. (2009). m-TAG: a PCR-based genomic integration method to visualize the localization of specific endogenous mRNAs in vivo in yeast. *Nat Protoc* 4, 1274–1284.

Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Mol. Cell* 31, 925–932.

Ho, A.K., Racznik, G.A., Ives, E.B., and Went, S.R. (1998). The integral membrane protein snl1p is genetically linked to yeast nuclear pore complex function. *Mol Biol Cell* 9, 355–373.

Ho, A.K., Shen, T.X., Ryan, K.J., Kiseleva, E., Levy, M.A., Allen, T.D., and Went, S.R. (2000). Assembly and preferential localization of Nup116p on the cytoplasmic face of the nuclear pore complex by interaction with Nup82p. *Mol. Cell. Biol.* 20, 5736–5748.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O’Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686–691.

Hurwitz, M.E., Strambio-de-Castillia, C., and Blobel, G. (1998). Two yeast nuclear pore complex proteins involved in mRNA export form a cytoplasmically oriented subcomplex. *Proc Natl Acad Sci U S A* 95, 11241–11245.

Irvin, J.D., and Pugh, B.F. (2006). Genome-wide transcriptional dependence on TAF1 functional domains. *J. Biol. Chem.* 281, 6404–6412.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., et al. (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* 21, 947–962.

Kim, S.J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B.D., Hogan, J.A., Upla, P., et al. (2018). Integrative structure and functional anatomy of a nuclear pore complex. *Nature* 555, 475–482.

Lin, D.H., Stuwe, T., Schilbach, S., Rundlet, E.J., Perriches, T., Mobbs, G., Fan, Y., Thierbach, K., Huber, F.M., Collins, L.N., et al. (2016). Architecture of the symmetric core of the nuclear pore. *Science* 352, aaf1015.

Loeillet, S., Palancade, B., Cartron, M., Thierry, A., Richard, G.F., Dujon, B., Doye, V., and Nicolas, A. (2005). Genetic network interactions among replication, repair and nuclear pore deficiencies in yeast. *DNA Repair (Amst)* 4, 459–468.

Long, R.M., Singer, R.H., Meng, X., Gonzalez, I., Nasmyth, K., and Jansen, R.P. (1997). Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* 277, 383–387.

Longtine, M.S., McKenzie, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* 14, 953–961.

Lutzmann, M., Kunze, R., Stangl, K., Stelter, P., Tóth, K.F., Böttcher, B., and Hurt, E. (2005). Reconstitution of Nup157 and Nup145N into the Nup84 complex. *J. Biol. Chem.* 280, 18442–18451.

Maekawa, H., Usui, T., Knop, M., and Schiebel, E. (2003). Yeast Cdk1 translocates to the plus end of cytoplasmic microtubules to regulate bud cortex interactions. *EMBO J.* 22, 438–449.

Mandon, E.C., Trueman, S.F., and Gilmore, R. (2009). Translocation of proteins through the Sec61 and SecYEG channels. *Curr. Opin. Cell Biol.* 21, 501–507.

Mans, B.J., Anantharaman, V., Aravind, L., and Koonin, E.V. (2004). Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* 3, 1612–1637.

Marelli, M., Aitchison, J.D., and Wozniak, R.W. (1998). Specific binding of the karyopherin Kap121p to a subunit of the nuclear pore complex containing Nup53p, Nup59p, and Nup170p. *J. Cell Biol.* 143, 1813–1830.

Mészáros, N., Cibulka, J., Mendiburo, M.J., Romanauska, A., Schneider, M., and Köhler, A. (2015). Nuclear pore basket proteins are tethered to the nuclear envelope and can regulate membrane curvature. *Dev. Cell* 33, 285–298.

Niepel, M., Strambio-de-Castillia, C., Fasolo, J., Chait, B.T., and Rout, M.P. (2005). The nuclear pore complex-associated protein, Mlp2p, binds to the yeast spindle pole body and promotes its efficient assembly. *J Cell Biol* 170, 225–235.

Niepel, M., Molloy, K.R., Williams, R., Farr, J.C., Meinema, A.C., Vecchiotti, N., Cristea, I.M., Chait, B.T., Rout, M.P., and Strambio-De-Castillia, C. (2013). The nuclear basket proteins Mlp1p and Mlp2p are part of a dynamic interactome including Esc1p and the proteasome. *Mol Biol Cell* 24, 3920–3938.

Oeffinger, M., Wei, K.E., Rogers, R., DeGrasse, J.A., Chait, B.T., Aitchison, J.D., and Rout, M.P. (2007). Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat Methods* 4, 951–956.

Onischenko, E., Stanton, L.H., Madrid, A.S., Kieselbach, T., and Weis, K. (2009). Role of the Ndc1 interaction network in yeast nuclear pore complex assembly and maintenance. *J Cell Biol* 185, 475–491.

Onischenko, E., Tang, J.H., Andersen, K.R., Knockenhauer, K.E., Vallotton, P., Derrer, C.P., Kralt, A., Mugler, C.F., Chan, L.Y., Schwartz, T.U., et al. (2017). Natively Unfolded FG Repeats Stabilize the Structure of the Nuclear Pore Complex. *Cell* 171, 904-917.e19.

Panse, V.G., Kuster, B., Gerstberger, T., and Hurt, E. (2003). Unconventional tethering of Ulp1 to the transport channel of the nuclear pore complex by karyopherins. *Nat Cell Biol* 5, 21–27.

Ren, Y., Seo, H.-S., Blobel, G., and Hoelz, A. (2010). Structural and functional analysis of the interaction between the nucleoporin Nup98 and the mRNA export factor Rae1. *Proc Natl Acad Sci U S A* 107, 10406–10411.

Rout, M.P., Aitchison, J.D., Suprpto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol* 148, 635–651.

Rouviere, J.O., Bulfoni, M., Tuck, A., Cosson, B., Devaux, F., and Palancade, B. (2018). A SUMO-dependent feedback loop senses and controls the biogenesis of nuclear pore subunits. *Nature Communications* 9, 1665.

Solsbacher, J., Maurer, P., Vogel, F., and Schlenstedt, G. (2000). Nup2p, a yeast nucleoporin, functions in bidirectional transport of importin alpha. *Mol. Cell. Biol.* 20, 8468–8479.

Winey, M., Yarar, D., Giddings, T.H., and Mastronarde, D.N. (1997). Nuclear pore complex number and distribution throughout the *Saccharomyces cerevisiae* cell cycle by three-dimensional reconstruction from electron micrographs of nuclear envelopes. *Mol Biol Cell* 8, 2119–2132.

Wolf, J.J., Dowell, R.D., Mahony, S., Rabani, M., Gifford, D.K., and Fink, G.R. (2010). Feed-forward regulation of a cell fate determinant by an RNA-binding protein generates asymmetry in yeast. *Genetics* 185, 513–522.

Yoshida, K., Seo, H.-S., Debler, E.W., Blobel, G., and Hoelz, A. (2011). Structural and functional analysis of an essential nucleoporin heterotrimer on the cytoplasmic face of the nuclear pore complex. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16571–16576.

ARTICLE 2

A R-loop sensing pathway mediates the relocation of transcribed genes to nuclear pore complexes

The following manuscript reports the study of the reciprocal relationships between nuclear organisation and transcription-dependent genetic instability. In light of the interconnections between transcription, damage formation and loci localization with respect to the nuclear pore, we investigated whether R-loop formation could be a trigger for loci localization at NPCs, characterized the factors involved in mediating R-loops-NPC interaction, and assessed the consequences of this association on R-loop fate and genome homeostasis.

This work is a result of a collaboration with other two groups; Karine Dubrana's team (CEA, Fontenay-aux-Roses) provided the live imaging data, while Vincent Géli team (CRCM, Marseille) performed the NPC-ChIP-seq experiment. My role in this project has consisted in (i) building and setup of the genetic systems described in this manuscript, including most yeast strains and plasmids (Figs. 1-5 and S1-4) (ii) the determination of R-loop dependency and the screen of the candidate factors involved in the relocation mechanism through biochemical assays (Fig. 2-4, S2) , (iii) the assessment of the implication of NPC association on R-loop dependent genetic instability (Fig. 5, S1; S4), and (iv) data visualization and writing of the manuscript first draft.

A R-loop sensing pathway mediates the relocation of transcribed genes to nuclear pore complexes

Arianna Penzo¹, Marion Dubarry², Clémentine Brocas³, Myriam Zheng¹, Ophélie Lautier¹, Marie-Noëlle Simon², Vincent Geli², Karine Dubrana³ & Benoit Palancade¹

¹ Université Paris Cité, CNRS, Institut Jacques Monod, F-75013 Paris, France.

² Marseille Cancer Research Center (CRCM), U1068, Institut National de la Santé et de la Recherche Médicale (INSERM), UMR7258, Centre National de la Recherche Scientifique (CNRS), Aix Marseille University, Institut Paoli-Calmettes, Equipe Labélisée Ligue, 13273 Marseille, France.

³ iRCM, Institut Jacob, CEA, INSERM U1274, UMR Stabilité Génétique Cellules Souches et Radiations, Université Paris-Saclay, Université Paris Cité, 92265 Fontenay-aux-Roses, France.

Correspondence: Benoit Palancade, benoit.palancade@ijm.fr

ABSTRACT

Nuclear organization has emerged as a critical player in the control of genomic processes, including transcription. In this context, nuclear pore complexes (NPCs) have increasingly recognized interactions with the genome, as exemplified in yeast, where they bind inducible genes and damaged genomic regions, positively impacting their fate. To investigate the pathways fostering chromatin association with NPCs, we have combined genome-wide approaches with live imaging of individually-tagged model loci. Strikingly, ChIP-seq analyses of NPC-bound genes revealed a strong correlation between pore association and the propensity to accumulate co-transcriptional R-loops, which are genotoxic structures forming through hybridization of nascent RNAs with their DNA templates. Manipulating cis- or trans-acting regulators of hybrid formation further demonstrated that R-loop accumulation *per se*, rather than high transcription or R-loop-associated genetic instability, is the primary trigger for relocation to NPCs. Mechanistically, R-loop-dependent repositioning involves the recognition of displaced ssDNA moieties by the ssDNA-binding protein RPA, and SUMO-dependent interactions between RPA and NPC-associated factors. Preventing R-loop-dependent NPC localization leads to lethality, while permanent NPC tethering of a model hybrid-prone sequence attenuates R-loop-dependent genetic instability. Remarkably, this novel relocation pathway involves similar molecular factors as those required for the association of stalled replication forks or eroded telomeres with NPCs, suggesting the existence of convergent mechanisms for sensing transcriptional and genotoxic stresses.

INTRODUCTION

The three-dimensional organization of the nucleus plays a central role in the regulation of several genomic transactions, including transcription and DNA repair, thus contributing to the maintenance of genome homeostasis. Among the structural components of the nucleus that shape its compartmentalization are nuclear pore complexes (NPCs), which are conserved, megadalton-sized multiprotein assemblies embedded within the nuclear envelope and built from multiple copies of ~30 subunits called nucleoporins (Nups; Beck and Hurt, 2017). While scaffold Nups sub-complexes delineate a central channel in which nucleo-cytoplasmic exchanges occur, peripheral components, *i.e.* the cytoplasmic filaments and the nuclear basket, extend contacts towards the cytoplasm and the nucleoplasm. In this way, nuclear pores notably establish interactions with specific regions of the genome, beyond their canonical role in the selective transport of proteins and RNAs (Fernandez-Martinez and Rout, 2021). This is well exemplified in budding yeast, where several inducible loci, including galactose-activated and heat shock genes, relocate to the nuclear periphery and associate with NPCs upon transcriptional activation (Brickner and Walter, 2004; Casolari et al., 2004, 2005; Abruzzi et al., 2006; Cabal et al., 2006; Dieppois et al., 2006; Schmid et al., 2006; Taddei et al., 2006, reviewed in Dieppois and Stutz, 2010; Ibarra and Hetzer, 2015; Sumner and Brickner, 2021). While proximity to the pore may couple transcription with mRNA processing and export, thus positively impacting gene expression (Brickner and Walter, 2004; Taddei et al., 2006; Saik et al., 2020), the physiological significance of gene repositioning a.k.a. “gene gating” (Blobel, 1985) still remains debated. Strikingly, DNA lesions, *e.g.* unreparable DNA double-strand breaks (DSBs; Nagai et al., 2008), eroded telomeres (Khadaroo et al., 2009), or challenged replication forks (Su et al., 2015; Kramarz et al., 2020), also cause the relocalisation of genomic regions to nuclear pores in yeast cells. In these situations, NPC association has been shown to locally impact DNA repair pathway choices, thus contributing to the maintenance of genetic integrity (reviewed in Freudenreich and Su, 2016). Remarkably, NPC repositioning events similarly target transcribed or damaged loci in several distant species, in which loss-of-function of nucleoporins can trigger changes in gene expression or DNA damage, pointing to the functional importance of pore-chromatin interactions (reviewed in Pascual-Garcia and Capelson, 2021; Lamm et al., 2021; Freudenreich and Su, 2016).

Repositioning typically involves diffusive or active motion of chromatin domains within the nucleus (Lamm et al., 2021; Pascual-Garcia and Capelson, 2021). Anchoring of specific genomic regions to NPCs is further achieved through protein-protein contacts involving DNA- and NPC-bound factors, for instance the mediator and TREX-2 complexes, whose association bridges the promoter of activated *GAL* genes with the nuclear basket (Cabal et al., 2006; Schneider et al., 2015). The interactions between chromatin and NPCs also involve SUMOylation, a post-translational modification that relies on the covalent addition of the SUMO (small ubiquitin-like modifier) polypeptide to its protein targets. SUMO moieties are covalently coupled to lysine residues through an enzymatic cascade

involving an E1 activating enzyme, an E2 conjugating enzyme and an E3 ligase; removed through the action of SUMO-proteases; and recognized by SIM (SUMO interaction-motifs)-containing proteins, including SUMO-targeted ubiquitin ligases (STUbLs), which direct modified substrates to proteasomal degradation (reviewed in Geiss-Friedlander and Melchior, 2007; Vertegaal, 2022). Remarkably, the SUMO pathway is itself compartmentalized within the yeast nucleus, with the three SUMO ligases (Siz1, Siz2, Mms21) being localized in the nucleoplasm, while the essential SUMO-protease Ulp1 and the Slx5/Slx8 STUbL are restricted to NPCs (Palancade and Doye, 2008; Nagai et al., 2008). In this respect, highly-expressed or inducible genes harbor high-levels of SUMOylation (Rosonina et al., 2010), and their repositioning to NPCs requires both the SUMO E3 ligase Siz2 and the SUMO-protease Ulp1 (Texari et al., 2013; Saik et al., 2020; Ptak et al., 2021). Similarly, SUMOylation waves occur at DNA lesions (Cremona et al., 2012; Psakhye and Jentsch, 2012), and the relocalisation of damaged chromatin to NPCs involves SUMO ligases and SIM-containing NPC-associated factors (Horigome et al., 2016; Churikov et al., 2016; Ryu et al., 2016; Whalen et al., 2020; Kramarz et al., 2020). Beyond these common signals, it remains however to be understood whether gene gating and damage relocalisation utilize redundant or overlapping pathways.

Another process connecting high transcriptional activity to genetic instability is the formation of R-loops, which are three-stranded structures involving the annealing of nascent RNAs onto their DNA templates, thus displacing single-strand DNA moieties. In yeast, R-loops preferentially form at highly-expressed loci (Wahba et al., 2016) and their unscheduled accumulation ultimately leads to replication stress and DSBs accumulation (reviewed in García-Muse and Aguilera, 2019). While the formation of R-loops is sterically prevented by the coating of the transcripts with RNA-binding factors, such as the THO complex (Huertas and Aguilera, 2003) and the spliceosome (Bonnet et al., 2017), their removal from the genome involves dedicated enzymes, including ribonucleases of the RNase H family (Cerritelli and Crouch, 2009) and DNA:RNA helicases (e.g. Sen1/Senataxin; Mischo et al., 2011). How R-loops are detected and handled within the nuclear environment is however poorly understood. Notably, coating by the ssDNA-binding complex RPA (replication protein A) has been proposed to sense R-loops and possibly promote their removal through recruitment of RNase H1 in human cells (Nguyen et al., 2017). The idea that the NPC could also function in hybrid metabolism further arose from genetic screens which scored R-loop accumulation and R-loop-dependent genetic instability in nucleoporin mutants (Chan et al., 2014; García-Benítez et al., 2017). Strikingly, artificial tethering of an R-loop forming reporter to the nuclear pore can attenuate hybrid levels (García-Benítez et al., 2017). Altogether, these findings prompted us to explore whether R-loops themselves could act as a signal for repositioning to NPCs. By combining biochemical and live imaging approaches, we further examined the signals and pathways potentially mediating the association of R-loops with nuclear pores, in light of our knowledge of chromatin-NPC interactions. Finally, we investigated the functional impact that proximity to the pore could exert on R-loop metabolism.

RESULTS

Genome-wide association between R-loop formation and NPC localisation

In order to investigate the relationships between R-loop formation and gene positioning, we set out to compare the localization of genomic NPC contact sites to available maps of DNA:RNA hybrid distribution (Wahba et al., 2016; Aiello et al., 2022). For this purpose, we performed chromatin immunoprecipitation coupled to sequencing (ChIP-seq) using as a bait a functional, myc-tagged version of the scaffold nucleoporin Nic96 (**Fig. 1A**). Several NPC contact sites were observed within protein-coding genes, which were in average more highly-transcribed than the rest of the genome (**Fig. S1A**). Since DNA:RNA hybrid formation also correlates with transcription (**Fig. S1B**, Wahba et al., 2016), we restricted our analysis to the most highly-expressed genes, which were further categorized according to their intron content, a *cis*-acting modulator of R-loop formation (Bonnet et al., 2017). In this way, we were able to compare NPC association between two equally-sized groups of genes with similar transcription rate, gene length and base content (**Fig. S1C**, **Table S1**), but distinctive R-loop levels (**Fig. S1D-E**). Strikingly, R-loop-prone, intronless loci displayed higher Nic96 occupancy over their gene bodies as compared to their R-loop-depleted, intron-containing counterparts (**Fig. 1B-C**). Furthermore, the extent of Nic96 enrichment over intronless loci correlated with their propensity to form R-loops (**Fig. 1B**, *top panel*, genes ranked by R-loop levels). To confirm this finding, the same analysis was performed on independent datasets of NPC-bound genes, which were previously obtained using distinct scaffold nucleoporins as baits in ChIP-on-chip experiments (**Fig. 1D-E**; Van de Vosse et al., 2013). Similarly, highly-expressed, R-loop-forming intronless genes showed enhanced association with Nup170 and Nup157 as compared to the intron-containing group (**Fig. 1D-E**). Of note, the fact that NPC association is less pronounced for intron-containing loci, albeit similarly transcribed as the intronless group, rules out that the detected signals reflect the intrinsic bias of ChIP experiments for highly-transcribed regions (Teytelman et al., 2013). Overall, these genome-wide analyses rather indicate that the propensity of genes to form R-loops correlates with their association with NPCs.

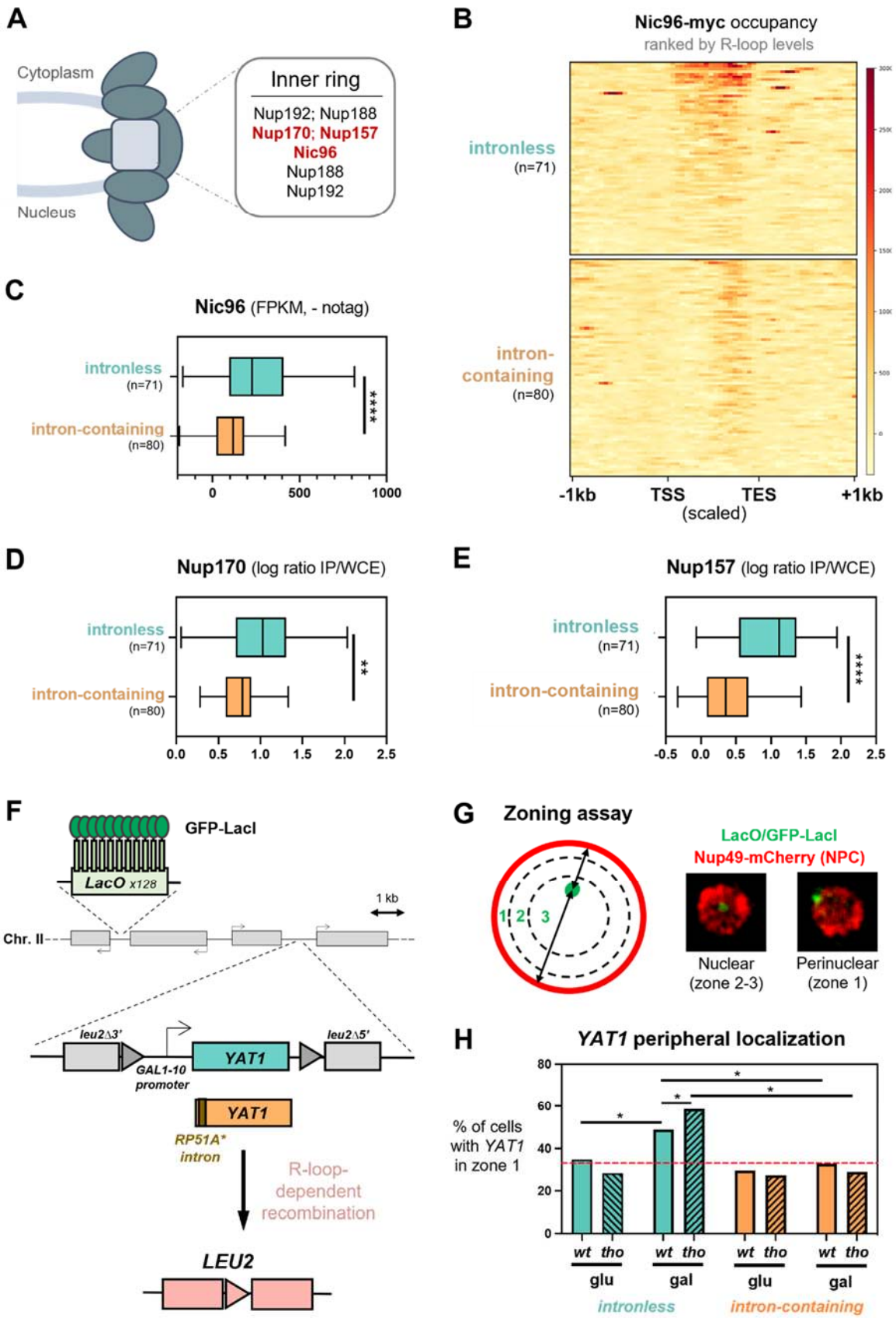


Figure 1. R-loops can be a signal for repositioning to NPCs. **A**, Schematic representation of the yeast nuclear pore complex (NPC) showing sub-complexes of nucleoporins as bubbles. The inner ring nucleoporins highlighted in red were used as baits in ChIP-seq (Nic96, this study) and ChIP-on-chip (Nup170, Nup157; Van de Vosse et al., 2013). **B**, Heatmap analysis of Nic96 occupancy at highly-transcribed intronless and intron-containing genes, aligned at their Transcription Start Site (TSS) and Transcription End Site (TES). Only the regions between the TSS and the TES are scaled. Genes are grouped based on their intron content and ranked according to their R-loop levels (as measured in Wahba et al., 2016). **C**, Average Nic96 occupancy at highly-transcribed intronless and intron-containing genes (FPKM, no tag control subtracted). **D-E**, Nup170 or Nup157 enrichments at intronless and intron-containing highly-transcribed genes, represented as the average \log_2 (IP/whole cell extract) for all the probes covering a given genomic feature. **F**, Principle of the *YAT1* integrated reporter construct. Either the intronless or the intron-containing version of the *YAT1* transgene, under the control of the *GAL1-10* promoter, are flanked by direct *leu2* repeats to allow quantification of R-loop-dependent recombination events. The reporter is integrated at the chromosome II *GAL* locus, which also contains an array of Lac operator (LacO) repeats for microscopy visualization. **G**, Principle of the zoning assay. GFP-LacI binding to the LacO array allows to measure the distance of the locus of interest, which appears as a bright green dot, relative to the nuclear envelope, stained by the Nup49-mCherry nucleoporin. The nucleus is divided in three equivolometric regions and only the dots localizing at the outermost region are scored as peripheral (zone 1), while others are considered nucleoplasmic (zones 2-3). **H**, Fraction of cells (%) showing intronless or intron-containing *YAT1* in zone 1, in wt or *tho* (*mft1* Δ) mutant cells grown in glycerol-lactate medium and further treated with either glucose (glu) or galactose (gal) for 5h. Between 388 and 672 cells were counted in 2 independent experiments for each condition. The red dashed bar indicates the expected value for a randomly-distributed locus.

* $p < 0.05$; ** $p < 0.01$; **** $p < 0.0001$; Mann-Whitney-Wilcoxon test. See also **Fig. S1**.

A reporter assay to probe R-loop-dependent relocalisation to nuclear pore complexes

To directly assess whether hybrid formation triggers gene localization at NPCs, we engineered a reporter locus in which R-loop accumulation can be locally modulated, and further recorded its nuclear position through live imaging. For this purpose, we inserted the GC-rich, *bona fide* R-loop forming *YAT1* ORF within the chromosome II *GAL* locus, under the control of the inducible *GAL1-10* promoter (**Fig. 1F**). High levels of transcription were previously shown to trigger R-loop formation on the *YAT1* sequence, both *in vitro* and *in vivo* (Bonnet et al., 2017). To further enhance hybrid accumulation at the reporter locus, we introduced the same construction in a mutant of the THO complex (*mft1Δ*, hereafter labelled *tho*), which triggers R-loops and R-loop-dependent genetic instability on the *YAT1* gene (Chávez et al., 2001; Bonnet et al., 2017). Conversely, to locally attenuate hybrid formation, we inserted at the 5' of the *YAT1* transgene a short artificial intron, which alleviates R-loop formation and R-loop-associated genotoxicity in *cis* (Bonnet et al., 2017). In these different strains, direct repeats flanking the *YAT1* reporter permitted the quantification of R-loop-dependent recombination events, which reconstitute a functional *LEU2* prototrophy marker (**Fig. 1F**). Importantly, this assay confirmed that the integrated reporter exhibits increased R-loop-dependent genetic instability in *tho* mutants and that this phenotype is rescued by the insertion of the intron (**Fig. S1F**), in agreement with our previous observations using plasmid-borne versions of the same constructs (Bonnet et al., 2017). An array of LacI-GFP-bound tandem repeats of the bacterial Lac operator, inserted at the same locus, finally allowed to visualize the position of the reporter gene with respect to NPCs, which are detected owing to a mCherry-tagged version of the Nup49 nucleoporin (**Fig. 1F**). Peripheral localization was assigned to the loci positioning in the most external of the three equivolumetric zones in which the nucleus is segmented for image analysis (**Fig. 1G**), as previously described (see Material and Methods). As expected, in the absence of transcription (glucose-containing medium), the reporter gene appeared randomly distributed in the nucleus, in *wt* and *tho* mutant cells, regardless of its intron content (**Fig. 1H**). However, upon transcriptional activation (galactose-containing medium), the fraction of cells in which intronless *YAT1* localises at the nuclear periphery increased significantly in *wt* cells, a phenotype further enhanced in the R-loop-accumulating *tho* mutant (**Fig. 1H**). Strikingly, the presence of the intron completely abolished *YAT1* relocation to the nuclear envelope in both *wt* and *tho* mutant cells (**Fig. 1H**). Altogether, these results indicate that R-loop accumulation triggers relocalisation of a model inducible locus to NPCs, mirroring the genome-wide observations reported above for constitutively expressed genes.

Stress-induced transcriptional activation leads to R-loop dependent relocation to the NPC

To expand our findings, we wondered whether stress situations involving the coordinated transcriptional induction of multiple responsive loci would similarly result in their R-loop-dependent repositioning to NPCs. To this aim, we focused our attention on the heat shock (HS) response since it induces high loads of transcription at heat shock genes (e.g. *HSP104*), some of which were

previously reported to relocate to NPCs upon activation (Dieppo et al., 2006; Brickner et al., 2012). Intriguingly, it had been reported that association of heat shock-activated loci with NPCs is enhanced in cells lacking the THO complex (Rougemaille et al., 2008; Mouaikel et al., 2013). To interfere with hybrid accumulation at HS genes in *trans*, we combined the R-loop-accumulating *tho* mutant with the overexpression of human RNase H1 (RNH1), a classical strategy to probe R-loop-dependent phenotypes (Wahba et al., 2011; Bonnet et al., 2017). The nuclear localisation of the HS-induced *HSP104* locus was further scored by microscopy as above, using strains in which LacO arrays are inserted downstream of the gene (Rougemaille et al., 2008; **Fig. 2A**). In this assay, HS triggered the relocalisation of the *HSP104* gene to the nuclear periphery in a fraction of the *wt* cells, and this phenotype was enhanced in the *tho* mutant (**Fig. 2B**), in accordance with previous studies (Dieppo et al., 2006; Rougemaille et al., 2008; Brickner et al., 2012). Remarkably, RNase H overexpression suppressed HS-induced *HSP104* relocalisation in both *wt* and *tho* mutant cells (**Fig. 2B**).

To complement these findings, we took advantage of previous observations reporting that the enhanced peripheral localization of HS genes in *tho* mutants is reflected by their biochemical co-fractionation with NPCs in heavy chromatin isolates (Rougemaille et al., 2008, **Fig. 2C**). We indeed confirmed the increased occurrence of *HSP104* in heavy chromatin fractions obtained from *tho* mutants, and further observed that this phenotype was suppressed by RNH1 overexpression (**Fig. 2D**), in agreement with our microscopy observations. Notably, the co-fractionation appeared to be specific of HS-induced loci, as it was not detected for a constitutively transcribed gene (*ACT1*) or an intergenic region (**Fig. S2A**). To ensure that the suppression of NPC relocalisation was not caused by indirect effects of RNH1 overexpression, we engineered an alternative construct in which the removal of a previously identified mitochondrial targeting signal (Cerritelli and Crouch, 2009) and the utilization of an inducible promoter alleviated the growth defects observed with the use of the complete sequence (**Fig. S2B**). Short-term induction of RNH1 overexpression similarly suppressed *HSP104* co-fractionation with NPCs in *tho* mutants (**Fig. S2C**), validating our previous observations.

To determine whether other genes of the HS regulon similarly display R-loop-induced repositioning, we further mapped HS-induced genomic NPC contact sites by ChIP-seq, using the same nucleoporin bait as above (Nic96). HS triggered the appearance of extended regions of contact between the NPC and gene bodies in *wt* cells, with increased Nic96 enrichment in the *tho* mutant, as exemplified in **Fig. 2E**. When specific NPC-associated peaks were ranked by size, this increase was deeply marked for long regions (>1kb; **Fig. 2F**). Gene ontology analysis of the genes interested by these extended contacts finally revealed an over-representation of heat-shock responsive loci (GO: “*protein folding*”, $p= 1.5e-4$; “*response to temperature stimulus*”, $p= 1.25e-3$). Altogether, these data indicate that R-loop formation can also act as a trigger for gene relocalisation to NPCs in the case of a coordinated transcriptional response impacting several distant loci.

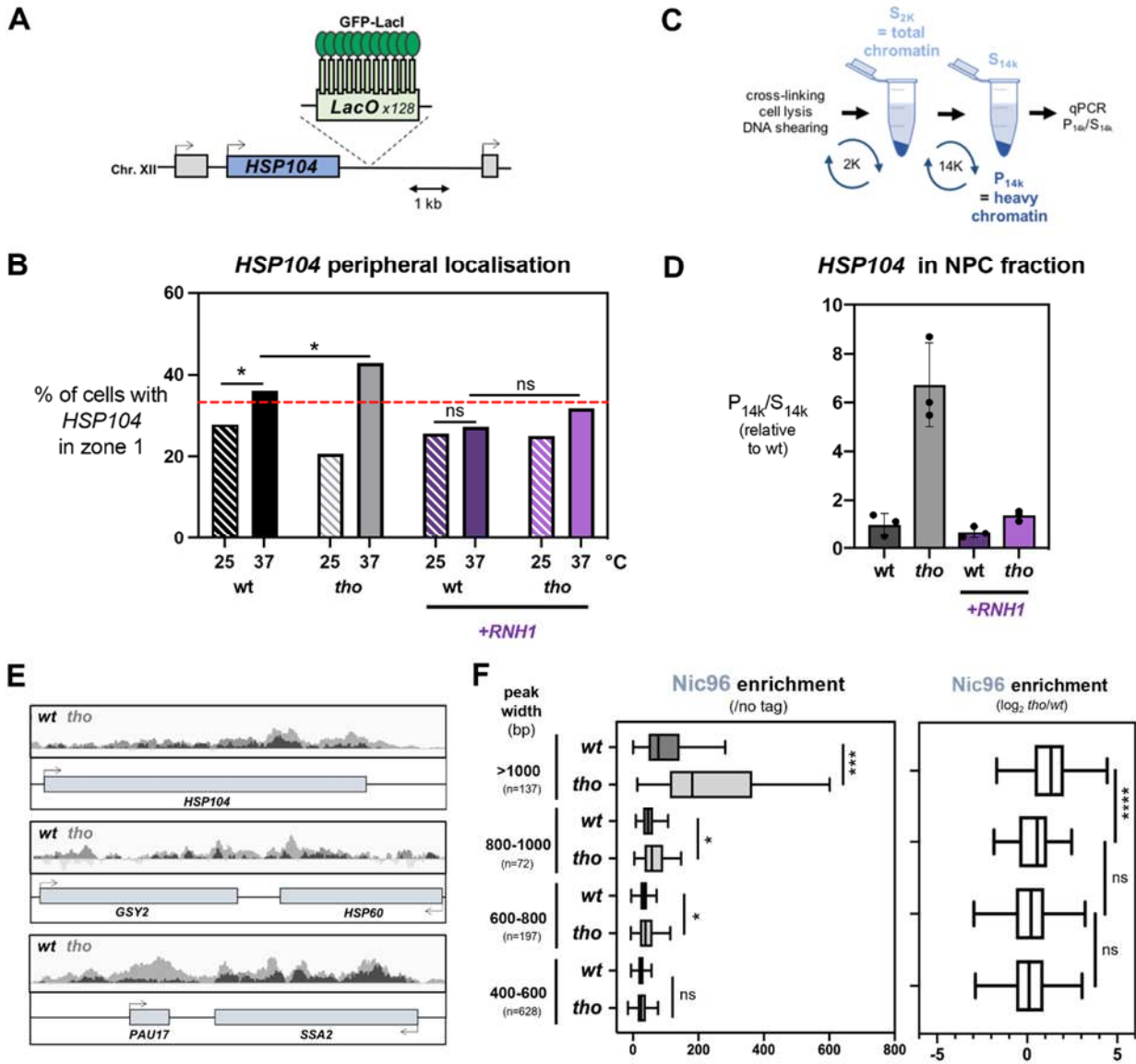


Figure 2. Heat shock-induced transcriptional activation leads to R-loop dependent relocalisation to the NPC. **A**, Schematic representation of the genomic *HSP104* locus with the integrated LacO array used for microscopy visualization. **B**, Fraction of cells (%) showing *HSP104* in zone 1, in wt or *tho* (*mft1Δ*) mutant cells transformed with either an empty vector or a GPD-hsRNH1 construct (+*RNH1*), grown at 25°C or heat-shocked at 37°C for 15min. Between 64 and 133 cells were counted for each condition. The red dashed bar indicates the expected value for a randomly-distributed locus. **C**, Principle of the differential chromatin fractionation procedure. The presence of the gene of interest in the pellet (P_{14k}) and supernatant (S_{14k}) fractions is evaluated by qPCR. **D**, qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from wt or *tho* (*mft1Δ*) mutant cells transformed with either an empty vector or a GPD-hsRNH1 construct (+*RNH1*), and heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean±SD, n=3, relative to wt). **E**, Integrative Genomics Viewer (IGV) representative screenshots of genes displaying Nic96 enrichment (no tag subtracted) in wt or *tho* (*mft1Δ*) mutant cells heat-shocked at 37°C for 15min. **F**, Nic96 enrichment (*left panel*, no tag subtracted; *right panel*, log₂[*tho*/wt]) in wt or *tho* mutant cells heat-shocked at 37°C for 15min. Nic96-bound regions identified through peak calling were categorized according to peak width (bp). The number of regions is indicated for each category.

*p<0.05; *** p<0.001; **** p<0.0001; ns, not significant; Mann-Whitney-Wilcoxon test. See also **Fig. S2**.

R-loop-dependent gating defines a novel NPC relocalisation pathway

We next wanted to investigate the relationships between this newly uncovered R-loop-dependent gene repositioning process and other described situations in which defined chromatin regions also interact with NPCs. Highly expressed and inducible genes were previously reported to associate with nuclear pores during the course of transcriptional activation, in a gene gating pathway requiring the NPC-bound TREX-2 (Transcription and Export) complex (Dieppo and Stutz, 2010; Schneider et al., 2015). However, *tho* mutants, in which we scored increased association of R-loop-forming loci with NPCs, have globally reduced transcription rates (Huertas and Aguilera, 2003; Rondón et al., 2003), as notably reported for the *YAT1* gene (Chávez et al., 2001; Bonnet et al., 2017). Furthermore, TREX-2 mutants (e.g. *sus1* Δ , *sac3* Δ , *sem1* Δ) triggered by themselves association of the *HSP104* locus with NPCs, as revealed by chromatin fractionation (Mathieu Rougemaille and Domenico Libri, personal communication). In light of the reported role of TREX-2 in preventing RNA-dependent genetic instability (González-Aguilera et al., 2008), including at the *YAT1* gene (Bonnet et al., 2017), it is likely that R-loop formation also acts as a signal for relocalisation to NPCs in these mutants. The fact that R-loop-dependent-repositioning still occurs in the absence of TREX-2 further supports that this relocalisation pathway is genetically distinguishable from canonical gene gating.

We next wondered whether R-loop-dependent DNA breakage or replication impairment could be the actual trigger for NPC repositioning, since both DSBs and blocked replication forks were previously reported to relocate to nuclear pores (Freudenreich and Su, 2016). To investigate the involvement of R-loop processing into DSBs in relocation, we combined the *tho* R-loop-accumulating mutant with the inactivation of enzymes described to trigger R-loop cleavage in yeast, i.e. the nucleotide excision repair factor Rad2 (Sollier et al., 2014), the DNA mismatch repair protein Mlh3 and the cytosine deaminase Fcy1 (Su and Freudenreich, 2017). None of the analyzed double mutants showed differences in *HSP104*-NPC co-fractionation as compared to the single *tho* mutant (**Fig. 3A**), indicating that R-loop-dependent damage does not cause R-loop-induced *HSP104* peripheral localization. Noteworthy, the genotoxicity of R-loops mainly arises from their encounter with the replication machinery (San Martín-Alonso et al., 2021). To assess whether R-loop-induced repositioning could stem from interferences with replication, we repeated the chromatin fractionation assay in cells synchronized in G0 through alpha-factor treatment. In these conditions, *tho* mutant cells still displayed an enhanced occurrence of the *HSP104* gene in the NPC fraction, similar to asynchronous cultures (**Fig. 3B**). Moreover, microscopy analyses revealed that a fraction of the cells displaying peripheral *HSP104* are unbudded, i.e. in the G1 phase of the cell cycle (data not shown), confirming that R-loop-dependent loci relocation to the periphery occurs independently from replication.

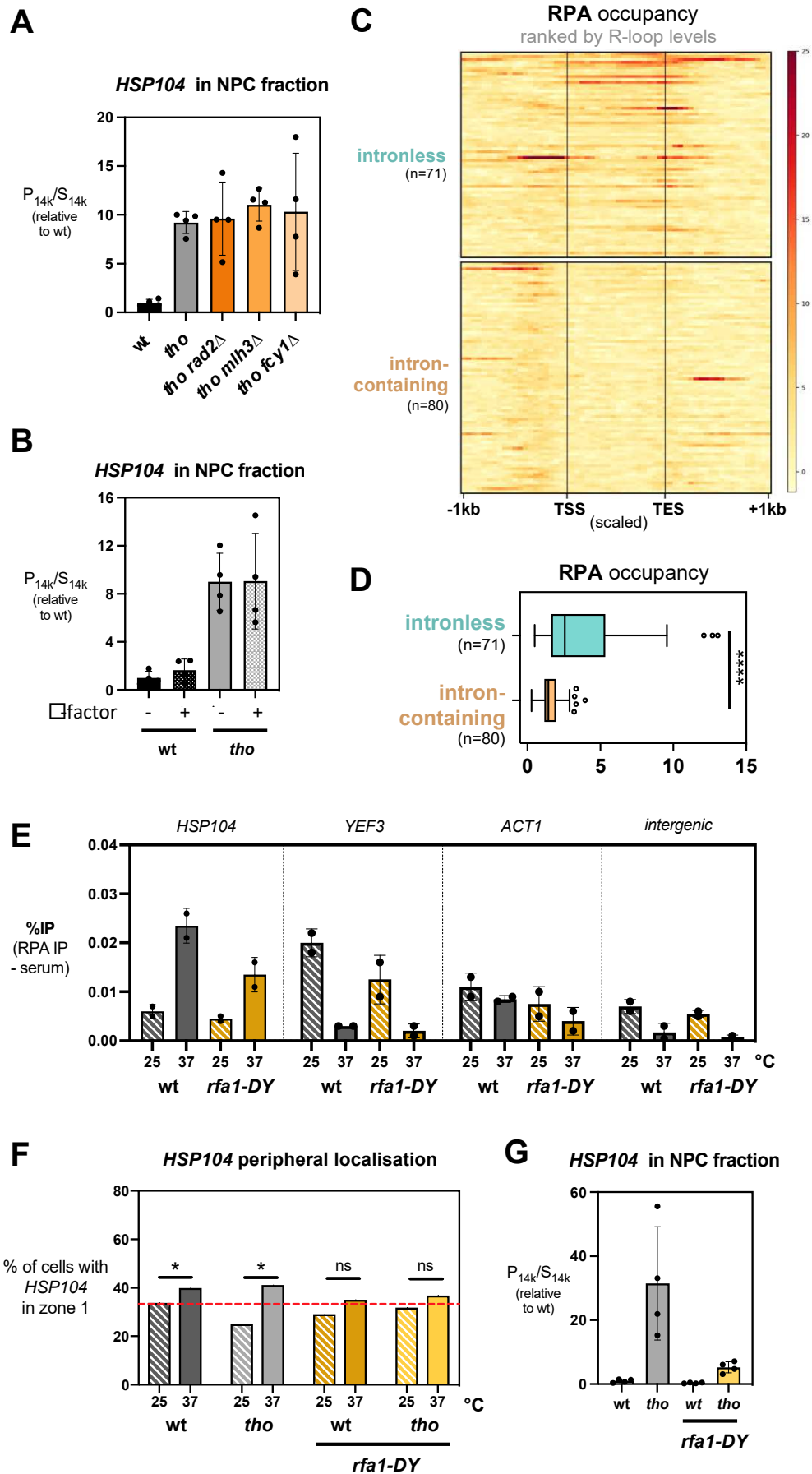


Figure 3. R-loop-dependent relocalisation to NPCs requires ssDNA coating by RPA. **A**, qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from the indicated strains heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean±SD, $n \geq 3$, relative to wt). **B**, qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from wt or *tho* (*mft1Δ*) mutant cells heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean±SD, $n=4$, relative to wt). When indicated, cells were arrested in G0 through alpha-factor treatment (α -factor). **C**, Heatmap analysis of RPA occupancy at the forward strand of highly-transcribed intronless and intron-containing genes, aligned at their Transcription Start Site (TSS) and Transcription End Site (TES), in wt cells arrested in G0 (strand-specific RPA ChIP-seq dataset from Reuswig et al., 2021). Only the regions between the TSS and the TES are scaled. Genes are grouped based on their intron content and ranked according to their R-loop levels (as measured in Wahba et al., 2016). **D**, Average RPA occupancy at the forward strand of highly-transcribed intronless and intron-containing genes (input subtracted). **E**, RPA occupancy analyzed by ChIP-qPCR on the indicated loci for wt or *rfa1-D228Y* (*rfa1-DY*) mutant cells, grown at 25°C or heat-shocked at 37°C for 15min (% of immunoprecipitation, control serum subtracted, mean±SD, $n=2$). **F**, Fraction of cells (%) showing *HSP104* in zone 1, in the indicated strains grown at 25°C or heat-shocked at 37°C for 15min. Between 70 and 226 cells were counted in 2 independent experiments for each condition. The red dashed bar indicates the expected value for a randomly-distributed locus. **G**, qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from the indicated strains heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean±SD, $n=4$, relative to wt).

* $p < 0.05$; **** $p < 0.0001$; ns, not significant; Mann-Whitney-Wilcoxon test. See also **Fig. S3**.

R-loop-dependent relocalisation to the nuclear pore requires ssDNA coating by RPA

Our data support that R-loop-dependent gating defines a novel NPC relocalisation pathway, in which increased transcription, damage formation or interference with replication are not the primary causes for repositioning. Among the distinctive structural features of R-loops that could rather be recognized prior to relocalisation are the displaced ssDNA moieties within these three-stranded structures. We therefore directed our attention to the main cellular ssDNA-binding complex, RPA (Replication Protein A), which was previously localized to transcribed genes in yeast and further associated with R-loop sensing and resolution in mammalian cells (Sikorski et al., 2011; Nguyen et al., 2017). To assess the specific presence of RPA at R-loop forming genes without the confounding effect of its replication-dependent recruitment, we took advantage of an available strand-specific RPA ChIP-seq dataset obtained from non-cycling yeast cells (Reuswig et al., 2021). By restricting our analysis of RPA occupancy to intronless and intron-containing highly-transcribed genes (same groups as **Fig. 1B**), we scored the presence of RPA stretches along intronless gene bodies (**Fig. 3C-D**), correlating with their R-loop levels and their association with NPCs (**Fig. 1B-C**). In contrast, RPA binding was nearly undetectable on their R-loop-depleted, intron-containing counterparts (**Fig. 3C-D**). Of note, RPA occupancy was similarly detected at forward and antisense strands in this ChIP assay (**Fig. 3C-D; Fig. S3A-B**), an expected pattern since yeast R-loops are typically smaller (~150bps; García-Pichardo et al., 2017) than sheared chromatin fragments (**Fig. S3C**).

To further confirm the correlation between RPA recruitment and R-loop-dependent gating, we used ChIP-qPCR to monitor RPA association to model loci in control and heat-shocked cells. RPA recruitment was indeed observed onto *YEF3*, a member of the intronless gene group used above, in control conditions, yet was abolished upon HS (**Fig. 3E**), which reportedly represses its transcription (Scholes and Lewis, 2020). In contrast, HS triggered RPA recruitment onto *HSP104* (**Fig. 3E**), concomitantly with the activation of this inducible locus. Importantly, RPA recruitment was reduced at both genes in a mutant impairing its association to ssDNA (*rfa1-D228Y* - Smith and Rothstein, 1995; Audry et al., 2015), testifying the specificity of the detected signal (**Fig. 3E**).

To investigate whether RPA recruitment to R-loops is required for NPC repositioning, we combined the *rfa1-D228Y* mutation with the inactivation of the THO complex and assessed *HSP104* localization upon transcriptional induction. Strikingly, impairing RPA association to ssDNA virtually suppressed the increase in *HSP104* peripheral localization scored by microscopy in *tho* mutant cells (**Fig. 3F**; compare *tho* and *tho rfa1-D228Y*). Consistently, the co-fractionation of *HSP104* with the nuclear pore was nearly abrogated upon decreased RPA binding (**Fig. 3G**). These data thus establish that ssDNA coating by RPA senses the formation of R-loops and mediates their relocalisation at nuclear pores.

SUMOylation events mediate R-loop-dependent NPC association

The establishment of contacts between transcribed chromatin and nuclear pores was previously reported to involve random sub-diffusion of the targeted locus within the nucleus, followed by its capture at the nuclear periphery by virtue of gene-NPC interactions (Sumner et al., 2021). We thereby asked whether dedicated factors could mediate the interaction between nuclear pores and RPA-bound R-loop-forming genes undergoing repositioning. Of note, RPA subunits were not previously identified in our proteomic analyses of nuclear pores (Bretes et al., 2014; Lautier et al., 2021), suggesting the existence of indirect or transient, labile interactions between this ssDNA-binding complex and NPCs. In light of the multiple reports indicating that SUMOylation, a highly-reversible post-translational modification, can contribute to NPC-chromatin interactions (Texari et al., 2013; Ptak et al., 2021; Horigome et al., 2016; Kramarz et al., 2020) and target the RPA subunit Rfa1 (Churikov et al., 2016; Whalen et al., 2020), we thereby wondered whether RPA could be SUMOylated concomitantly with R-loop gating. To test this hypothesis, we expressed a polyhistidine-tagged version of SUMO (His-SMT3) under the control of its endogenous promoter and purified SUMO-conjugates from yeast by denaturing affinity chromatography. Western-blot detection using Rfa1-specific antibodies did not reveal any modified bands in *wt* cells, yet uncovered a slower-migrating version of Rfa1 in a mutant of the NPC-associated deconjugating enzyme Ulp1 (*ulp1-333*), with a molecular weight compatible with mono-SUMOylation (**Fig. 4A**). Performing the same assay in a mutant strain expressing a non-SUMOylatable version of Rfa1, Rfa1-4KR (Dhingra et al., 2019), further confirmed that this species corresponds to mono-SUMOylated Rfa1 (**Fig. 4A**). Strikingly, RPA SUMOylation increased upon heat-shock induction (**Fig. 4A**), suggesting that this modification occurs concomitantly with R-loop relocalisation.

To further characterise the involvement of SUMOylation in this process, we assessed the impact of the inactivation of several components of the SUMO pathway (**Fig. 4B**) on R-loop-NPC association, with the same combination of biochemical and live imaging approaches as above. Remarkably, *HSP104* peripheral localization was completely suppressed in a mutant of the SUMO-ligase Mms21 (**Fig. 4C**), a subunit of the cohesin-like Smc5/6 complex (Zhao and Blobel, 2005). Consistently, *HSP104* co-fractionation with the nuclear pore was strongly alleviated in *mms21-11* mutant cells (**Fig. 4D**), while it remained unperturbed upon the double inactivation of the two main yeast SUMO-ligases Siz1 and Siz2 (**Fig. 4E**). Conversely, *HSP104* localization to NPCs was unchanged in a SUMO mutant unable to form poly-SUMO chains (*smt3-3KR*; **Fig. 4F**), suggesting that mono-SUMOylation events, as detected for Rfa1, are sufficient for repositioning. Consistently, preventing Rfa1 SUMOylation reduced the extent of *HSP104* peripheral localization (**Fig. 4G**) and co-fractionation with the nuclear pore (**Fig. 4H**). The fact that loss of Rfa1 SUMOylation (*rfa1-4KR*, **Fig. 4H**) does not fully phenocopy the RPA ssDNA-binding mutant (*rfa1-D228Y*, **Fig. 3G**) or the SUMO-ligase inactivation (*mms21-11*; **Fig. 4D**) suggests the existence of additional SUMOylation events, occurring downstream of RPA binding and involving Mms21 activity towards other factors, possibly associated with R-loops.

Unfortunately, we could not assess whether Rfa1 SUMOylation depends itself on the Mms21 SUMO-ligase, in light of the strict co-lethality between *MMS21* and *ULP1* inactivation (Zhao and Blobel, 2005).

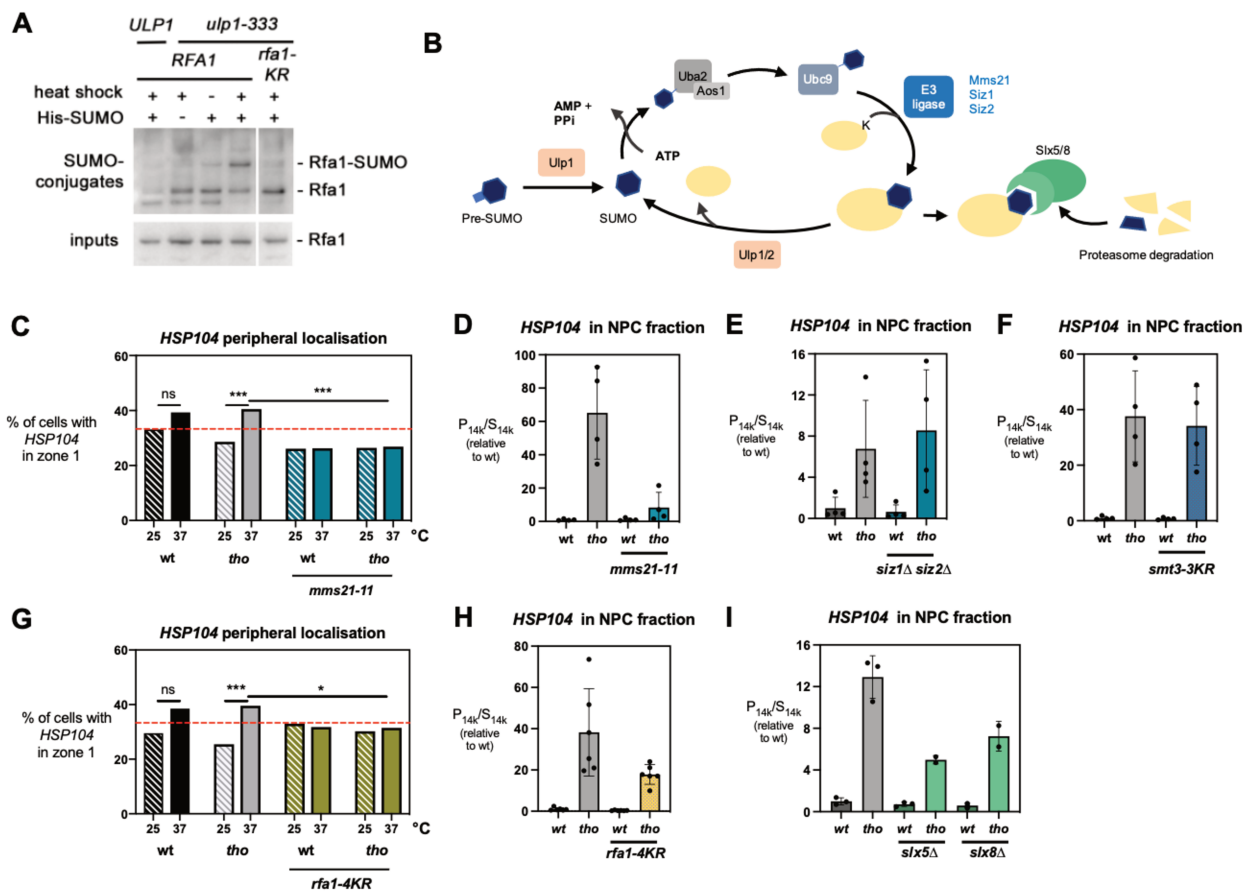


Figure 4. The SUMOylation pathway is involved in R-loop-dependent repositioning to NPCs. **A**, Western blot detection of Rfa1 in input fractions (*bottom panel*) or purified SUMO-conjugates (*top panel*) obtained from the indicated strains. Cells carrying the *His-SMT3* (*His-SUMO*) construct as indicated were grown at 25°C or heat-shocked at 37°C for 15min (heat shock). The positions of unmodified and mono-SUMOylated Rfa1 are indicated. **B**, Overview of the components of SUMO pathway in *S. cerevisiae*. **C**, Fraction of cells (%) showing *HSP104* in zone 1 in the indicated strains grown at 25°C or heat-shocked at 37°C for 15min. Between 283 and 523 cells were counted in 2 independent experiments for each condition. The red dashed bar indicates the expected value for a randomly-distributed locus. **D-F**, qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from the indicated strains heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean \pm SD, n=4, relative to wt). **G**, Fraction of cells (%) showing *HSP104* in zone 1 in the indicated strains grown at 25°C or heat-shocked at 37°C for 15min. Between 166 and 364 cells were counted in 2 independent experiments for each condition. The red dashed bar indicates the expected value for a randomly-distributed locus. **H-I**, qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from the indicated strains heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean \pm SD, n \geq 2, relative to wt).

*p<0.05; *** p<0.001; ns, not significant; Mann-Whitney-Wilcoxon test.

Finally, to investigate the mechanisms by which R-loop-bound, SUMOylated RPA interacts with nuclear pores, we assessed whether repositioning required SUMO-interaction motifs (SIM)-containing NPC-associated factors, *i.e.* Slx5/Slx8, which were previously found to contribute to damage relocation to the nuclear periphery (Freudenreich and Su, 2016). Remarkably, inactivation of either of these two factors caused a decrease in *HSP104* co-fractionation with the nuclear pore (**Fig. 4I**).

Altogether, this body of evidence demonstrates a requirement for the SUMOylation pathway in mediating R-loop relocalisation and suggests that the anchoring of R-loop-forming genes at nuclear pores involves interactions between R-loop-bound, sumoylated RPA complexes and NPC SUMO-interaction motifs.

Gene repositioning to the nuclear pore has a protective effect against R-loop toxicity

We next wondered whether relocalisation of hybrid-forming genes at NPCs could affect R-loop fate and genetic stability. To this aim, we first assessed the fitness of double mutants combining the hybrid-accumulating *tho* mutation and the inactivation of the different factors uncovered here as mediating R-loop repositioning, *i.e.* the ssDNA-binding complex RPA, the SUMO-ligase Mms21 and the NPC-associated SUMO-interacting factor Slx8. Growth assays revealed a synergic growth defect of the *tho rfa1-D228Y*, *tho mms21-11* and *tho slx8Δ* double mutants as compared to each single inactivation at 30°C, which was even exacerbated at 37°C (**Fig. 5A**). In contrast, simultaneous loss-of-function of the two SUMO-ligases Siz1 and Siz2, which detectably impact cell fitness but do not impair R-loop relocalisation (**Fig. 4E**), did not aggravate the growth defects of the R-loop-forming *tho* mutant (**Fig. 5A**). Since none of these mutations appears to affect *HSP104* expression levels (data not shown), these observations point to a protective effect of the R-loop relocalisation pathway in conditions of R-loop accumulation.

To further investigate the consequences of NPC association on R-loop metabolism, we set out to monitor R-loop-dependent genetic instability upon persistent peripheral localization of a hybrid-forming locus. To this aim, we took advantage of the presence of LexA-binding sites (LexA-BS) downstream of the *YAT1* reporter system used above (**Fig. 1F**), and co-expressed a fusion of the bacterial LexA protein to the basket nucleoporin Nup60 to tether the locus to the nuclear pore (**Fig. 5B**), as previously achieved (Texari et al., 2013). Microscopy analyses validated high levels of peripheral localization for *YAT1* in LexA-Nup60-expressing cells, independently of the transcriptional level or the presence of the intron (**Fig. S4A**), confirming the efficiency of the tethering system. Of note, control cells expressing the LexA protein alone displayed increased localization of the reporter to the nuclear envelope upon transcriptional activation (galactose medium), a phenotype enhanced in the *tho* mutant (**Fig. S4A**), as expected from our previous observations (**Fig. 1H**). We thereby performed the recombination assay in glucose-containing medium, in which the *YAT1* locus was randomly localized in LexA cells. In these conditions, *wt* cells displayed low but detectable levels of

recombination, which were enhanced in the *tho* mutant (Fig. 5C, left panel). This recombination phenotype was also suppressed by the presence of the intron (Fig. S4B), arguing it arises from R-loop-dependent genetic instability, possibly driven by basal transcription from the *LEU2* promoter. Strikingly, recombination frequencies were significantly reduced when the gene was persistently attached to the nuclear pore (Fig. 5C, right panel). Altogether, these observations support a protective effect of the nuclear pore against R-loop dependent genetic instability.

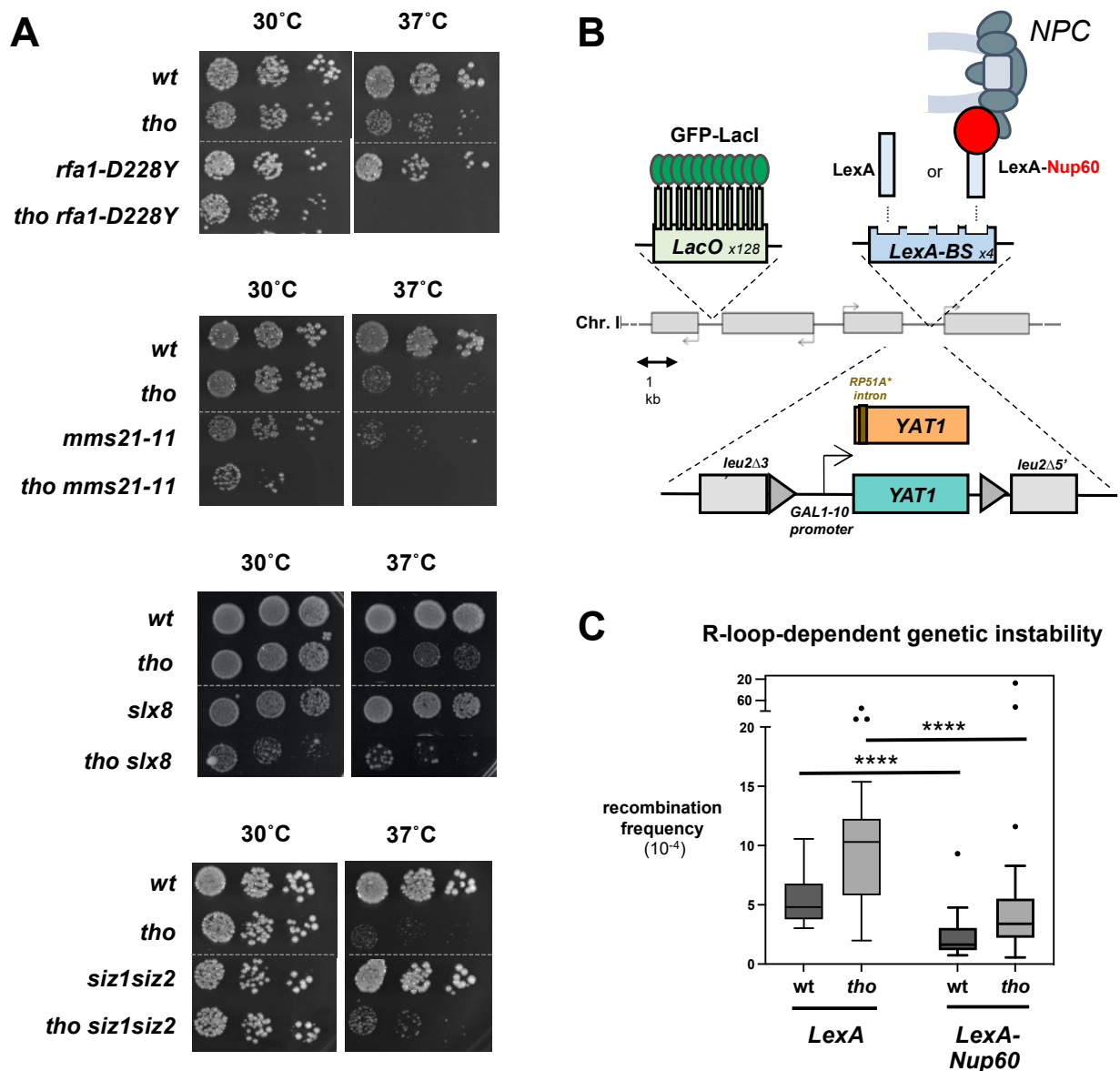


Figure 5. NPC association alleviates R-loop toxicity. **A**, Serial dilutions of the indicated strains were grown at the indicated temperatures on rich medium (YPD). **B**, Principle of the tethering assay. Either the intronless or the intron-containing version of the *YAT1* transgene, under the control of the *GAL1-10* promoter, are flanked by direct *leu2* repeats to allow quantification of R-loop-dependent recombination events. The reporter is integrated at the chromosome II *GAL* locus, which also contains an array of *LacO* repeats for microscopy visualization and LexA-binding sites. Expression of the LexA-Nup60 fusion ensures the permanent tethering of the locus to NPCs. **C**, Recombination frequencies (fraction of Leu⁺ prototrophs, x 10⁴; *wt* n=16, *tho* n=20) were calculated for the indicated strains as described in Material and Methods.

**** p<0.0001; Mann-Whitney-Wilcoxon test. See also Fig. S4.

DISCUSSION

In this study, we demonstrate that the co-transcriptional formation of R-loops can trigger the relocalisation of highly-expressed or inducible genes to NPCs (**Fig. 1, 2**). Our data support a model in which the coating of ssDNA by RPA and Mms21-dependent mono-sumoylation events allow the sensing of R-loops and their association to nuclear pores, where they are bound by SIM-containing NPC partners (**Fig. 3, 4**). Proximity to the pore would then alleviate R-loop formation and/or genotoxicity (**Fig. 5**), thus allowing to maintain high transcription levels while preserving genetic integrity (**Fig. S4C**).

Our genome-wide analyses of NPC-associated genes first unveiled a correlation between nuclear pore association and the propensity to form R-loops (**Fig. 1B-E; Fig. 2E-F**). The generation of high-resolution R-loop maps, especially in conditions of stress or metabolic shift in which loci relocation is detected, could provide further insights into the mechanisms of hybrid sensing and NPC targeting. In this respect, recently described RNase H-based R-loop capture methods (Chédin et al., 2021; Aiello et al., 2022) shall provide greater sensitivity in mapping short-lived DNA:RNA hybrids, particularly in *tho* mutants where transcription is lower at certain loci (Huertas and Aguilera, 2003). Interestingly, the reported polarity of R-loops signals shows a bias towards the 3' end of yeast gene bodies (Wahba et al., 2016; Aiello et al., 2022), which is also observed in *tho* cells for Nic96 occupancy (**Fig. 2E, top and medium panels**) and NPC-associated heavy chromatin (Rougemaille et al., 2008; Mouaikel et al., 2013). Together with our observation that highly-expressed, R-loop-deprived intron-containing genes do not associate with NPCs (**Fig. 1**), these findings support the idea that NPC association is driven by R-loop patterns rather than transcriptional activity. Consistently, interfering with R-loop formation at model loci in *cis*, through the insertion of an intron (**Fig. 1H**) or in *trans*, via RNase H over-expression (**Fig. 2B, D**), similarly abrogated relocalisation at nuclear pores without down-regulating transcription. Although indirect effects of splicing or RNase H activity cannot be excluded, these observations support that R-loops are the primary cause of NPC repositioning. Whether R-loop formation also partakes in the multiple situations where transcriptional activation drives NPC relocalisation (Brickner et al., 2019) thus remains to be investigated.

Our genetic dissection of the signals and pathways underlying the relocalisation of hybrid-forming genes further supports that “R-loop gating” does not occur through DNA damage formation (**Fig. 3A**) or replication impairment (**Fig. 3B**). While we cannot formally exclude that DSBs could arise at R-loop-forming loci independently of the nucleases assessed here (**Fig. 3A**), it should be noted that the peripheral localization of DNA breaks requires Siz2-dependent poly-sumoylation (Horigome et al., 2016), in contrast with R-loop repositioning, which specifically involves Mms21-dependent mono-sumoylation (**Fig. 4C-E**). Our data rather suggest that it is the direct sensing of the R-loop structure itself through the ssDNA-binding RPA complex that mediates NPC relocalisation. Indeed, RPA is detected onto R-loop forming genes with a dependency for high transcription levels and DNA-binding

activity (**Fig. 3C-E**), in the absence of replication (**Fig. 3C-D**), and decreased RPA recruitment hinders NPC association (**Fig. 3F-G**). The labeling of RPA-coated R-loops by Mms21-dependent mono-sumoylation may further distinguish them from other types of ssDNA-exposing structures, providing competence for binding to the nuclear pore via SIM-containing NPC partners (**Fig. 4I**). Different mechanisms could restrict the activity of this SUMO-ligase to R-loop-forming regions. Mms21 is part of the cohesin-like Smc5/6 complex, which was shown to be recruited to DNA in a R-loop-dependent manner in the context of Epstein-Barr Virus infection (Yiu et al., 2022). Moreover, Mms21 enzymatic activity has been shown to be enhanced by ssDNA binding *in vitro* (Varejão et al., 2018). In line with the *modus operandi* of SUMO-ligases, which typically lack substrate specificity and trigger protein group SUMOylation once recruited (Psakhye and Jentsch, 2012), Mms21 could thereby target several distinct, yet-to-be-identified R-loop-bound factors at hybrid-forming loci. The fact that RPA sumoylation increases concomitantly with R-loop gating (**Fig. 4A**), and that loss of Rfa1 sumoylation substantially diminishes NPC association (**Fig. 4G-H**) suggest that RPA is one of the main targets in this process, although it remains to be determined whether its modification relies on Mms21 enzymatic activity. In the future, assessing the sumoylation of RPA or other R-loop-associated factors in distinct genetic situations leading to hybrid accumulation shall shed light on the pattern of modifications specifically leading to R-loop repositioning. Remarkably, Mms21-dependent mono-sumoylation is also required for the relocalisation of replication forks spanning triplet nucleotide repeat regions (TNRs), where stalled intermediates associate with NPCs prior to damage formation and checkpoint activation, thus alleviating repeat instability (Su et al., 2015; Whalen et al., 2020). RPA sumoylation has also been found to occur during the course of senescence-associated telomere repositioning to NPCs (Churikov et al., 2016). Although the R-loop gating pathway described here and alternative relocalisation processes are genetically distinguishable, they thus share common factors in the detection and the labeling of non-canonical ssDNA-containing structures, e.g. TNR-blocked replication forks or eroded telomeres.

Repositioning to NPCs is generally described as beneficial for gene expression and the maintenance of genome integrity. Tethering experiments indicate that the proximity to nuclear pores indeed alleviates R-loop accumulation (García-Benítez et al., 2017) and R-loop-dependent recombination (**Fig. 5C**). In contrast, preventing R-loop repositioning by interfering with the RPA/Mms21 pathway gives rise to synthetic lethality in hybrid-forming *tho* mutants (**Fig. 5A**). Similarly, inactivation of the Nse1 subunit of the Smc5/6 complex enhances the growth defect of R-loop-accumulating RNase H mutants (*rnh1Δ rnh201Δ*; Chang et al., 2018). It remains to be determined whether decreased cell fitness is actually caused by excess R-loop accumulation in these different situations. Remarkably, mutants of the Smc5/6 complex and of the Nup84 complex, which supposedly anchors Slx5/8 to NPCs (Nagai et al., 2008), similarly display increased levels of R-loops in yeast (Chan et al., 2014; Chang et al., 2018). How the NPC environment ultimately influences R-loop fate also requires future investigation. The vicinity to the pore could allow the mRNA to engage more rapidly in its export path,

facilitating its eviction from the transcription site and thus preventing R-loop formation, as previously proposed (García-Benítez et al., 2017). Alternatively, the association with NPCs could give access to dedicated R-loop-resolving enzymes, or other factors protecting these structures from breakage. However, none of our previous proteomic analyses of nuclear pores identified interactors related to R-loop metabolism, at least in *wt* cells (Bretes et al., 2014; Lautier et al., 2021). Finally, recruitment to NPCs could allow the removal of R-loop-bound proteins stabilizing the hybrids or promoting their processing into genotoxic intermediates. In this respect, Ulp1-mediated deSUMOylation, Slx5/8-dependent ubiquitination and degradation by the proteasome, which also resides at the nuclear basket (Niepel et al., 2013), could ensure such clearance events. Whether RPA removal from R-loops requires its proteolysis and further destabilize these three-stranded structures at NPCs remains to be explored. In a scenario combining these different models, a “pioneering” R-loop would form during early transcription cycles and rapidly engage the induced gene in pore association. This event would both allow the local destabilization of the R-loop and prevent the subsequent formation of additional DNA:RNA hybrids at this locus. R-loop-dependent repositioning would thus be particularly critical for inducible genes undergoing several rounds of transcription in a short timeframe, ensuring the high rate of RNA production necessary to sustain viability.

ACKNOWLEDGEMENTS

We are very grateful to Andrès Aguilera, Catherine Dargemont, Catherine Freudenreich, Doug Koshland, Stéphane Marcand, Françoise Stutz and Xiaolan Zhao for sharing yeast strains and plasmids; to Florent Dobé, Léa Doré and Simon Morin for technical help; to members of the ImagoSeine@IJM facility (Université Paris Cité, CNRS, Institut Jacques Monod) for flow cytometry analyses; to Mathieu Rougemaille and Domenico Libri for sharing unpublished results; and to Michel Werner and other members of the Palancade and Libri lab for fruitful discussions.

This work was supported by Agence Nationale pour la Recherche (ANR-18-CE12-0003, to B.P.), Fondation ARC pour la recherche contre le Cancer (projet ARC, to B.P.), Ligue Nationale contre le Cancer (comité de Paris, to B.P.), the IdEx Université de Paris (ANR-18-IDEX-0001, to B.P.), the BioSPC PhD program (to A.P. and M.Z.), Fondation pour la Recherche Médicale (to A.P.) and E.U.R.Gene (to A.P.).

MATERIAL AND METHODS

Yeast strains, plasmids and growth conditions

All *S. cerevisiae* yeast strains used in this study (listed in **Table S2**) were obtained by homologous recombination and/or successive crosses. The construction of the plasmids used in this study (listed in **Table S3**) was performed using standard PCR-based molecular cloning techniques and was checked by sequencing.

Cells were grown at the indicated temperature in standard yeast extract peptone dextrose (YPD) or synthetic complete (SC) medium supplemented with the required nutrients. For heat shock, cells were grown at 25°C in the appropriate medium to $OD_{600nm}=0.4-0.5$, quickly shifted at 37°C by addition of one volume of medium prewarmed at 49°C or one half-volume of medium pre-warmed at 61°C and further maintained at 37°C for 15 min in a water-bath. For experiments involving *GAL* promoter induction, cells were grown at 30°C in glycerol-lactate (GGL: 0.17% YNB, 0.5% ammonium sulfate, 0.05% glucose, 2% lactate and 2% glycerol) supplemented with the required nutrients prior to induction with glucose or galactose (2%) for 5h. For experiments involving tet-OFF RNase H1 induction, cells transformed with the RNH1-overexpressing construct were grown in SC medium supplemented with doxycycline (5µg/mL, Sigma) and induction was achieved by transferring cells in fresh medium without doxycycline for 16h. G0 cell cycle arrest was triggered at 25°C by three sequential additions of alpha-factor (2µg/ml, Biotem) spaced by 1h, prior to heat shock ; effective synchronization was verified by microscopy observation of cell morphology and flow cytometry. Growth assays were performed by spotting serial dilutions of exponentially-growing cells on solid medium and incubating the plates at the indicated temperatures.

Chromatin immunoprecipitation

For Nic96 ChIP, cells were crosslinked for 10 min with 1% formaldehyde at RT under agitation. Excess formaldehyde was quenched with glycine 0.25M, cells were washed with cold TBS, and pellets were frozen and conserved at -80°C. Cell pellets were resuspended in lysis buffer (50 mM HEPES pH7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) supplemented with 1 mM PMSF and anti-protease (cOmplete Tablet, Roche, 505649001) and lysed by bead-beating (Precellys® 24, Bertin). The lysate was sonicated with a Bioruptor (Diagenode) and centrifuged at 2'000 g for 15 min at 4°C. The supernatant was incubated with anti-Myc (9E10, Santa Cruz Biotechnology, sc-40) on a rotating wheel overnight at 4°C. Protein G coated magnetic beads (Dynabeads, ThermoFischer Scientific) were equilibrated in lysis buffer and 30 µl were added per sample and incubated on a rotating wheel for 2.5h at 4°C. Beads washes were as follow: twice with lysis buffer, twice with lysis buffer supplemented with 360 mM NaCl, twice with wash buffer (10 mM Tris-HCl pH8, 0.25 M LiCl, 0.5% IGEPAL, 1 mM EDTA, 0.1% Na-deoxycholate) and once with TE

(10 mM Tris-HCl pH8, 1 mM EDTA). Antibodies were uncoupled from beads with 100 μ l of Elution Buffer (50 mM Tris-HCl pH8, 10 mM EDTA, 1% SDS) for 10 min at 65°C. Decrosslinking was performed at 65°C overnight. After 30 min of RNase A treatment (20 μ g, Roche), proteins were digested by addition of 100 μ g of Proteinase K (Sigma) and incubated for 1.5h at 37°C. DNA was purified using the kit InnuPrep PCRpure (Eurobio) and eluted into 35 μ l of H₂O prior to library preparation and deep-sequencing.

For RPA ChIP, cells (25 OD) were cross-linked with formaldehyde 1% for 10 min at the same temperature used for the growth, in the presence of KPi 100 mM pH 7.5. Excess formaldehyde was quenched with glycine 0.27M, cells were washed with cold TBS, and pellets were frozen in liquid nitrogen. Frozen cells were lysed by bead beating in 1 mL of lysis buffer (50 mM HEPES pH7.4, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na deoxycholate, 4 μ g/mL pepstatin A, 180 μ g/mL PMSF and 0.25X protease inhibitor cocktail, Roche). Chromatin sonication was achieved using a Bioruptor (Diagenode) and the fragmented chromatin was recovered in the supernatant after a 5 min 2'500g centrifugation at 4°C. An aliquot was taken as an input fraction (2%) and the remaining sample was mixed overnight at 4°C with either a RPA-specific polyclonal antiserum available in the lab or a control serum (5 μ L each in the presence of 0.5% (w/v) bovine serum albumin and 47.5 μ g/mL salmon testes DNA. Dynabeads Protein G (10 μ L, ThermoFischer Scientific) were pre-coated for 1h in blocking buffer (lysis buffer as above containing 0.5% (w/v) bovine serum albumin and 47.5 μ g/mL salmon testes DNA) and mixed with the immunoprecipitation mixtures for 1h. Beads washes were as follow: twice with lysis buffer; twice with lysis buffer supplemented with 360 mM NaCl; twice with 10 mM Tris pH 8, 250 mM LiCl, 0.5% Nonidet-P40, 0.5% deoxycholate, 1 mM EDTA, and once with 10 mM Tris-HCl pH 8, 1 mM EDTA. Immunoprecipitated complexes were eluted for 10 min at 65°C in 100 μ L 50 mM Tris pH 8, 10 mM EDTA, 1% SDS deproteinized with 16 μ g proteinase K in the presence of 250 mM NaCl for 1h at 42 °C and for 30 min at 65 °C. Input and immunoprecipitated DNAs were purified with the Nucleospin Gel and PCR Clean-up kit (Macherey Nagel) and further quantified by real-time PCR.

Genome-wide sequencing

DNA libraries were prepared using NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs) according to manufacturer's specifications. Each library was quantified on Qubit with Qubit dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA, USA) and size distribution was examined on the Bioanalyser with High Sensitivity DNA chip (Agilent, Santa Clara, CA, USA), to ensure that the samples have the proper size, no adaptor contamination and to estimate sample molarity. Each library was diluted to 1 nM and then pulled together at equimolar ratio. Libraries were denatured according to the manufacturer's instruction and sequenced on a mid-output flow cell (130 M clusters) using the NextSeq 500/550 Mid Output kit v2.5 150 cycles kit (Illumina), in paired-end 75/75 nt mode, according to the manufacturer's instructions.

Bioinformatic analyses

Highly-expressed intronless and intron-containing gene groups were defined as before (Bonnet et al., 2017), with the exception that genes encompassing repeated sequences leading to ambiguous mapping were excluded from the analysis (see **Table S1** for the list of the 71 intronless genes and 80 intron-containing genes considered here).

Nic96 ChIP-Seq data quality was assessed using FastQC (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). Paired-end reads were mapped to *S. cerevisiae* genome (2011, SacCer3) with Bowtie2 (Langmead and Salzberg, 2012) and duplicated reads were removed using SAMTools rmdup (Li et al., 2009) to obtain Binary Alignment Mapped (BAM) file. Normalized FPKM (Fragments Per Kilobase per Million mapped fragments) were subjected to peaks calling using MACS2 (Zhang et al., 2008) with a q-value < 0.05. Peak annotation was done with the BEDTools ClosestBed (Quinlan and Hall, 2010) by determining the closest genomic feature to the summit position of the MACS2 peak. Normalized bigwig files (subtracting the no tag ChIP from the Nic96 ChIP) and heatmaps were obtained using deepTools2 (Ramírez et al., 2016).

Nup170/Nup157 Chip-on-chip datasets (Van de Vosse et al., 2013) and RPA ChIP-seq data from control, alpha-factor-arrested cells (Reusswig et al., 2021) were retrieved from the Gene Expression Omnibus database (accession numbers GSE36795 and GSE182203). Nup170 or Nup157 enrichments were represented as the average log₂ (IP/input) for all the probes covering a given genomic feature. Normalized bigwig files (subtracting the input from the RPA ChIP) and heatmaps were obtained using deepTools2 (Ramírez et al., 2016).

Chromatin fractionation

Differential chromatin fractionation was performed as previously described (Rougemaille et al., 2008). Cells (25 OD) were cross-linked with formaldehyde 0,9% for 10 min at 37°C in the presence of KPi 100 mM pH 7.5. Excess formaldehyde was quenched with glycine 0,27M, cells were washed with cold TBS, and pellets were frozen in liquid nitrogen. Cell pellets were resuspended in Lysis buffer (50mM Hepes pH 7.5; 150mM NaCl; 1mM EDTA pH8; 1% Triton X-100; 0.1% Na deoxycholate; 0.1% NP40; 0.1% SDS; 4 µg/mL pepstatin A; 180 µg/mL PMSF) and lysed by bead beating using a Fastprep (QBIOSYSTEMS). Following a centrifugation for 10 min at 12000rpm in a bench centrifuge at 4°C, the chromatin pellet was resuspended in lysis buffer and sonicated with a bioruptor (Diagenode). The lysate was then centrifuged 10 min at 2000rpm to remove cellular debris and 1mL of the supernatant, containing the chromatin, was further centrifuged for 10 min at 14000rpm to isolate the “heavy chromatin”. The pellet was washed in lysis buffer and resuspended in 100 µL elution buffer (50mM Tris pH8; 10mM EDTA; 1% SDS). To decrosslink, 50µg of proteinase K were added to 100µL of the supernatant (S14K) and the resuspended pellet (P14K), and the samples were incubated 30

min at 37°C and 1h at 65°C. DNA was purified with the QIAquick DNA purification kit (Qiagen) according to the manufacturer instructions and quantified by real-time PCR.

Nucleic acid analyses

DNA amounts in ChIP or chromatin fractionation samples were quantified by real-time PCR with a LightCycler 480 system (Roche) using SYBR Green incorporation according to the manufacturer's instructions. For ChIP experiments, the amount of DNA in the immunoprecipitated fraction was divided by the amount detected in the input to evaluate the percentage of immunoprecipitation (% of IP). For differential chromatin fractionation, the amounts of DNA in the P14K and S14K fractions were determined to evaluate the P14K/S14K ratios. The sequences of the primers used in this study are listed in **Table S4**.

Protein analyses

SUMO-conjugates were isolated from yeast cells expressing a polyhistidine-tagged version of SUMO using nickel agarose denaturing chromatography as previously described (Bretes et al., 2014), starting from 50mL of exponentially growing cells (OD_{600} =0.5-1). Protein samples were separated on 4–12% precast SDS-PAGE gels (ThermoFisher Scientific). Proteins were further detected by western-blot following transfer to PVDF membranes. The following validated antibodies were used: anti-RPA polyclonal antibody (same as for ChIP), 1:1000; anti-rabbit HRP secondary antibody (Jackson Immunoresearch), 1:5000. Images were acquired using chemiluminescent reagents (Supersignal, ThermoFisher) with a ChemiDoc MP Imaging System (Bio-Rad).

Live cell imaging

Exponentially-growing cells were harvested by centrifugation and mounted on slides for imaging. Live cell images were acquired using a wide-field inverted microscope (Leica DMI-6000B) equipped with Adaptive Focus Control to eliminate Z drift, a 100×/1.4 NA immersion objective with a Prior NanoScanZ Nanopositioning Piezo Z Stage System, a CMOS camera (ORCA-Flash4.0; Hamamatsu) and a solid-state light source (SpectraX, Lumencore), piloted by the MetaMorph software (Molecular Device). For GFP-mCherry two-colour images, 19 focal steps of 0.20 μ m were acquired sequentially for GFP and mCherry with an exposure time of 100 ms using solid-state 475- and 575-nm diodes and appropriate filters (GFP-mCherry filter; excitation: double BP, 450–490/550–590 nm and dichroic double BP 500–550/600–665 nm; Chroma Technology Corp.). Processing was achieved using the ImageJ software (National Institutes of Health). All the images shown are z projections of z-stack images.

Image analysis was realized with the FIJI software (Schindelin et al., 2012). Distances between loci and nuclear envelope were measured using either the PointPicker plugin (Meister et al., 2010) or a home-made macro. Cell cycle was determined on the basis of cellular morphology (unbudded cells:

G1; budded cells: S/G2). To determine the enrichment in zone 1 relative to a theoretical random position, or to compare enrichment in zone 1 of two different strains, we used a proportional analysis with a confidence limit of 95%.

Hyper-recombination assay

Independent clones were individually resuspended in 1mL glycerol-lactate medium, grown for at least 2h at 30°C and then induced with glucose or galactose (2%) for 5h. Cells were resuspended in water, appropriate dilutions were plated on SC medium lacking leucine to estimate the number of Leu+ recombinants, or SC medium to estimate cell survival, and plates were incubated for 3 days at 30°C. Hyper-recombination rates were defined as the proportion of Leu+ prototrophs estimated from at least 6 independent colonies.

BIBLIOGRAPHY

- Abruzzi, K.C., Belostotsky, D.A., Chekanova, J.A., Dower, K., Rosbash, M., 2006. 3'-end formation signals modulate the association of genes with the nuclear periphery as well as mRNP dot formation. *EMBO J.* 25, 4253–4262. <https://doi.org/10.1038/sj.emboj.7601305>
- Aiello, U., Challal, D., Wentzinger, G., Lengronne, A., Appanah, R., Pasero, P., Palancade, B., Libri, D., 2022. Sen1 is a key regulator of transcription-driven conflicts. <https://doi.org/10.1101/2022.02.09.479708>
- Audry, J., Maestroni, L., Delagoutte, E., Gauthier, T., Nakamura, T.M., Gachet, Y., Saintomé, C., Géli, V., Coulon, S., 2015. RPA prevents G-rich structure formation at lagging-strand telomeres to allow maintenance of chromosome ends. *EMBO J.* 34, 1942–1958. <https://doi.org/10.15252/embj.201490773>
- Beck, M., Hurt, E., 2017. The nuclear pore complex: understanding its function through structural insight. *Nat. Rev. Mol. Cell Biol.* 18, 73–89. <https://doi.org/10.1038/nrm.2016.147>
- Blobel, G., 1985. Gene gating: a hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 82, 8527–8529. <https://doi.org/10.1073/pnas.82.24.8527>
- Bonnet, A., Grosso, A.R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S.C., Janbon, G., Géli, V., de Almeida, S.F., Palancade, B., 2017. Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability. *Mol. Cell* 67, 608–621.e6. <https://doi.org/10.1016/j.molcel.2017.07.002>
- Bretes, H., Rouviere, J.O., Leger, T., Oeffinger, M., Devaux, F., Doye, V., Palancade, B., 2014. Sumoylation of the THO complex regulates the biogenesis of a subset of mRNPs. *Nucleic Acids Res.* 42, 5043–5058. <https://doi.org/10.1093/nar/gku124>
- Brickner, D.G., Ahmed, S., Meldi, L., Thompson, A., Light, W., Young, M., Hickman, T.L., Chu, F., Fabre, E., Brickner, J.H., 2012. Transcription factor binding to a DNA zip code controls interchromosomal clustering at the nuclear periphery. *Dev. Cell* 22, 1234–1246. <https://doi.org/10.1016/j.devcel.2012.03.012>
- Brickner, D.G., Randise-Hinchliff, C., Lebrun Corbin, M., Liang, J.M., Kim, S., Sump, B., D'Urso, A., Kim, S.H., Satomura, A., Schmit, H., Coukos, R., Hwang, S., Watson, R., Brickner, J.H., 2019. The Role of Transcription Factors and Nuclear Pore Proteins in Controlling the Spatial Organization of the Yeast Genome. *Dev. Cell* 49, 936–947.e4. <https://doi.org/10.1016/j.devcel.2019.05.023>
- Brickner, J.H., Walter, P., 2004. Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS Biol.* 2, e342. <https://doi.org/10.1371/journal.pbio.0020342>
- Cabal, G.G., Genovesio, A., Rodriguez-Navarro, S., Zimmer, C., Gadal, O., Lesne, A., Buc, H., Feuerbach-Fournier, F., Olivo-Marin, J.-C., Hurt, E.C., Nehrbass, U., 2006. SAGA interacting factors confine subdiffusion of transcribed genes to the nuclear envelope. *Nature* 441, 770. <https://doi.org/10.1038/nature04752>
- Casolari, J.M., Brown, C.R., Drubin, D.A., Rando, O.J., Silver, P.A., 2005. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. *Genes Dev.* 19, 1188–1198. <https://doi.org/10.1101/gad.1307205>
- Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., Silver, P.A., 2004. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* 117, 427–439.
- Cerritelli, S.M., Crouch, R.J., 2009. Ribonuclease H: the enzymes in eukaryotes. *FEBS J.* 276, 1494–1505. <https://doi.org/10.1111/j.1742-4658.2009.06908.x>
- Chan, Y.A., Aristizabal, M.J., Lu, P.Y.T., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., Hieter, P., 2014. Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet.* 10, e1004288. <https://doi.org/10.1371/journal.pgen.1004288>
- Chang, E.Y., Wells, J.P., Tsai, S.-H., Coulombe, Y., Chan, Y.A., Zhu, Y.D., Fournier, L.-A., Hieter, P., Masson, J.-Y., Stirling, P.C., 2018. MRE11-RAD50-NBS1 activates Fanconi Anemia R-loop suppression at transcription-replication conflicts (preprint). *Cell Biology*. <https://doi.org/10.1101/472654>

- Chávez, S., García-Rubio, M., Prado, F., Aguilera, A., 2001. Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 21, 7054–7064. <https://doi.org/10.1128/MCB.21.20.7054-7064.2001>
- Chédin, F., Hartono, S.R., Sanz, L.A., Vanoosthuysse, V., 2021. Best practices for the visualization, mapping, and manipulation of R-loops. *EMBO J.* 40, e106394. <https://doi.org/10.15252/embj.2020106394>
- Churikov, D., Charifi, F., Eckert-Boulet, N., Silva, S., Simon, M.-N., Lisby, M., Géli, V., 2016. SUMO-Dependent Relocalization of Eroded Telomeres to Nuclear Pore Complexes Controls Telomere Recombination. *Cell Rep.* 15, 1242–1253. <https://doi.org/10.1016/j.celrep.2016.04.008>
- Cremona, C.A., Sarangi, P., Yang, Y., Hang, L.E., Rahman, S., Zhao, X., 2012. Extensive DNA damage-induced sumoylation contributes to replication and repair and acts in addition to the mec1 checkpoint. *Mol. Cell* 45, 422–432. <https://doi.org/10.1016/j.molcel.2011.11.028>
- Dhingra, N., Wei, L., Zhao, X., 2019. Replication protein A (RPA) sumoylation positively influences the DNA damage checkpoint response in yeast. *J. Biol. Chem.* 294, 2690–2699. <https://doi.org/10.1074/jbc.RA118.006006>
- Dieppois, G., Iglesias, N., Stutz, F., 2006. Cotranscriptional Recruitment to the mRNA Export Receptor Mex67p Contributes to Nuclear Pore Anchoring of Activated Genes. *Mol. Cell. Biol.* 26, 7858–7870. <https://doi.org/10.1128/MCB.00870-06>
- Dieppois, G., Stutz, F., 2010. Connecting the transcription site to the nuclear pore: a multi-tether process that regulates gene expression. *J. Cell Sci.* 123, 1989–1999. <https://doi.org/10.1242/jcs.053694>
- Freudenreich, C.H., Su, X.A., 2016. Relocalization of DNA lesions to the nuclear pore complex. *FEMS Yeast Res.* 16. <https://doi.org/10.1093/femsyr/fow095>
- García-Benítez, F., Gaillard, H., Aguilera, A., 2017. Physical proximity of chromatin to nuclear pores prevents harmful R loop accumulation contributing to maintain genome stability. *Proc. Natl. Acad. Sci. U. S. A.* 114, 10942–10947. <https://doi.org/10.1073/pnas.1707845114>
- García-Muse, T., Aguilera, A., 2019. R Loops: From Physiological to Pathological Roles. *Cell* 179, 604–618. <https://doi.org/10.1016/j.cell.2019.08.055>
- García-Pichardo, D., Cañas, J.C., García-Rubio, M.L., Gómez-González, B., Rondón, A.G., Aguilera, A., 2017. Histone Mutants Separate R Loop Formation from Genome Instability Induction. *Mol. Cell* 66, 597–609.e5. <https://doi.org/10.1016/j.molcel.2017.05.014>
- Geiss-Friedlander, R., Melchior, F., 2007. Concepts in sumoylation: a decade on. *Nat. Rev. Mol. Cell Biol.* 8, 947–956. <https://doi.org/10.1038/nrm2293>
- González-Aguilera, C., Tous, C., Gómez-González, B., Huertas, P., Luna, R., Aguilera, A., 2008. The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability. *Mol. Biol. Cell* 19, 4310–4318. <https://doi.org/10.1091/mbc.e08-04-0355>
- Horigome, C., Bustard, D.E., Marcomini, I., Delgosaie, N., Tsai-Pflugfelder, M., Cobb, J.A., Gasser, S.M., 2016. PolySUMOylation by Siz2 and Mms21 triggers relocation of DNA breaks to nuclear pores through the Slx5/Slx8 STUbL. *Genes Dev.* 30, 931–945. <https://doi.org/10.1101/gad.277665.116>
- Huertas, P., Aguilera, A., 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* 12, 711–721. <https://doi.org/10.1016/j.molcel.2003.08.010>
- Ibarra, A., Hetzer, M.W., 2015. Nuclear pore proteins and the control of genome functions. *Genes Dev.* 29, 337–349. <https://doi.org/10.1101/gad.256495.114>
- Khadaroo, B., Teixeira, M.T., Luciano, P., Eckert-Boulet, N., Germann, S.M., Simon, M.N., Gallina, I., Abdallah, P., Gilson, E., Géli, V., Lisby, M., 2009. The DNA damage response at eroded telomeres and tethering to the nuclear pore complex. *Nat. Cell Biol.* 11, 980–987. <https://doi.org/10.1038/ncb1910>

- Kramarz, K., Schirmeisen, K., Boucherit, V., Ait Saada, A., Lovo, C., Palancade, B., Freudenreich, C., Lambert, S.A.E., 2020. The nuclear pore primes recombination-dependent DNA synthesis at arrested forks by promoting SUMO removal. *Nat. Commun.* 11, 5643. <https://doi.org/10.1038/s41467-020-19516-z>
- Lamm, N., Rogers, S., Cesare, A.J., 2021. Chromatin mobility and relocation in DNA repair. *Trends Cell Biol.* 31, 843–855. <https://doi.org/10.1016/j.tcb.2021.06.002>
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lautier, O., Penzo, A., Rouvière, J.O., Chevreux, G., Collet, L., Loïodice, I., Taddei, A., Devaux, F., Collart, M.A., Palancade, B., 2021. Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis. *Mol. Cell* 81, 2417–2427.e5. <https://doi.org/10.1016/j.molcel.2021.03.030>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., Proudfoot, N.J., 2011. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell* 41, 21–32. <https://doi.org/10.1016/j.molcel.2010.12.007>
- Mouaikel, J., Causse, S.Z., Rougemaille, M., Daubenton-Carafa, Y., Blugeon, C., Lemoine, S., Devaux, F., Darzacq, X., Libri, D., 2013. High-frequency promoter firing links THO complex function to heavy chromatin formation. *Cell Rep.* 5, 1082–1094. <https://doi.org/10.1016/j.celrep.2013.10.013>
- Nagai, S., Dubrana, K., Tsai-Pflugfelder, M., Davidson, M.B., Roberts, T.M., Brown, G.W., Varela, E., Hediger, F., Gasser, S.M., Krogan, N.J., 2008. Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science* 322, 597–602. <https://doi.org/10.1126/science.1162790>
- Nguyen, H.D., Yadav, T., Giri, S., Saez, B., Graubert, T.A., Zou, L., 2017. Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Mol. Cell* 65, 832–847.e4. <https://doi.org/10.1016/j.molcel.2017.01.029>
- Niepel, M., Molloy, K.R., Williams, R., Farr, J.C., Meinema, A.C., Vecchiotti, N., Cristea, I.M., Chait, B.T., Rout, M.P., Strambio-De-Castillia, C., 2013. The nuclear basket proteins Mlp1p and Mlp2p are part of a dynamic interactome including Esc1p and the proteasome. *Mol. Biol. Cell* 24, 3920–3938. <https://doi.org/10.1091/mbc.E13-07-0412>
- Palancade, B., Doye, V., 2008. Sumoylating and desumoylating enzymes at nuclear pores: underpinning their unexpected duties? *Trends Cell Biol.* 18, 174–183. <https://doi.org/10.1016/j.tcb.2008.02.001>
- Pascual-Garcia, P., Capelson, M., 2021. The nuclear pore complex and the genome: organizing and regulatory principles. *Curr. Opin. Genet. Dev.* 67, 142–150. <https://doi.org/10.1016/j.gde.2021.01.005>
- Pelechano, V., Chávez, S., Pérez-Ortín, J.E., 2010. A complete set of nascent transcription rates for yeast genes. *PLoS One* 5, e15442. <https://doi.org/10.1371/journal.pone.0015442>
- Psakhye, I., Jentsch, S., 2012. Protein Group Modification and Synergy in the SUMO Pathway as Exemplified in DNA Repair. *Cell* 151, 807–820. <https://doi.org/10.1016/j.cell.2012.10.021>
- Ptak, C., Saik, N.O., Premashankar, A., Lapetina, D.L., Aitchison, J.D., Montpetit, B., Wozniak, R.W., 2021. Phosphorylation-dependent mitotic SUMOylation drives nuclear envelope–chromatin interactions. *J. Cell Biol.* 220, e202103036. <https://doi.org/10.1083/jcb.202103036>
- Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. Oxf. Engl.* 26, 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., Manke, T., 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–165. <https://doi.org/10.1093/nar/gkw257>

- Reusswig, K.-U., Bittmann, J., Peritore, M., Wierer, M., Mann, M., Pfander, B., 2021. Unscheduled DNA replication in G1 causes genome instability through head-to-tail replication fork collisions. <https://doi.org/10.1101/2021.09.06.459115>
- Rondón, A.G., Jimeno, S., García-Rubio, M., Aguilera, A., 2003. Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation. *J. Biol. Chem.* 278, 39037–39043. <https://doi.org/10.1074/jbc.M305718200>
- Rosonina, E., Duncan, S.M., Manley, J.L., 2010. SUMO functions in constitutive transcription and during activation of inducible genes in yeast. *Genes Dev.* 24, 1242–1252. <https://doi.org/10.1101/gad.1917910>
- Rougemaille, M., Dieppois, G., Kisseleva-Romanova, E., Gudipati, R.K., Lemoine, S., Blugeon, C., Boulay, J., Jensen, T.H., Stutz, F., Devaux, F., Libri, D., 2008. THO/Sub2p Functions to Coordinate 3'-End Processing with Gene-Nuclear Pore Association. *Cell* 135, 308–321. <https://doi.org/10.1016/j.cell.2008.08.005>
- Ryu, T., Bonner, M.R., Chiolo, I., 2016. Cervantes and Quijote protect heterochromatin from aberrant recombination and lead the way to the nuclear periphery. *Nucl. Austin Tex* 7, 485–497. <https://doi.org/10.1080/19491034.2016.1239683>
- Saik, N.O., Park, N., Ptak, C., Adames, N., Aitchison, J.D., Wozniak, R.W., 2020. Recruitment of an Activated Gene to the Yeast Nuclear Pore Complex Requires Sumoylation. *Front. Genet.* 11, 174. <https://doi.org/10.3389/fgene.2020.00174>
- San Martin-Alonso, M., Soler-Oliva, M.E., García-Rubio, M., García-Muse, T., Aguilera, A., 2021. Harmful R-loops are prevented via different cell cycle-specific mechanisms. *Nat. Commun.* 12, 4451. <https://doi.org/10.1038/s41467-021-24737-x>
- Schmid, M., Arib, G., Laemmli, C., Nishikawa, J., Durussel, T., Laemmli, U.K., 2006. Nup-PI: the nucleopore-promoter interaction of genes in yeast. *Mol. Cell* 21, 379–391. <https://doi.org/10.1016/j.molcel.2005.12.012>
- Schneider, M., Hellerschmied, D., Schubert, T., Amlacher, S., Vinayachandran, V., Reja, R., Pugh, B.F., Clausen, T., Köhler, A., 2015. The Nuclear Pore-Associated TREX-2 Complex Employs Mediator to Regulate Gene Expression. *Cell* 162, 1016–1028. <https://doi.org/10.1016/j.cell.2015.07.059>
- Scholes, A.N., Lewis, J.A., 2020. Comparison of RNA isolation methods on RNA-Seq: implications for differential expression and meta-analyses. *BMC Genomics* 21, 249. <https://doi.org/10.1186/s12864-020-6673-2>
- Sikorski, T.W., Ficarro, S.B., Holik, J., Kim, T., Rando, O.J., Marto, J.A., Buratowski, S., 2011. Sub1 and RPA associate with RNA polymerase II at different stages of transcription. *Mol. Cell* 44, 397–409. <https://doi.org/10.1016/j.molcel.2011.09.013>
- Smith, J., Rothstein, R., 1995. A mutation in the gene encoding the *Saccharomyces cerevisiae* single-stranded DNA-binding protein Rfa1 stimulates a RAD52-independent pathway for direct-repeat recombination. *Mol. Cell. Biol.* 15, 1632–1641. <https://doi.org/10.1128/mcb.15.3.1632>
- Sollier, J., Stork, C.T., García-Rubio, M.L., Paulsen, R.D., Aguilera, A., Cimprich, K.A., 2014. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol. Cell* 56, 777–785. <https://doi.org/10.1016/j.molcel.2014.10.020>
- Su, X.A., Dion, V., Gasser, S.M., Freudenreich, C.H., 2015. Regulation of recombination at yeast nuclear pores controls repair and triplet repeat stability. *Genes Dev.* 29, 1006–1017. <https://doi.org/10.1101/gad.256404.114>
- Su, X.A., Freudenreich, C.H., 2017. Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 114, E8392–E8401. <https://doi.org/10.1073/pnas.1711283114>
- Sumner, M.C., Brickner, J., 2021. The Nuclear Pore Complex as a Transcription Regulator. *Cold Spring Harb. Perspect. Biol.* a039438. <https://doi.org/10.1101/cshperspect.a039438>

- Sumner, M.C., Torrisi, S.B., Brickner, D.G., Brickner, J.H., 2021. Random sub-diffusion and capture of genes by the nuclear pore reduces dynamics and coordinates inter-chromosomal movement. *eLife* 10, e66238. <https://doi.org/10.7554/eLife.66238>
- Taddei, A., Van Houwe, G., Hediger, F., Kalck, V., Cubizolles, F., Schober, H., Gasser, S.M., 2006. Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* 441, 774–778. <https://doi.org/10.1038/nature04845>
- Texari, L., Dieppo, G., Vinciguerra, P., Contreras, M.P., Groner, A., Letourneau, A., Stutz, F., 2013. The Nuclear Pore Regulates GAL1 Gene Transcription by Controlling the Localization of the SUMO Protease Ulp1. *Mol. Cell* 51, 807–818. <https://doi.org/10.1016/j.molcel.2013.08.047>
- Teytelman, L., Thurtle, D.M., Rine, J., van Oudenaarden, A., 2013. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proc. Natl. Acad. Sci.* 110, 18602–18607.
- Van de Vosse, D.W., Wan, Y., Lapetina, D.L., Chen, W.-M., Chiang, J.-H., Aitchison, J.D., Wozniak, R.W., 2013. A role for the nucleoporin Nup170p in chromatin structure and gene silencing. *Cell* 152, 969–983. <https://doi.org/10.1016/j.cell.2013.01.049>
- Varejão, N., Ibars, E., Lascorz, J., Colomina, N., Torres-Rosell, J., Reverter, D., 2018. DNA activates the Nse2/Mms21 SUMO E3 ligase in the Smc5/6 complex. *EMBO J.* 37. <https://doi.org/10.15252/emboj.201798306>
- Vertegaal, A.C.O., 2022. Signalling mechanisms and cellular functions of SUMO. *Nat. Rev. Mol. Cell Biol.* <https://doi.org/10.1038/s41580-022-00500-y>
- Wahba, L., Amon, J.D., Koshland, D., Vuica-Ross, M., 2011. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell* 44, 978–988. <https://doi.org/10.1016/j.molcel.2011.10.017>
- Wahba, L., Costantino, L., Tan, F.J., Zimmer, A., Koshland, D., 2016. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* 30, 1327–1338. <https://doi.org/10.1101/gad.280834.116>
- Whalen, J.M., Dhingra, N., Wei, L., Zhao, X., Freudenreich, C.H., 2020. Relocation of Collapsed Forks to the Nuclear Pore Complex Depends on Sumoylation of DNA Repair Proteins and Permits Rad51 Association. *Cell Rep.* 31, 107635. <https://doi.org/10.1016/j.celrep.2020.107635>
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>
- Zhao, X., Blobel, G., 2005. A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proc. Natl. Acad. Sci. U. S. A.* 102, 4777–4782. <https://doi.org/10.1073/pnas.0500537102>

SUPPLEMENTAL MATERIAL

A R-loop sensing pathway mediates the relocation of transcribed genes to nuclear pore complexes

Arianna Penzo, Marion Dubarry, Clémentine Brocas, Myriam Zheng, Ophélie Lautier, Marie-Noëlle Simon, Vincent Geli, Karine Dubrana & Benoit Palancade

includes:

Figures S1-4

Tables S1-4

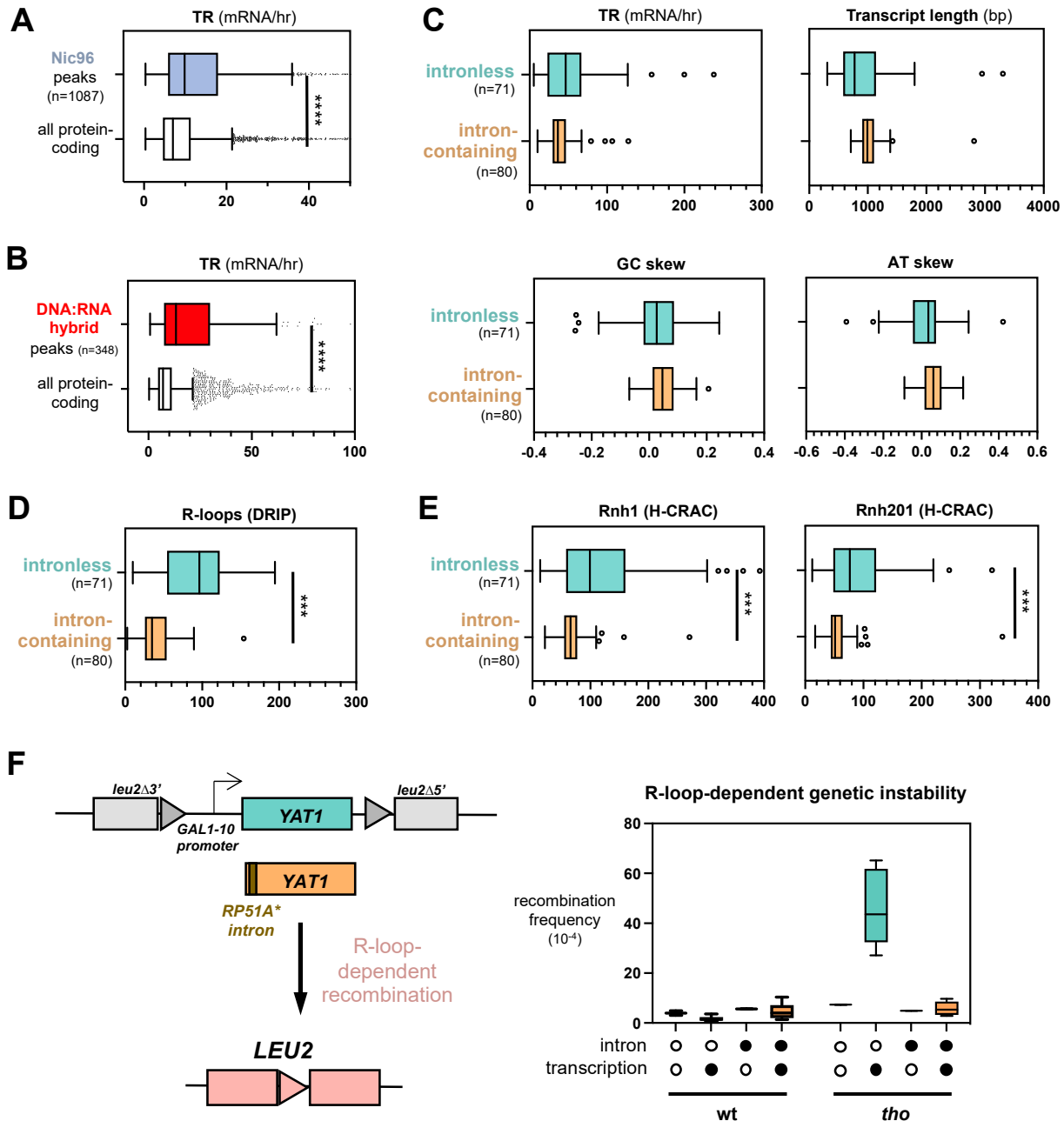


Figure S1 (related to Figure 1). **Validation of the gene datasets and reporter systems used to analyze the relationships between NPC association and R-loop levels.** **A-B**, Transcription rates (TR, mRNA/hr, from Pelechano et al., 2010) for the genes associated with Nic96 peaks (this study), DNA:RNA hybrid peaks (Wahba et al., 2016) or all protein-coding genes. The number of considered peaks is indicated. **C**, Transcription rates (TR, mRNA/hr, from Pelechano et al., 2010), transcript length (pre-mRNA, bp), GC-skew and AT-skew for the intronless and intron-containing highly-expressed genes defined in Bonnet et al. 2017 (listed in **Table S1**). **D-E**, R-loops levels for the same groups of intronless and intron-containing highly-expressed genes as scored by DNA:RNA hybrid immunoprecipitation (**D**, DRIP, Wahba et al., 2016), RNase H1 CRAC (**E**, left panel, Aiello et al., 2022) and RNase H2 CRAC (**E**, Rnh201, right panel, Aiello et al., 2022). **F**, Principle of the hyper-recombination assay. Either the intronless or the intron-containing version of the *YAT1* transgene are inserted within chromosome II under the control of the *GAL1-10* promoter and flanked by direct *leu2* repeats. Recombination leads to the reconstitution of a functional *LEU2* marker. Recombination frequencies (fraction of Leu⁺ prototrophs, x 10⁴; n=2-6) were calculated as described in Material and Methods for *YAT1* or *intron-YAT1* strains, either wt or *tho* (*mft1*Δ). When indicated, transcription was induced for 5h in the presence of galactose.

*** p<0.001; **** p<0.0001; Mann-Whitney-Wilcoxon test.

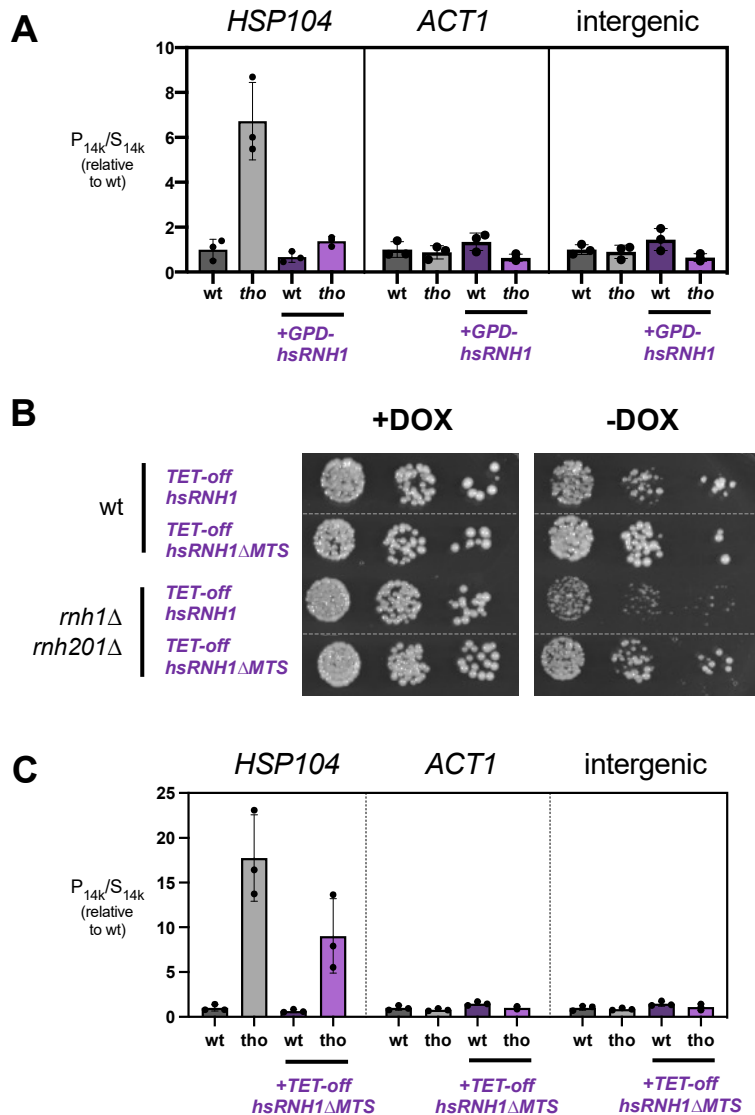


Figure S2 (related to Figure 2). **RNase H1 over-expression suppresses *HSP104* co-fractionation with NPCs.** **A**, qPCR-based quantification of the amount of DNA from the indicated loci in heavy chromatin fractions from wt or *tho* (*mft1* Δ) mutant cells transformed with either an empty vector or a GPD-hsRNH1 construct (+GPD-hsRNH1), and heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean \pm SD, n=3, relative to wt). **B**, Serial dilutions of wt or *rnh1* Δ *rnh201* Δ mutant cells carrying either the TET-off-hsRNH1 or the TET-off-hsRNH1- Δ MTS constructs were grown at 30°C on selective medium (SC-Ura). When indicated, the medium was supplemented with doxycycline (DOX, 5 μ g/mL) to repress the *TET-off* promoter. Note that while induction of the hsRNH1 construct triggers growth defects in both strains (-DOX panel), over-expression of the version deleted for the mitochondrial targeting sequence (MTS) does not detectably impact cell fitness. **C**, qPCR-based quantification of the amount of DNA from the indicated loci in heavy chromatin fractions from wt or *tho* (*mft1* Δ) mutant cells transformed with the TET-off-hsRNH1- Δ MTS construct and heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean \pm SD, n=3, relative to wt). RNH1 induction was achieved by growing the cells in the absence of doxycycline for 16h prior to HS (+*TET-off* hsRNH1 Δ MTS).

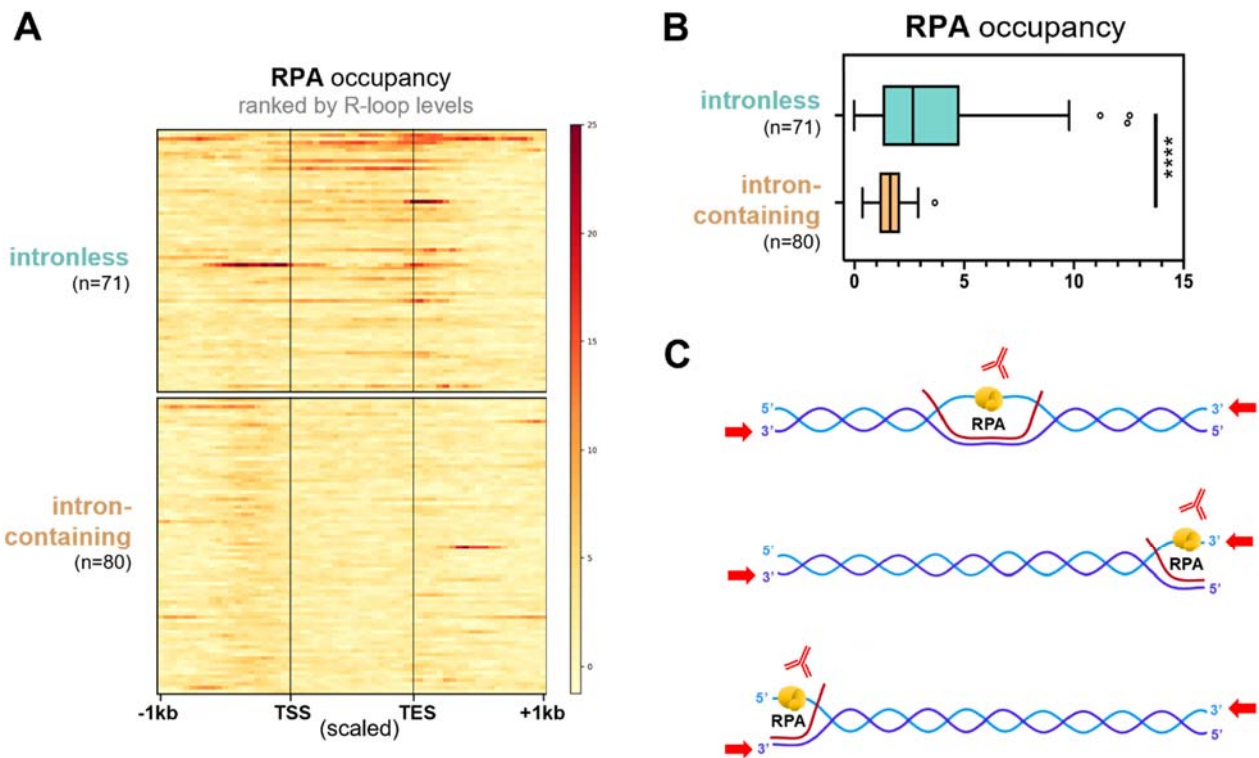


Figure S3 (related to Figure 3). **RPA ChIP-seq analysis.** **A**, Heatmap analysis of RPA occupancy at the antisense strand of highly-transcribed intronless and intron-containing genes, aligned at their Transcription Start Site (TSS) and Transcription End Site (TES), in wt cells arrested in G0 (strand-specific RPA ChIP-seq dataset from Reusswig et al., 2021). Only the regions between the TSS and the TES are scaled. Genes are grouped based on their intron content and ranked according to their R-loop levels (as measured in Wahba et al., 2016). **B**, Average RPA occupancy at the antisense strand of highly-transcribed intronless and intron-containing genes (input subtracted). **C**, Immunoprecipitation of R-loop-bound RPA is not expected to retrieve directional signals in strand-specific ChIP-seq. In the procedure used by Reusswig et al. (2021), immunoprecipitated DNA is denatured and 3'-specific adaptors are ligated (red arrows) prior to library amplification. Since R-loops (~150bps; García-Pichardo et al., 2017) are expected to be smaller than sonicated DNA fragments (200-500bp), this treatment will similarly tail both strands with the adaptors.

**** $p < 0.0001$; Mann-Whitney-Wilcoxon test.

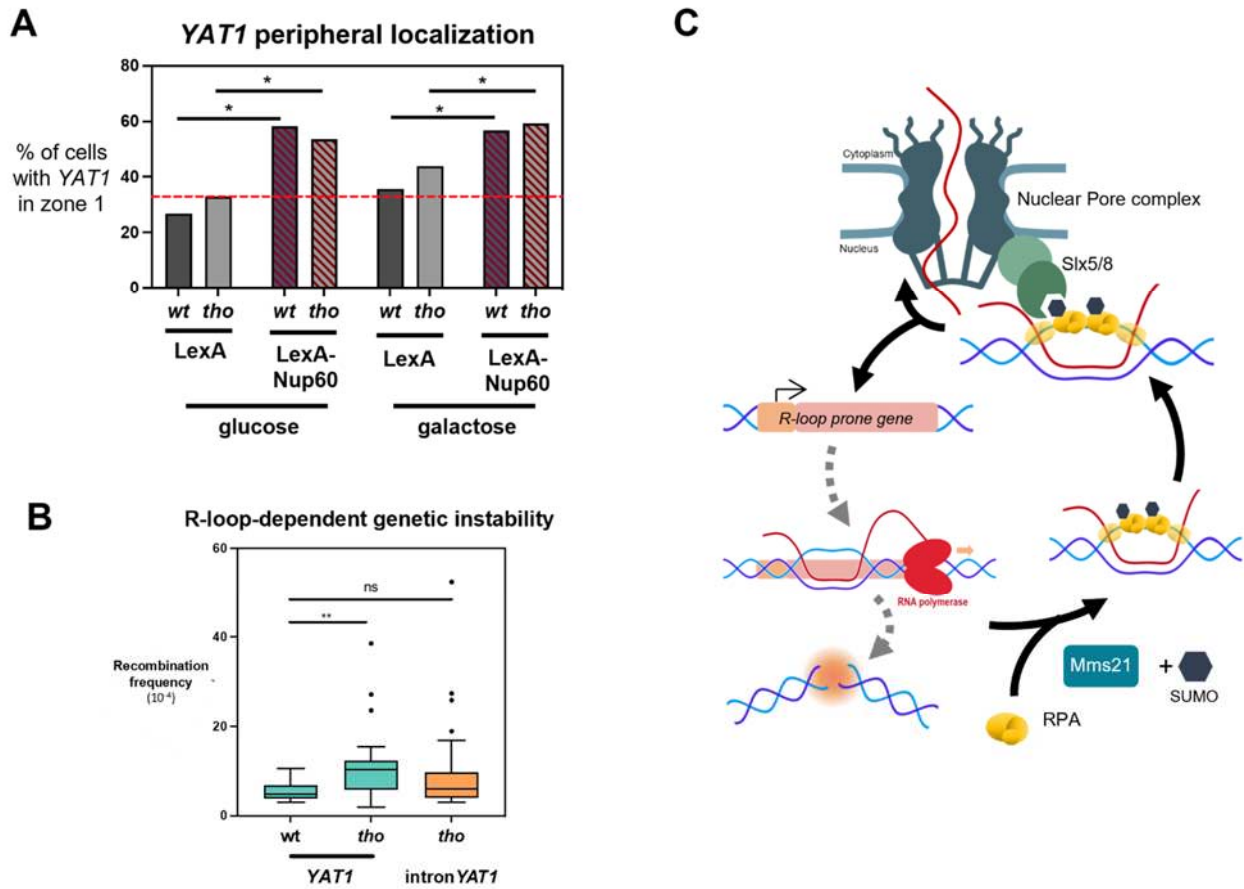


Figure S4 (related to Figure 5 and Discussion). **NPC association alleviates R-loop toxicity.** **A**, Fraction of cells (%) showing intronless *YAT1* in zone 1, in wt or *tho* (*mft1Δ*) mutant cells carrying either LexA- or LexA-Nup60-expressing constructs, grown in glycerol-lactate medium and further treated with glucose (glu) or galactose (gal) for 5h. Between 260 and 500 cells were counted in 2 independent experiments for each condition. The red dashed bar indicates the expected value for a randomly-distributed locus. **B**, Recombination frequencies (fraction of Leu⁺ prototrophs, x 10⁴; wt *YAT1* n=16, *tho* *YAT1* n=20, *tho* intron*YAT1* n=24) were calculated as described in Material and Methods for *YAT1* or intron-*YAT1* strains, either wt or *tho* (*mft1Δ*), grown in glucose-containing medium. **C**, Proposed model for the relocalisation of R-loop forming loci to nuclear pores.

*p<0.05; **p<0.01; ns, not significant; Mann-Whitney-Wilcoxon test.

Table S1. Intronless and intron-containing highly-expressed genes considered in this study
(related to Fig. 1 and 3).

intronless (n=71)		intron-containing (n=80)	
YLR029C	YHR193C	YGR034W	YNL302C
YLR110C	YNL031C	YLR061W	YOR182C
YJL158C	YDL055C	YAL003W	YIL133C
YLR044C	YJR009C	YBL092W	YPR132W
YKL060C	YGR234W	YGL103W	YPL143W
YKL152C	YEL009C	YNL112W	YML026C
YGL008C	YLR441C	YGL030W	YML024W
YEL027W	YBL003C	YDL130W	YDL083C
YNL145W	YGL147C	YLR048W	YLR388W
YLR249W	YJL190C	YJL189W	YIL148W
YCR012W	YBR010W	YPR043W	YJL177W
YLR075W	YDR134C	YBR048W	YNL162W
YDR382W	YDR225W	YOL127W	YOL120C
YKL056C	YDR418W	YHR010W	YLR406C
YOR063W	YGL253W	YDL061C	YIL052C
YPL131W	YKL216W	YIL069C	YOR096W
YAL038W	YLR300W	YMR116C	YDR447C
YJR123W	YLL045C	YHR021C	YLR185W
YHL015W	YNL030W	YGL031C	YLR333C
YOL086C	YDR155C	YDL075W	YKR094C
YOL040C	YNL067W	YML073C	YDR064W
YDR050C	YLR264W	YNL069C	YDL082W
YDR224C	YDL192W	YIL018W	YDR471W
YGL123W	YEL054C	YJR145C	YDR025W
YDR461W	YGR037C	YDR500C	YLR448W
YOR369C	YPL037C	YKR057W	YBR189W
YLR167W	YHR089C	YBL087C	YOR312C
YDL081C	YDL014W	YGL189C	YKL156W
YLR340W	YEL034W	YBR191W	YNL301C
YOR167C	YBR106W	YPL079W	YCR031C
YNL178W		YGR118W	YMR143W
YOL039W		YGR148C	YBL027W
YOL109W		YJL136C	YHR141C
YHR174W		YOR234C	YER074W
YGR060W		YHR203C	YGR214W
YLR325C		YER117W	YNL096C
YGR192C		YKL180W	YMR142C
YDR276C		YOR293W	YLR344W
YDR033W		YDR450W	YMR230W
YIL053W		YOL121C	
YML028W		YHL001W	

Table S2. Strains used in this study (related to Material and Methods).

CODE	NAME	GENOTYPE	SOURCE	USAGE (Figures)
YBP539	wt (BY4742)	MATalpha <i>ura3 his3 leu2 lys2</i>	Euroscarf	Fig. 1B-C, 2D-F, S2A-C, 3A, 4E, 4I, 5A
YBP936	wt (BY4741)	MATa <i>ura3 his3 leu2 met15</i>	Euroscarf	Fig. 3B
YBP1525	wt (W303)	MATalpha <i>ade2 ura3 his3 trp1 leu2 can1</i>	Gift from C. Dargemont	Fig. 3G, 4A, 4D, 4F, 4H, 5A
YKD2204	wt (<i>HSP104</i>)	(W303) <i>LacO@HSP104 3'::TRP1 ade2-1::GFP-LacI-ADE2 pRS316-NUP49-mCherry</i>	This Study	Fig. 2B, 3F, 4C, 4G
YPB2102	wt (<i>YAT1</i>)	(W303) <i>NUP49::NUP49-mCherry-HphMX his3::LacI-GFP::HIS3 LacO@GAL10 3'::TRP1 LexA-BS@GAL1 3' pLEU2-leu2Δ3'-pGAL1-YAT1-leu2Δ5'::KanMX@GAL1 3'</i>	This Study	Fig. 1H, S1F, 5C, S5A-B
YPB2103	wt (intron- <i>YAT1</i>)	(W303) <i>NUP49::NUP49-mCherry-HPHmx his3::LacI-GFP::HIS3 LacO@GAL10 3'::TRP1 LexA-BS@GAL1 3' pLEU2-leu2Δ3'-pGAL1-RPL51A*intron-YAT1-leu2Δ5'::KanMX@GAL1 3'</i>	This Study	Fig. 1H, S1F, S5A
YBP1501	<i>tho</i> (BY4742)	(BY4742) <i>mft1::KanMX</i>	Euroscarf	Fig. 2D-F, S2A, S2C, 3A, 4E, 4I, 5A
YBP2109	<i>tho</i> (BY4741)	(BY4741) <i>mft1::KanMX</i>	Euroscarf	Fig. 3B
YPB2006	<i>tho</i> (W303)	(W303) <i>mft1::KanMX</i>	Gift from D. Libri	Fig. 3G, 4D, 4F, 4H, 5A
YKD2373	<i>tho</i> (<i>HSP104</i>)	(YKD2204) <i>mft1::KanMX</i>	This Study	Fig. 2B, 3F, 4C, 4G
YBP2105	<i>tho</i> (<i>YAT1</i>)	(YPB2102) <i>mft1::NatMX</i>	This Study	Fig. 1H, S1F, 5C, S5A-B
YBP2106	<i>tho</i> (intron- <i>YAT1</i>)	(YPB2103) <i>mft1::NatMX</i>	This Study	Fig. 1H, S1F, S5A-B
YBP2307	Nic-96-myc	(BY4742) <i>NIC96-13Myc::KanMX</i>	This Study	Fig. 1B-C, 2E-F, 3E
YBP2308	<i>tho</i> Nic-96-myc	(BY4742) <i>NIC96-13Myc::KanMX mft1::NatMX</i>	This Study	Fig. 2E-F

YBP1978	<i>rnh1</i> Δ <i>rnh201</i> Δ	<i>rnh1::KanMX rnh201::NatMX</i>	Derived from LW5031, gift from D. Koshland	Fig. S2B
YBP2332	<i>tho fc1</i> Δ	(BY4742) <i>fcy1::KanMX mft1::NatMX</i>	This Study	Fig. 3A
YBP2331	<i>tho mlh3</i> Δ	(BY4742) <i>mlh3::KanMX mft1::NatMX</i>	This Study	Fig. 3A
YBP2018	<i>tho rad2</i> Δ	(BY4742) <i>rad2::KanMX mft1::KanMX</i>	This Study	Fig. 3A
YBP2315	<i>rfa1-D228Y</i> (BY4742)	(BY4742) <i>rfa1-D228Y::NatMX</i>	This Study	Fig. 3E
YBP1478	<i>siz1</i> Δ <i>siz2</i> Δ	(BY4742) <i>siz1::KanMX siz2::KanMX</i>	This study	Fig. 4E, 5A
YBP2238	<i>tho siz1</i> Δ <i>siz2</i> Δ	(BY4742) <i>siz1::KanMX siz2::KanMX mft1::NatMX</i>	This study	Fig. 4E, 5A
YBP2277	<i>slx5</i>	(BY4742) <i>slx5::NatMX</i>	This study	Fig. 4I
YBP2237	<i>tho slx5</i>	(BY4742) <i>slx5::NatMX mft1::KanMX</i>	This study	Fig. 4I
YBP1167	<i>slx8</i>	(BY4742) <i>slx8::HphMX</i>	This study	Fig. 4I, 5A
YBP2166	<i>tho slx8</i>	(BY4742) <i>slx8::HphMX mft1::NatMX</i>	This study	Fig. 4I, 5A
YBP2202	<i>rfa1-D228Y</i>	(W303) <i>rfa1-D228Y</i>	Audry et al, 2015	Fig. 3G, 5A
YBP2219	<i>tho rfa1-D228Y</i>	(W303) <i>rfa1-D228Y mft1::KanMX</i>	This study	Fig. 3G, 5A
YBP1769	<i>ulp1-333</i>	(W303) <i>ULP1::HIS3 YCpLac22-ulp1-333-TRP1</i>	Gift from F. Stutz	Fig. 4A
YBP2333	<i>ulp1-333 rfa1-4KR</i>	(W303) <i>rfa1- K170R, K180R, K411R, K427R; ULP1::KanMX; YCpLac11-LEU2-ulp1-333</i>	This Study	Fig. 4A
YBP2243	<i>rfa1-KR</i>	(W303) <i>rfa1- K170R, K180R, K411R, K427R mft1::KanMX</i>	Dhingra et al, 2019	Fig. 4H
YBP2280	<i>tho rfa1-KR</i>	(W303) <i>rfa1- K170R, K180R, K411R, K427R mft1::KanMX</i>	This Study	Fig. 4H
YBP1079	<i>mms21-11</i>	(W303) <i>mms21-11-LEU2</i>	Gift from X. Zhao	Fig. 4D, 5A
YBP2278	<i>tho mms21-11</i>	(W303) <i>mms21-11-LEU2 mft1::KanMX</i>	This Study	Fig. 4D, 5A
YBP2282	<i>smt3-KR</i>	(W303) <i>smt3- K11R, K15R, K19R::TRP1</i>	Gift from S. Marcand	Fig. 4F
YBP2290	<i>tho smt3-KR</i>	(W303) <i>smt3- K11R, K15R, K19R::TRP1 mft1::KanMX</i>	This study	Fig. 4F
YKD2206	<i>rfa1-D228Y</i> (HSP104)	(YKD2204) <i>rfa1-D228Y</i>	This Study	Fig. 3F

YKD2238	<i>tho rfa1-D228Y</i> (HSP104)	(YKD2204) <i>rfa1-D228Y mft1::KanMX</i>	This Study	Fig. 3F
YKD2427	<i>rfa1-KR</i> (HSP104)	(YKD2204) <i>rfa1- K170R, K180R, K411R, K427R</i>	This Study	Fig. 4C
YBP2380	<i>tho rfa1-KR</i> (HSP104)	(YKD2204) <i>rfa1- K170R, K180R, K411R, K427R mft1::KanMX</i>	This Study	Fig. 4C
YKD2378	<i>mms21-11</i> (HSP104)	(YKD2204) <i>mms21-11::LEU2</i>	This Study	Fig. 4G
YKD2275	<i>tho mms21-11</i> (HSP104)	(YKD2204) <i>mms21-11::LEU2 mft1::KanMX</i>	This Study	Fig. 4G

Table S3. Plasmids used in this study (related to Material and Methods).

CODE	NAME	SEQUENCE	REFERENCE	Usage
BP1932	GPD-hsRNH1	AmpR/HIS3/2 μ : GPD _{prom} -myc-hsRNH1-CYC1 _{term}	Bonnet et al., 2017	Fig. 2B-D, S2A
BP2167	TET-off <i>hsRNH1ΔMTS</i>	AmpR/URA3/2 μ : CMV _{prom} -TetR-VP16, tetO ₇ - Δ MTS-hsRNH1-3xFlag	This study	Fig. S2B-C
BP2155	His ₆ -SMT3	AmpR/URA3/CEN: SMT3 _{prom} -His ₆ -SMT3-SMT3 _{term}	This study	Fig. 4
BP1882	LexA	AmpR/URA3/2 μ : LexA	Texari et al., 2013	Fig. 5C, S4A, S4C
BP1883	LexA-Nup60	AmpR/URA3/2 μ : LexA-Nup60	Texari et al., 2013	Fig. 5C, S4A, S4C

Table S4. Primers used in this study (related to Material and Methods).

HSP104 3' F	GTTCTACCAAATCACGAAGC
HSP104 3' R	TCTAGGTCATCATCAATTTCC
ACT1 3' F	ACGTTACCCAATTGAACACG
ACT1 3' R	AGAACAGGGTGTTCTTCTGG
YEF3 3' F	GATTGCCGGTGGTAAGAAGA
YEF3 3' R	CGTAAGCATCACCCAATTCC
Intergenic* F	GAAACCACGAAAAGTTCACCA
Intergenic* R	AGCTTCTGCAAACCTCATTG

* chrIV:43199..53262

DISCUSSION

1 | Mechanisms underlying compartmentalisation of mRNA-related processes at NPCs

In this manuscript, I describe two additional processes that enable the targeting of RNA-containing structures, i.e. mRNAs engaged in translation (Article 1) and R-loops (Article 2) to the nuclear pore, which result in the prevention of protein aggregates formation on the cytoplasmic side and the protection against transcription-dependent genetic instability on the nucleoplasmic side.

1.1 | Karyopherin-dependent localised translation at nuclear pores

In our study aiming at deciphering the co-translational regulation of NPC assembly, we identified a specific subset of mRNAs whose translation occurs at the proximity of the nuclear pore (Article 1). The association of these mRNAs to the NPC requires the assembly of polysomes (Article 1, Fig. 4E, 4F), and is mediated by the recognition of the N-terminal portion of the nascent polypeptide by karyopherins (Kaps) (Article 1, Fig. 3B, 3C). The transport receptors mediate the association of the mRNA with the NPC and may favour the rapid co-translational import of the nuclear protein inside the nucleus. Additional analyses are ongoing to further characterise the pathway leading to relocation, in particular regarding the requirement for specific *cis*- or *trans*- acting factors for targeting this specific subset of mRNAs to the pore. Given the variety of canonical and non-canonical NLS characterised (Lu et al., 2021), we could hypothesize that Kaps-mediated recognition of nascent polypeptides may concern a specific NLS type, which may be recognised by dedicated transport receptors. Moreover, the observed bias for long genes (Article 2, Fig. 3E) prompts to pursue the investigation of the role of translational kinetics in this process, in particular concerning the presence of ribosome pausing sites in the mRNAs of interest.

1.2 | R-loops as the signal for active genes relocation to the nuclear pore

In our study investigating the causes driving the relocation of active genes to the nuclear pore (Article 2), we demonstrated how the repositioning of inducible loci to the nuclear periphery can be dependent on R-loop formation. To prove this, we took advantage of the *tho* mutant that shows an increased relocalisation of heat-shock genes and galactose-inducible reporters, in which R-loop accumulation could be modulated in *cis*, by insertion of an intron (Article 2, Fig. 1H) or in *trans*, by overexpressing hsRNase H1 (Article 2, Fig. 2B, D).

1.2.1 | R-loop gating in R-loop accumulating mutants

If R-loops are the intrinsic cause of loci relocation to the nuclear periphery, can we expect to see the same enhanced localization phenotype in other R-loop accumulating mutants?

***HSP104* relocalisation in absence of RNase H**

Interestingly, when assessing *HSP104* relocation to the nuclear pore by chromatin fractionation in the *rnh1Δ rnh201Δ* R-loop accumulating mutant, the locus does not show any difference in NPC association as compared to the wt (Fig. 1). This suggests that the relocalisation phenotype could be specific of the *tho* mutant, and may cast doubt on the actual contribution of R-loops in triggering relocalisation. Indeed, the study first showing differential fractionation of *HSP104* in *tho* mutants (Rougemaille et al., 2008a) attributed this phenomenon to defects in the processing of the 3' end of the RNA (further discussed in section 1.3.1). However, the failure to detect such an interaction in the *rnh1Δ rnh201Δ* mutant could be explained by the different composition of the mRNP in *THO+* cells, which could form more dynamic interactions with the pore, undetectable with our biochemical approach. This would be consistent with the idea that mRNPs-NPC interactions would have a role in stabilising, rather than triggering relocation, as observed for transcription-dependent gene gating, which is preferentially promoter-dependent (see Introduction, paragraph 1.2.4). Live imaging observations of *HSP104* localisation in the *rnh1Δ rnh201Δ* mutant would be therefore necessary to clarify this point.

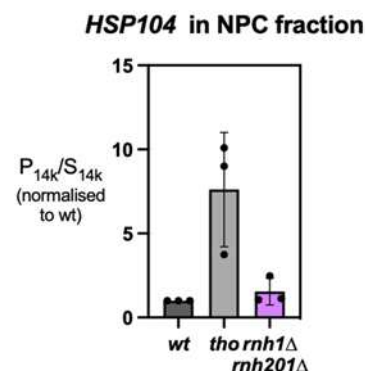


Figure 1. *HSP104* co-fractionation with the nuclear pore in absence of RNase H1 and 2. qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from wt, *tho* (*mft1Δ*) or *rnh1Δ rnh201Δ* cells heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean \pm SD, n=3, relative to wt; Methods as in Article2, except that the strains were maintained at 30°C prior to heat-shock).

Moreover, since there are no available informations about the extent of accumulation and the position of R-loops in heat-shock condition, the possibility that R-loops are different in the two mutants cannot be excluded. While RNases H act by degrading already formed R-loops, the THO complex has a pivotal role in preventing their formation. In *rnh1Δ rnh201Δ* cells, the active THO complex could therefore be efficient in avoiding R-loop accumulation at stress-induced loci, thus preventing the formation of the substrate responsible for *HSP104* relocation. It would be reasonable to suppose that the action of the THO complex is dominant in containing R-loop accumulation compared to RNases H, both in terms of the timing of intervention and the use of energy and resources. Relying primarily on a pathway that causes RNA degradation could in fact prove extremely costly, especially in emergency conditions such as heat shock stress, which requires a rapid response.

Role of Sen1 in R-loop gating

The helicase Sen1 has been shown to predominantly act on R-loops in S-phase, by promoting resolution of transcription-replication conflicts (San Martin-Alonso et al., 2021; Aiello et al., 2022). Sen1 inactivation could lead to an increase in loci relocation to the nuclear pore, by stabilising R-loops involved in transcription-replication conflicts, which would remain unresolved. However, we could expect this increase to be limited to the proportion of genes actively transcribed in concomitance with fork passage. Moreover, relocation of such loci could be also subjected to stalled fork-induced replication, rather than R-loop gating. Finally, the same considerations exposed above about the predominance of THO-dependent R-loop prevention mechanisms over R-loop resolution are valid also in this case.

Systematically assessing NPC relocation in several distinct mutant situations associated with R-loop accumulation will be required to shed additional light on the dynamics of R-loop prevention and/or removal at induced loci and its influence on gene position in the nucleus.

1.2.2 | R-loop detection upon stress-induced transcription

To investigate R-loop contribution to loci association to the nuclear pore, we used two model loci: the ectopically-inserted, galactose-inducible *YAT1* reporter, and the endogenous heat-shock responsive *HSP104* gene. While the *YAT1* construct has already been characterised as a *bona fide* R-loop forming sequence (Bonnet et al., 2017), no evidence is currently available in the literature of R-loop profiles obtained upon heat-shock activation. The heat-shock response causes substantial changes in the transcription program, both by upregulating the Hsf1 regulon, which constitutes 3% of the yeast genome (Hahn et al., 2004), and downregulating the genes for ribosome biogenesis (Castells-Roca et al., 2011). The generation of genome-wide R-loop maps upon heat-shock induction is therefore pivotal to provide reliable R-loop profiles in this condition, particularly with respect to the determination of R-loops position within coding regions, which in maps generated in standard growth conditions shows a bias for the 3' end of ORFs (Wahba et al., 2016; Aiello et al., 2022).

Our attempts at R-loop detection at the *HSP104* gene by DRIP-qPCR have up to now failed to reproducibly detect DNA:RNA hybrids at *HSP104* 3' (data not shown), the position at which the chromatin fractionation signal peaks (Rougemaille et al., 2008a; Mouaikel et al., 2013). R-loop detection has also been particularly challenging in the *tho* mutant, which displays a strong transcription impairment (Huertas and Aguilera, 2003, Fig. 2). For a more accurate quantification, normalisation of the R-loop signal to the nascent transcript levels or polII occupancy, as in (Bonnet et al., 2017), may be necessary to be able to account for these difference in transcription efficiency.

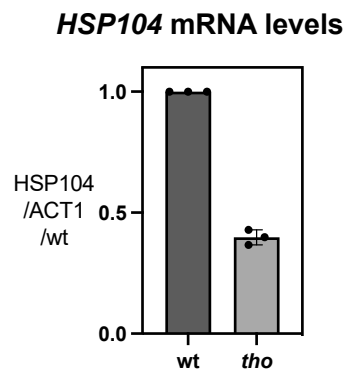


Figure 2. ***HSP104* transcript levels in wt and *tho* mutant.** qPCR-based quantification of *HSP104* mRNA levels from wt and *tho* (*mft1Δ*) cells heat-shocked at 37°C for 15min (Method as in Article1, heatshock condition, strains and primers as in Article2). RNA levels are expressed relative to *ACT1* and to the wt (mean±SD, n=3).

Furthermore, in light of the two different classes of R-loops known to be differentially detected by antibody-based or RNase H-based methods (Castillo-Guzman and Chédin, 2021; Miller et al., 2022, discussed in introduction, paragraph 2.2.3), employing the latter, for example through dead-scrNH1-ChIP, could be more suitable for the detection of relocation-triggering R-loops. In view of the rapidity of the relocation, these structures may have a very short half-life, especially if we speculate that the proximity to the pore could prevent formation of other hybrids after the pioneering one driving repositioning (further discussed in paragraph 2.2.2). Indeed, *HSP104* chromatin fractionation is detectable at 2' upon heat-shock induction, when the genes already underwent around 40 transcription cycles (Mouaikel et al., 2013). In addition, it will be critical to assess R-loop formation at *HSP104* in mutants preventing relocation (e.g. *rfa1-D228Y*; *mms21-11*).

1.3 | R-loop sensing mechanisms: role of RPA and SUMOylation

The mechanism of R-loop relocation to the nuclear pore identified in our study relies on the coating of the displaced ssDNA by RPA, and on mono-SUMOylation events specifically dependent on the Mms21 SUMO-ligase, a subunit of the cohesin-like Smc5/6 complex.

1.3.1 | RNA metabolism and R-loop sensing by RPA

THO-dependent *HSP104* co-fractionation with the nuclear pore was previously shown to be suppressed in mutants of the nuclear exosome (*rrp6*) and of the cleavage and polyadenylation factor 1 complex (*rna14-3* and *rna15-2*) (Rougemaille et al., 2008a). This shows the requirement of these factors for preventing the prolonged formation of the protein-DNA complex which mediates the association of the gene to the NPC, pointing to a dependency of *HSP104* localisation to commitment to 3' end processing and termination. However, a different interpretation could be proposed: in light of the accumulation of unprocessed RNA species characteristic of the *rrp6* mutant (Wyers et al., 2005), we could hypothesise that in this context, RNA-binding factors could be titrated by these excess of RNA molecules, thus interfering with the establishment or the stabilisation of R-loop-driven gene-NPC interactions. Interestingly, it had been previously reported that *rrp6* mutants can display aberrant termination phenotypes due to the titration of the RNA-binding Nrd1-Nab3-Sen1 termination factor (Villa et al., 2020).

Assessing the effect of excess RNA on R-loop gating

To disentangle the effect of RNA accumulation versus the physical requirement for 3' end processing factors, we sought to increase nuclear RNA levels in a manner that would not directly affect *HSP104* transcription efficiency or mRNA metabolism. To achieve this, we took advantage of a mutant for the debranching enzyme Dbr1, involved in the linearisation of the intron lariat after pre-mRNA splicing, in which excessive unprocessed lariats accumulate in the nucleus, without apparent consequences for the efficiency of the splicing process or cell fitness (Chapman and Boeke, 1991). *DBR1* inactivation is unlikely to directly affect *HSP104*, which does not contain an intron. Strikingly, combination of the *dbr1* and *tho* mutations completely suppresses *HSP104*-NPC association (Fig. 3A), showing that the presence of excess RNA in the cell could indeed perturb R-loop gating, likely by titrating RNA-binding factors which are required for the NPC interaction.

Tools to investigate the titration of gating factors by RNA

Further experiments are required to understand how the excess RNA can hinder *HSP104*-cofractionation with the nuclear pore. In light of the requirement of RPA for *HSP104*-NPC association (Article2, Fig. 3F-G), and its recently re-evaluated affinity for RNA (Mazina et al., 2020), titration of RPA by accumulated or extended transcripts could explain the decrease in relocation observed in *tho dbr1*, *tho rrp6*, *tho rna14-3* and *tho rna15-2* cells. Interestingly, yeast cells lacking Rrp6 or Trf4, a subunit of the TRAMP (Trf4/5-Air1/2-Mtr4) Polyadenylation complex, display phenotypes evoking RPA loss-of-function (Manfrini et al., 2015; Brown et al., 2021). To attempt to rescue *HSP104* co-fractionation with the NPC in *tho dbr1* cells, a multicopy plasmid was constructed, carrying the coding sequences of the three subunits of the RPA trimer, which allow to achieve a ~40 fold increase in their RNA

levels (data not shown, consistent with Brown et al., 2021), and up to a ≈ 4 -fold increase in Rfa1 protein levels (Fig. 3B-C). The use of a similar construct already successfully increased resilience of human cells against replication catastrophe (Toledo et al., 2013) and suppressed CAG tract fragility in yeast cells lacking Trf4 (Brown et al., 2021).

Preliminary chromatin fractionation assays, however, did not show any impact of RPA overexpression on *HSP104* cofractionation in *tho dbr1* cells (data not shown). Further controls would be necessary to ensure that the increase in RPA subunits levels effectively results in the formation of functional trimers, and that the achieved overexpression is sufficient to counteract the extent of RNA accumulation in this mutant. At this stage, this result suggests that excessive RNA lariats could rather impact other factors involved in R-loop gating. Live imaging confirmation of *HSP104* localisation in the *tho dbr1* double mutant will help to elucidate whether excessive RNA accumulation impacts R-loop sensing, therefore abolishing *HSP104* peripheral localisation, supporting the RPA titration hypothesis, or only the subsequent stabilisation of the interaction, by affecting mRNP composition.

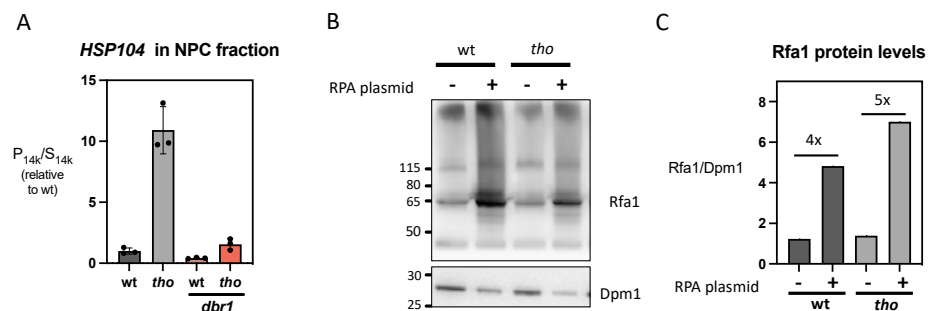


Figure 3. **Tools to investigate the titration of gating factors by RNAs.** A. qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from the indicated strains heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean \pm SD, n=3, relative to wt; Methods as in Article2). B. Whole cell extracts from wt and *tho* cells carrying either the RPA overexpressing construct or an empty plasmid were analyzed by SDS-PAGE and western blotting using anti-RPA and anti-Dpm1 antibodies (methods as in Article1). Dpm1 is used as a loading control. The position of molecular weights is indicated (kDa). C. Quantification of the protein levels from (B). The mean integrated densities of the bands corresponding to the proteins of interest (Rfa1 and Dpm1) were quantified using ImageJ. After subtracting the background, Rfa1 values were normalised to the corresponding Dpm1 amount. Fold increase between *tho* and wt levels are indicated.

1.3.2 | Is the Smc5/6 complex being specifically recruited at R-loops?

The complete suppression of *HSP104* peripheral localisation in the *rfa1-D228Y* and *mms21-11* mutants (Article2, Fig. 3F, 3G, 4C, 4D) demonstrated how targeting of R-loops to nuclear pores results from a combined action of RPA and the Smc5/6 complex, of which the SUMO-ligase Mms21 is a subunit. R-loop-dependent recruitment of the Smc5/6 complex to DNA has been observed in the context of host immune response against Epstein-Barr Virus infection (Yiu et al., 2022). Moreover, Mms21 sumoylation activity has been shown to be enhanced by ssDNA binding in vitro (Varejão et al., 2018), and conditional inactivation of the Nse4 subunit in budding yeast causes increased levels of R-loops (Chang et al., 2018). In addition, chromatin fractionation assays in a mutant of another Smc5/6 subunit,

Nse3, which was shown to decrease recruitment of the complex to the DNA (Moradi-Fard et al., 2016), also shows a decrease in *HSP104*-NPC association (Fig. 4), further supporting the importance of Smc5/6 recruitment at the site of R-loop formation to trigger the relocation.

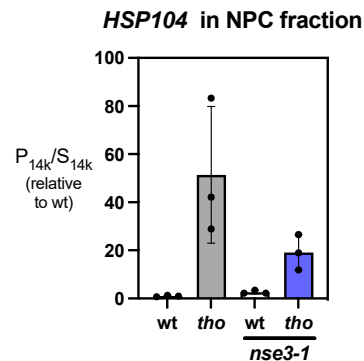


Figure 4. *HSP104* co-fractionation with the nuclear pore in mutants affecting Smc5/6 complex recruitment to chromatin. qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from the indicated strains heat-shocked at 37°C for 15min (P14k/S14k, mean±SD, n=3, relative to wt; Methods as in Article2).

The Smc5/6 complex has been shown to be required for NPC relocation of TNRS-spanning collapsed forks, in a manner dependent on Smc5 SUMOylation (Whalen et al., 2020), and for the exclusion of repair factors from repeat-rich regions, i.e. heterochromatin in *Drosophila* cells (Chiolo et al., 2011) and the nucleolus in yeast (Torres-Rosell et al., 2005), to prevent unscheduled recombination events. Its involvement in the relocation of R-loops to the nuclear pore identified here is in line with its pivotal role in genome stability maintenance and contributes to the growing body of evidence pointing to a role of the Smc5/6 complex in R-loop metabolism.

1.4 | Factors mediating the interactions between SUMO-RPA-coated-R-loops and the nuclear pore

The nuclear side of the nuclear pore forms a platform available for the docking of numerous factors exerting NPC-related functions. Since none of the yeast nucleoporins possesses a DNA-binding domain, intermediate DNA-bound factors are required to mediate chromatin-NPC interactions. To this aim, RPA-coated ssDNA and SUMOylation constitute the recognition “label” that make the R-loop competent for the formation of physical interactions with NPC-associated factors.

1.4.1 | SIM-containing NPC-partners anchor R-loops to the nuclear pore

On the nuclear pore side, the anchoring of R-loops coated with SUMOylated proteins can be mediated by SIM-containing NPC partners, i.e. the STUbL Slx5/8. Indeed, the deletion of either of its two subunits leads to a decrease in *HSP104* co-fractionation with the nuclear pore (Article 2, Fig. 4I). The SUMO-protease Ulp1 is also able to recognize SUMOylated proteins through a distinct fold, i.e. its SUMO-binding domain (SBD, Elmore et al., 2011). Preliminary chromatin fractionation assays employing a thermosensitive Ulp1 mutant (*ulp1-333*), which abolishes both its activity and peripheral localisation (Li and Hochstrasser, 2003), were non

conclusive (data not shown). Moreover, the mutant allele used affects both functions of Ulp1, i.e. the maturation of the SUMO precursor, and the removal of the modifier from SUMOylated substrates (see Introduction, Box 1). This experiment should thereby be repeated in cells constitutively expressing processed SUMO, in order to distinguish between a requirement for Ulp1 SUMO-binding or deconjugating activities, rather than an indirect effect due to SUMO unavailability.

1.4.2 | Additional intermediate factors reinforce R-loop-NPC association?

Whether the STUbL directly recognises SUMO-RPA, or whether other intermediate factors or remodelling of the locus are involved, is still a matter of study. Indeed, the only partial suppression of *HSP104*-NPC co-fractionation in the absence of RPA SUMOylation (Article 2, Fig. 4H), or in *slx5/8* STUbL mutants (Article 2, Fig. 4I), would suggest the involvement of additional factors or events in the process. Moreover, in light of the typical *modus operandi* of E3 SUMO-ligases, which, lacking substrate specificity, rely mostly on their specific recruitment at a given locus, where they trigger a wave of SUMOylation which synergically interest protein groups (Psakhye and Jentsch, 2012), the presence of multiple SUMOylated substrates is to be expected. Notably, several transcription elongation and processing factors have been shown to be substrates of SUMOylation, e.g. Yra1 and Rrp6, together with the nuclear basket nucleoporin Mlp2 (Wohlschlegel et al., 2004). Proteomic analysis of SUMO-conjugates would allow the identification of eventual other factors whose sumoylation is increased in *tho*/heatshocked cells, shedding additional light on the mechanism and the consequences of R-loop forming loci relocation at the NPC. Notably, our preliminary proteomic observations identified an enrichment in sumoylated transcription related factors in *tho* cells, including the nucleoporins Mlp1-2 (data not shown). However, these preliminary assays were performed using a copper-induced tagged version of Smt3 for the affinity purification of SUMO-conjugates, which causes high levels of SUMO overexpression, even in absence of copper induction (Ulrich and Davies, 2009), thus likely introducing biases in the analysis. In the future, the utilization of our newly optimised protocol (see Article 2, Material and methods), employing a His-Smt3 construct under the control of its endogenous promoter, combined with the *ulp1-333* allele for the stabilisation of SUMO-conjugates, will provide more reliable experimental conditions to decipher the SUMOylation events occurring during R-loop gating.

1.5 | Comparison of NPC relocation pathways

As thoroughly discussed in the introduction of this manuscript, several mechanisms mediating DNA loci association to the nuclear pore have been characterised over the years, prompted either by high levels of transcription, or by the generation of stalled damaged intermediates unable to be taken over by canonical repair pathways. An important point of our study has been the comparison of the R-loop gating mechanism we identified with the other relocation pathways already known in the literature. Indeed, we were able to show that *HSP104*-NPC association is mediated by a stand-alone mechanism since we were able to detect this interaction in conditions in which other pathways are abolished, i.e. upon deletion of transcription and export factors (M. Rougemaille, D. Libri, personal communication), by preventing R-loop processing into DSBs (Article 2, Fig. 3A), and in absence of replication (Article 2, Fig. 3B).

1.5.1 | How does R-loop gating relates to transcription-dependent relocation?

Numerous studies have investigated the mechanisms underlying active gene relocation at the nuclear periphery, highlighting numerous, possibly overlapping or co-existing mechanisms. In this manuscript, I propose a categorisation of the studies carried out in yeast in three groups, highlighting three general mechanisms proposed to explain loci association to the NPC (see Introduction, paragraph 1.2.4, summarised in Table 1 below). Our pathway shows the highest similarity with studies proposing mRNP-mediated gene interaction with the NPC, supporting the fact that they may constitute overlapping pathways. However, further research needs to be devoted to formally investigate the involvement of promoter-proximal events in our model, to further understand the relationship between R-loop dependent and transcription dependent relocation.

Requirement for ...	Promoter driven model	Promoter de-repression model	mRNP mediated model	R-loop gating (our study)
Active transcription	No	No	Yes	Yes
Promoter sequences	Yes (UAS)	Yes (GRS)	Yes	?
Transcription factors	Yes	Yes	Yes	?
SAGA and TREX-2 complexes	Yes	No	No	No
3' end	No	No	Yes	Yes
RNA processing factors	No	No	Yes	?
SUMOylation	?	Yes	?	Yes
Ulp1/STUbL	?	Yes	?	Yes

Table 1. Comparison between the three proposed model for transcription-dependent relocation and the R-loop gating mechanism identified in this study. See Introduction, paragraph 1.2.4 for further details.

1.5.2 | Does R-loop gating requires damage formation and checkpoint activation?

To further demonstrate the distinction between R-loop and damage-dependent gene gating, we assessed the involvement of the checkpoint, which is required in the majority of the known damage-dependent pathways (see Introduction, paragraph 1.3.3 and Table 2 below). To do this, we employed the *mec1 sml1* double mutant, deleted for the Mitosis Entry Checkpoint kinase Mec1, required for cell cycle arrest and transcriptional responses to damaged or unreplicated DNA, and the ribonucleotide reductase inhibitor Sml1, which rescues the lethality of the single deletion of the essential *MEC1* gene (Zhao et al., 1998). Chromatin fractionation assays in the combined *tho mec1 sml1* triple mutant show only a mild effect on the *HSP104* phenotype (Fig. 5).

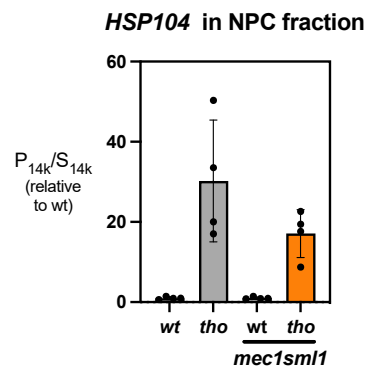


Figure 5. *HSP104* co-fractionation with the nuclear pore in a checkpoint mutant. qPCR-based quantification of the amount of DNA from the *HSP104* locus in heavy chromatin fractions from the indicated strains heat-shocked at 37°C for 15min (P_{14k}/S_{14k} , mean±SD, n=4, relative to wt; Methods as in Article 2).

This result seems to suggest a possible yet marginal involvement of checkpoint activation in R-loop gating, thus implying a contribution of R-loop processing into damage to trigger the relocation, which we excluded due to the lack of effect of nucleases deletions on the phenotype (Article 2, Fig. 3A). However, it is also important to note that Mec1-mediated phosphorylation has been shown to enhance Mms21 SUMOylation activity (Carlborg et al., 2015), raising the possibility that the mild decrease observed in the triple mutant could result from an indirect effect on Mms21 activity levels, rather than defects in checkpoint signalling.

This lack of requirement for checkpoint activation adds another common point with the mechanism described for the relocation of stalled replication forks encompassing CAG repeats tracks (Su et al., 2015). This suggests that the sensing of ssDNA-exposing structures and their compartmentalisation at the nuclear periphery is a general mechanism, even in the absence of effective breakage of the DNA. The rapid treatment of problematic structures could ensure their resolution and at the same time protect the neighbouring chromatin from potential disturbances that might fall upon them, preventing genetic instability.

When comparing the different mechanisms of damage relocation identified in yeast, it is evident how the main distinguishing factor is the requirement of the SUMO-ligase and the extent of SUMOylation to which the DNA-associated factors are subjected. This differential loci labelling likely determines the choice of the

pathway towards which such structures are channelled once they reach the nuclear pore microenvironment, to be resolved in the most conservative way possible.

Requirement for...	Unrepairable DSBs	Eroded Telomers	Stalled forks	R-loops (our study)
SUMO-ligases	Mms21 + Siz1/2	Siz1/2	Mms21	Mms21
SUMO-RPA	?	?	Yes	Yes
Poly-SUMOylation	Yes	?	No	No
Checkpoint activation	Yes	Yes	No	No
Slx5/8	Yes	Yes	Yes	Yes
NE anchor	Nup84 complex Or Mps3	?	Nup84 complex Nup1 C-term	?

Table 2. Comparison between damage-dependent and R-loop-dependent relocation mechanisms identified in budding yeast.

1.5.3 | Common features of NPC relocation pathways

When comparing all the mechanisms of targeting to the nuclear pore, it stands out that they all share a similar pattern, despite differences in triggers, factors involved, and functional consequences, with the following steps (depicted in Fig. 6):

1. The association of proteins (*in yellow*) to the nucleic acids of interest (e.g. transcription factors, RPA, or the ribosome),
2. Addition of other protein determinants (*in pink/dark blue*), i.e. recruitment of downstream factors, post-translational modifications, or translation initiation) which confer the competence for the interaction with NPC partners,
3. Recognition of the protein-nucleic acid complex by specific receptors (*in green*) characterized by increased residency in the nuclear pore microenvironment.

At this point, two possible fates are envisioned based on the nature of the relocated nucleic acid:

- A. For structures which require transit through the nuclear pore, i.e. mRNA or the nascent polypeptide coding for a nuclear protein, the receptor mediates the docking of protein-nucleic acid structures to the nuclear pore to facilitate the transit of the molecule across the channel.
- B. For stalled DNA damage intermediates that needs to recover the native DNA conformation, the NPC-associated receptor favours the remodelling of the molecule, e.g. by deSUMOylating or evicting DNA associated factors, to allow the locus to proceed towards the repair pathway of choice.

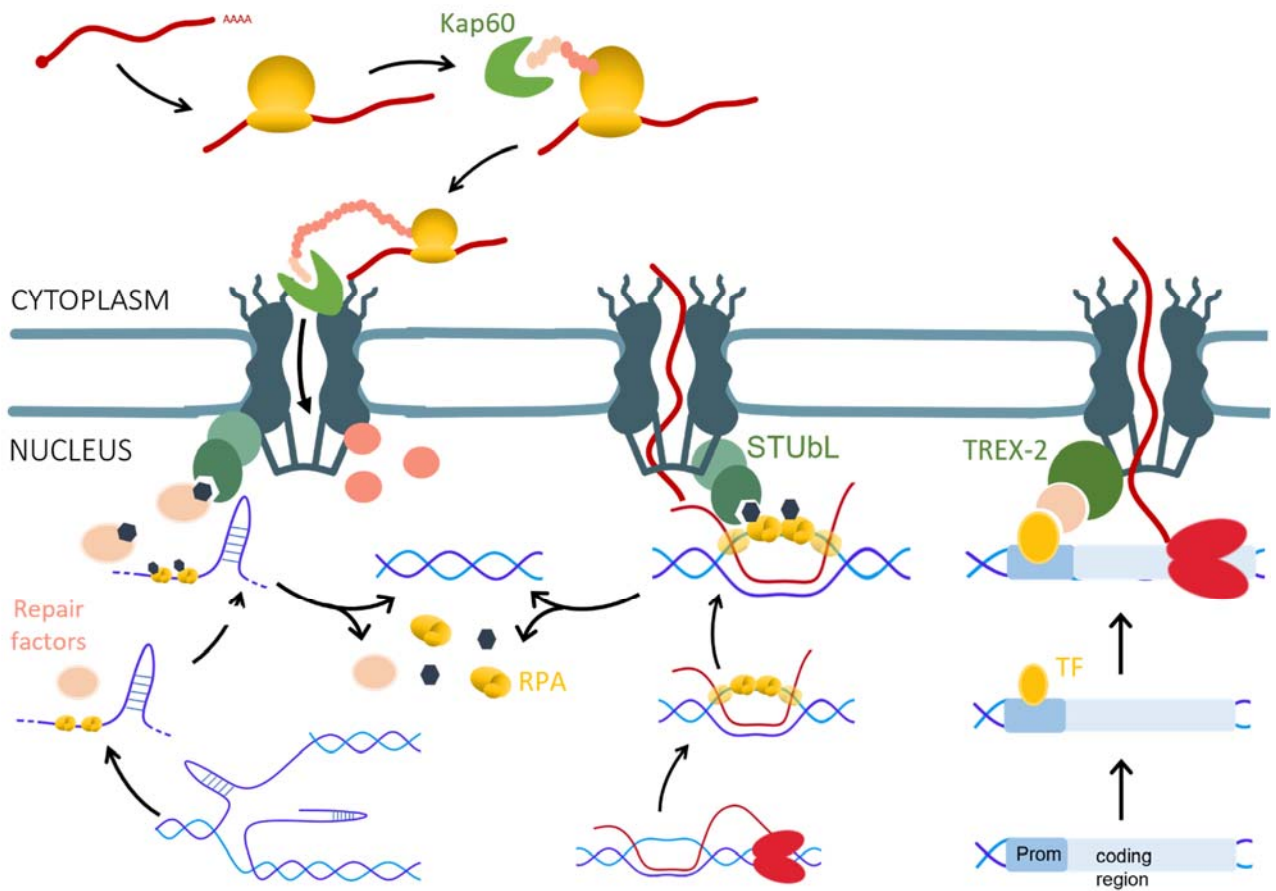


Figure 6. **General principle of nucleic acid relocation to the nuclear pore complex.** Four different relocation pathways are depicted in parallel: localized translation at the nuclear pore (top), damage dependent relocation (bottom left, exemplified by forks blocked at TNRs), R-loop-dependent relocation (center) and transcription dependent relocation (right, exemplified by the promoter-activation driven model).

2 | Functional consequences of gene / mRNA relocalisation to the pore

In this manuscript, I describe the mechanisms underlying localised translation of aggregate-prone proteins at the NPC, and R-loop dependent relocation to the nuclear periphery. We then proceeded to investigate how the proximity to the nuclear pore could influence such targets, in light of their potential toxicity for cell homeostasis.

2.1 | Physiological relevance of localised mRNA translation at the nuclear pore complex

My contribution to our study on co-translational assembly and localised translation of nucleoporins has been to set up strategies to determine the physiological relevance of mRNA association to the nuclear pore. To this aim, I employed two strategies to induce translation events of the Nup1 and/or Nup2 proteins at ectopic cellular locations, described in Article 1. Briefly, we fused the *NUP1* coding region with the terminator sequence of the *ASH1* mRNA, which is sufficient for the asymmetric distribution of mRNAs at the bud tip of dividing yeast cells (Long et al., 1997, Article 1, Fig. 4A-C). Secondly, the deletion of the translational repressor Hek2 was used to increase the number of *NUP1* mRNA molecules engaged in translation while not affecting their association to the NPC (Rouvière et al., 2018). Both strategies resulted in the formation of cytoplasmic aggregates containing Nup1 proteins (Article 1, Fig. 4D-E, S4B), which were able to trap the transport receptors required for NPC anchoring (Article 1, Fig. 4F). This supports the hypothesis that localised translation to the NPC has an important role in preventing the cytoplasmic aggregation of proteins, which is particularly relevant for FG Nups, like Nup1, due to their numerous hydrophobic domains.

Overall, our observation on the functional relevance of localised translation at the NPC highlights the importance of the faster import and assembly of proteins which could be prone to form potentially toxic cytoplasmic aggregates. Together, prolonged Hek2-dependent translational repression and localised translation would ensure a tight control on the production of these nucleoporins, limiting to the minimum the presence of the proteins in the cytoplasm, while maintaining a “reservoir” of ready-to-be-translated RNA molecules to rapidly supply new NPC components when NPC repair or assembly is required. In the future, it will be interesting to investigate how localised translation may be regulated in different circumstances, i.e. during the cell cycle or in response to stress or NPC damage. Moreover, the extension of our observations to the other factors whose RNAs have been found to associate to the nuclear pore (Article 1, Fig. 2A) may allow to better understand the general rules governing this phenomenon and its relevance in other aspects of nuclear homeostasis, beyond nuclear pore turnover.

2.2 | Consequences of loci relocation to the nuclear pore for R-loop metabolism

In our study, the artificial tethering of our R-loop forming reporter reduces the rate of R-loop dependent genetic instability in both wt and *tho* (*mft1Δ*) mutant cells (Article 2, Fig. 5C), supporting a protective effect of the nuclear pore against R-loop genotoxicity. Our result is consistent with the observed decrease of R-loop levels in loci artificially tethered to the NPC in nuclear basket (*mlp1Δ*) and *tho* (*hpr1Δ*) mutants (García-Benítez et al., 2017). Moreover, mutants for the factors required for R-loop relocation to the NPC display synthetic lethality with the *tho* mutant (Article 2, Fig. 5A), in agreement with the importance of this pathway for cell survival in stress conditions.

2.2.1 | Optimising the detection of R-loop dependent genetic instability

The protective effect of NPC against R-loop dependent genetic instability shown in our study is limited to conditions of low levels of transcription (Article 2, Fig. 5C). Indeed, the increased peripheral localisation of the reporter gene upon galactose induction (Article 2, Fig. S4A) complicates the comparison between tethered and untethered loci. The extent of transcription occurring in glucose-grown cells is anyway sufficient to generate R-loop dependent genetic instability (Article 2, Fig. S4B). Performing the tethering assay in the identified mutants in which R-loop-dependent relocation is abolished (*rfa1-D228Y*; *mms21-11*), should allow to genetically prevent the relocation of the locus even upon high levels of transcription. Another possible strategy could consist in the use of alternative transcription induction mechanisms, i.e. the VP16 activator, which was shown to promote the nucleoplasmic localisation of the *HXK1* locus (Taddei et al., 2006). However, the mechanism by which this localisation shift occurs has not been elucidated, and while VP16 activation is effective in preventing transcription-dependent relocalisation, this may not necessarily result in the suppression of R-loops-dependent repositioning.

2.2.2 | How does the nuclear pore act on R-loop metabolism?

Once acquiring the competency for the association to the nuclear pore, multiple scenarios can be envisioned for the fate of the (SUMO-)RPA-coated R-loop-forming locus relocated at the periphery.

Resolution of the pioneer R-loop?

Firstly, we could imagine that access to the nuclear pore interactome may modulate the fate of the pioneer R-loop whose coating by RPA triggers the spatial shift to the nuclear periphery. Facilitated access to ribonucleases or helicases may promote R-loop resolution, improving transcription efficiency. However, no particular enrichment of R-loop-related factors has been detected in previously published proteomic analyses of nuclear pores, including from the lab (Bretes et al., 2014; Lautier et al., 2021). A possible mechanism of R-loop resolution could be exerted by R-loop-associated RPA itself, which has been shown to interact with RNase H1 also in budding yeast (Maestroni et al., 2020). However, the ability of RPA to recruit RNase H1 would likely be unrelated to locus positioning.

RPA coating of the single strand DNA could also constitute an obstacle for R-loop resolution, preventing the re-annealing of the DNA even after the dissolution of the hybrid. Proximity to the nuclear pore could favour the eviction of RPA or other R-loop-associated SUMOylated factors, possibly through SUMO-dependent ubiquitylation.

Prevention of R-loop genotoxicity?

SUMOylation of factors associated to DNA lesions has been shown to prevent aberrant recombination and favour genome stability by channelling them towards alternative repair mechanisms (see Introduction, paragraph 1.3.3 and 1.3.5). Consistently, our tethering assay shows decreased genetic instability at NPC-tethered R-loop-forming loci (Article 2, Fig. 5C). This result could have multiple interpretations: firstly, rapid resolution of the R-loop, as speculated above, would prevent harmful outcomes from their accumulation. Secondly, proximity to the nuclear pore could isolate the R-loop from R-loop-processing factors, thus preventing its conversion into DSBs. Thirdly, R-loop repositioning to the nuclear pore would not affect the extent of their processing into breaks, but rather influence the choice of the repair pathway. Indeed, the nuclear pore microenvironment has been shown to be more permissive towards error-prone recombination pathways (*e.g.* NHEJ), which would not be detectable with our reporter scoring homologous recombination.

Prevention of further R-loop accumulation?

Finally, we could hypothesize that the relocation of R-loop forming loci to the nuclear pore would not directly affect the “pioneer” R-loop triggering relocation, but rather have a longer-term effect on the future potential R-loops that could form at the locus while transcription keeps occurring. The coupling of transcription with mRNA export could promote a more rapid eviction of the RNA from the transcription site, thus disadvantaging the RNA from annealing to the template DNA. Access to an environment unfavourable to excess R-loop formation may be particularly beneficial for inducible loci that are subjected to high transcription frequencies, with numerous transcripts produced in a short time-frame. Promoting stress-induced transcription while preventing the formation of obstacles or the “waste” of the newly produced transcript may result highly beneficial for cell viability and metabolic adaptation.

3 | Conservation of relocation mechanisms across organisms

3.1 | Conservation of R-loop gating

Although there are yet no direct evidences of R-loop-NPC interactions in other eukaryotes, a conservation of R-loop gating across organisms may be envisioned. Indeed, transcription-dependent relocation has been characterised for several loci in metazoans (see Introduction, paragraph 1.2.6). Moreover, light-regulated gene repositioning to the nuclear pore has been observed in *Arabidopsis* (Feng et al., 2014). The relationship between loci relocation and R-loops in these instances has not been investigated and cannot be excluded. A more systematic investigation of R-loop contribution to the establishment of loci-NPC interactions in the frame of transcriptional activation or transcriptional memory may shed additional light on the generality of this phenomenon. In light of the prevalent interaction of nucleoporins with stress-inducible, cell cycle-regulated or developmental genes (see Introduction, paragraph 1.2.7), NPC-related R-loop modulation at these loci may add an additional layer of regulation to safeguard the transcriptional programs necessary for cell survival and the establishment of cell identity.

Remarkably, chromatin-nucleoporin interactions have been shown to also occur in the nucleoplasm in metazoans (Maya Capelson et al., 2010; Kalverda et al., 2010). Whether the R-loop-Nup association would be similarly shifted to the nucleoplasm in these organisms is difficult to predict. Interestingly, the nuclear basket nucleoporins Nup50 and Nup153 have been identified in the R-loop proximal proteome of human cells (Mosler et al., 2021), while Wang et al. (2018) and Yan et al. (2022) identified other scaffold nucleoporins, among which the Y-complex nucleoporin Nup133. Furthermore, while the human nuclear basket nucleoporin TPR has been shown to prevent RNA-mediated replication stress, whether this action is specifically exerted at the nuclear periphery has not been investigated (Kosar et al., 2021). Finally, human cells depleted for the co-transcriptional pre-mRNA processing factor WDR33 show peripheral RPA foci, and an increase in R-loop levels, which is particularly enhanced in cells undergoing RPA exhaustion due to replication catastrophe (Teloni et al., 2019). All together, these data suggest a possible conservation of R-loop/RPA/Nucleoporin interactions in mammalian cells.

Contrary to the yeast genome, which is intron-poor, the majority of coding genes in metazoan genomes possess multiple introns. In this context, the preventing action of the spliceosome, both by RNA coating and in reducing the homology with the template by intron excision, may constitute the primary protective mechanism against R-loop dependent transcription perturbations, while nucleoporin function in R-loop metabolism, if confirmed, may be limited to specific subsets of genes, e.g. intronless genes or ORF with specific transcriptional requirements.

3.2 | Conservation of localised translation at NPCs

Major differences in nuclear pore assembly modalities between yeast and metazoan cells needs to be taken into account when speculating on the conservation of localised translation of nucleoporin-coding mRNAs across organisms. Indeed, while the only NPC assembly mechanism available in closed mitosis organisms is *de novo* assembly, metazoan cells additionally rely on post-mitotic assembly in concomitance with nuclear envelope reformation (see Introduction, paragraph 1.1.2). Moreover, *Drosophila* oocyte cells are equipped with a mechanism for storing nuclear pores in membranous compartments (*Annulate lamellae* - AL) that allows to maintain NPC density in dividing cells in the absence of active transcription. Notably, a subset of nucleoporin and transport factors-encoding mRNAs are enriched at the surface of AL (Hampoelz et al., 2019). It could therefore be speculated that, given the danger of storing soluble proteins that are highly hydrophobic and prone to aggregate formation, their storage in the form of mRNA, combined with their localised translation at the pore or at ALs, may ensure their safe supply at NPCs.

Interestingly, several screens systematically assessing the sub-cellular localization of mRNAs have been performed in fly and human cells (Lécuyer et al., 2007; Chouaib et al., 2020). While these studies did not identify nucleoporin transcripts as being locally translated at NPCs, the latter one identified two mRNAs (*ASPM* and *SPEN*) that localize at the nuclear envelope in a translation-dependent manner. Whether their localization also involves karyopherin-dependent association with NPCs remains to be addressed. This finding also illustrates that localised translation at NPCs could extend beyond NPC assembly, interesting other nuclear factors as shown in our study (Article 1. Fig, 2A). Further characterisation of these factors will be required to better understand whether NPC-associated translation may constitute a relevant, generalised mechanism for the establishment of the nuclear proteome and the maintenance of genome and cell homeostasis.

BIBLIOGRAPHY

- Abruzzi, K.C., Belostotsky, D.A., Chekanova, J.A., Dower, K., Rosbash, M., 2006. 3'-end formation signals modulate the association of genes with the nuclear periphery as well as mRNP dot formation. *EMBO J.* 25, 4253–4262. <https://doi.org/10.1038/sj.emboj.7601305>
- Aguilera, A., García-Muse, T., 2012. R loops: from transcription byproducts to threats to genome stability. *Mol. Cell* 46, 115–124. <https://doi.org/10.1016/j.molcel.2012.04.009>
- Aguilera, P., Whalen, J., Minguet, C., Churikov, D., Freudenreich, C., Simon, M.-N., Géli, V., 2020. The nuclear pore complex prevents sister chromatid recombination during replicative senescence. *Nature Communications* 11. <https://doi.org/10.1038/s41467-019-13979-5>
- Ahmed, S., Brickner, D.G., Light, W.H., Cajigas, I., McDonough, M., Froysheter, A.B., Volpe, T., Brickner, J.H., 2010. DNA zip codes control an ancient mechanism for gene targeting to the nuclear periphery. *Nat. Cell Biol.* 12, 111–118. <https://doi.org/10.1038/ncb2011>
- Ahn, J.H., Davis, E.S., Daugird, T.A., Zhao, S., Quiroga, I.Y., Uryu, H., Li, J., Storey, A.J., Tsai, Y.-H., Keeley, D.P., Mackintosh, S.G., Edmondson, R.D., Byrum, S.D., Cai, L., Tackett, A.J., Zheng, D., Legant, W.R., Phanstiel, D.H., Wang, G.G., 2021. Phase separation drives aberrant chromatin looping and cancer development. *Nature* 595, 591–595. <https://doi.org/10.1038/s41586-021-03662-5>
- Aiello, U., Challal, D., Wentzinger, G., Lengronne, A., Appanah, R., Pasero, P., Palancade, B., Libri, D., 2022. Sen1 is a key regulator of transcription-driven conflicts. *Mol Cell* S1097-2765(22)00604–9. <https://doi.org/10.1016/j.molcel.2022.06.021>
- Akey, C.W., Singh, D., Ouch, C., Echeverria, I., Nudelman, I., Varberg, J.M., Yu, Z., Fang, F., Shi, Y., Wang, J., Salzberg, D., Song, K., Xu, C., Gumbart, J.C., Suslov, S., Unruh, J., Jaspersen, S.L., Chait, B.T., Sali, A., Fernandez-Martinez, J., Ludtke, S.J., Villa, E., Rout, M.P., 2022. Comprehensive structure and functional adaptations of the yeast nuclear pore complex. *Cell* 185, 361-378.e25. <https://doi.org/10.1016/j.cell.2021.12.015>
- Akhtar, A., Gasser, S.M., 2007. The nuclear envelope and transcriptional control. *Nat. Rev. Genet.* 8, 507–517. <https://doi.org/10.1038/nrg2122>
- Albert, B., Léger-Silvestre, I., Normand, C., Gadal, O., 2012. Nuclear organization and chromatin dynamics in yeast: biophysical models or biologically driven interactions? *Biochim Biophys Acta* 1819, 468–481. <https://doi.org/10.1016/j.bbagr.2011.12.010>
- Albert, S., Schaffer, M., Beck, F., Mosalaganti, S., Asano, S., Thomas, H.F., Plitzko, J.M., Beck, M., Baumeister, W., Engel, B.D., 2017. Proteasomes tether to two distinct sites at the nuclear pore complex. *Proc Natl Acad Sci U S A* 114, 13726–13731. <https://doi.org/10.1073/pnas.1716305114>
- Alecki, C., Chiwara, V., Sanz, L.A., Grau, D., Arias Pérez, O., Boulier, E.L., Armache, K.-J., Chédin, F., Francis, N.J., 2020. RNA-DNA strand exchange by the *Drosophila* Polycomb complex PRC2. *Nat Commun* 11, 1781. <https://doi.org/10.1038/s41467-020-15609-x>
- Allegretti, M., Zimmerli, C.E., Rantos, V., Wilfling, F., Ronchi, P., Fung, H.K.H., Lee, C.-W., Hagen, W., Turoňová, B., Karius, K., Börmel, M., Zhang, X., Müller, C.W., Schwab, Y., Mahamid, J., Pfander, B., Kosinski, J., Beck, M., 2020. In-cell architecture of the nuclear pore and snapshots of its turnover. *Nature* 586, 796–800. <https://doi.org/10.1038/s41586-020-2670-5>
- Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., Liberi, G., 2012. Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell* 151, 835–846. <https://doi.org/10.1016/j.cell.2012.09.041>
- Amaral, N., Ryu, T., Li, X., Chiolo, I., 2017. Nuclear Dynamics of Heterochromatin Repair. *Trends Genet* 33, 86–100. <https://doi.org/10.1016/j.tig.2016.12.004>
- Andrulis, E.D., Neiman, A.M., Zappulla, D.C., Sternglanz, R., 1998. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* 394, 592–595. <https://doi.org/10.1038/29100>
- Appanah, R., Lones, E.C., Aiello, U., Libri, D., De Piccoli, G., 2020. Sen1 Is Recruited to Replication Forks via Ctf4 and Mrc1 and Promotes Genome Stability. *Cell Rep* 30, 2094-2105.e9. <https://doi.org/10.1016/j.celrep.2020.01.087>

- Arab, K., Karaulanov, E., Musheev, M., Trnka, P., Schäfer, A., Grummt, I., Niehrs, C., 2019. GADD45A binds R-loops and recruits TET1 to CpG island promoters. *Nature Genetics* 51, 217–223. <https://doi.org/10.1038/s41588-018-0306-6>
- Ariel, F., Lucero, L., Christ, A., Mammarella, M.F., Jegu, T., Veluchamy, A., Mariappan, K., Latrasse, D., Blein, T., Liu, C., Benhamed, M., Crespi, M., 2020. R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Molecular Cell* 77, 1055–1065.e4. <https://doi.org/10.1016/j.molcel.2019.12.015>
- Arudchandran, A., Cerritelli, S., Narimatsu, S., Itaya, M., Shin, D.Y., Shimada, Y., Crouch, R.J., 2000. The absence of ribonuclease H1 or H2 alters the sensitivity of *Saccharomyces cerevisiae* to hydroxyurea, caffeine and ethyl methanesulphonate: implications for roles of RNases H in DNA replication and repair. *Genes Cells* 5, 789–802. <https://doi.org/10.1046/j.1365-2443.2000.00373.x>
- Ashkenazy-Titelman, A., Shav-Tal, Y., Kehlenbach, R.H., 2020. Into the basket and beyond: the journey of mRNA through the nuclear pore complex. *Biochem J* 477, 23–44. <https://doi.org/10.1042/BCJ20190132>
- Aten, J.A., Stap, J., Krawczyk, P.M., van Oven, C.H., Hoebe, R.A., Essers, J., Kanaar, R., 2004. Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* 303, 92–95. <https://doi.org/10.1126/science.1088845>
- Auld, K.L., Silver, P.A., 2006. Transcriptional Regulation by the Proteasome as a Mechanism for Cellular Protein. *Cell Cycle* 5, 1503–1505. <https://doi.org/10.4161/cc.5.14.2979>
- Ayrapetov, M.K., Gursoy-Yuzugullu, O., Xu, C., Xu, Y., Price, B.D., 2014. DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin. *Proceedings of the National Academy of Sciences* 111, 9169–9174. <https://doi.org/10.1073/pnas.1403565111>
- Bader, A.S., Bushell, M., 2020. DNA:RNA hybrids form at DNA double-strand breaks in transcriptionally active loci. *Cell Death Dis* 11, 280. <https://doi.org/10.1038/s41419-020-2464-6>
- Bai, X.-T., Gu, B.-W., Yin, T., Niu, C., Xi, X.-D., Zhang, J., Chen, Z., Chen, S.-J., 2006. Trans-repressive effect of NUP98-PMX1 on PMX1-regulated c-FOS gene through recruitment of histone deacetylase 1 by FG repeats. *Cancer Res* 66, 4584–4590. <https://doi.org/10.1158/0008-5472.CAN-05-3101>
- Balk, B., Maicher, A., Dees, M., Klermund, J., Luke-Glaser, S., Bender, K., Luke, B., 2013. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* 20, 1199–1205. <https://doi.org/10.1038/nsmb.2662>
- Belotserkovskii, B.P., Hanawalt, P.C., 2022. Mechanism for R-loop formation remote from the transcription start site: Topological issues and possible facilitation by dissociation of RNA polymerase. *DNA Repair (Amst)* 110, 103275. <https://doi.org/10.1016/j.dnarep.2022.103275>
- Bermejo, R., Capra, T., Jossen, R., Colosio, A., Frattini, C., Carotenuto, W., Cocito, A., Doksani, Y., Klein, H., Gómez-González, B., Aguilera, A., Katou, Y., Shirahige, K., Foiani, M., 2011. The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. *Cell* 146, 233–246. <https://doi.org/10.1016/j.cell.2011.06.033>
- Bermejo, R., Kumar, A., Foiani, M., 2012. Preserving the genome by regulating chromatin association with the nuclear envelope. *Trends Cell Biol* 22, 465–473. <https://doi.org/10.1016/j.tcb.2012.05.007>
- Berto, A., Yu, J., Morchoisne-Bolhy, S., Bertipaglia, C., Vallee, R., Dumont, J., Ochsenein, F., Guerois, R., Doye, V., 2018. Disentangling the molecular determinants for Cenp-F localization to nuclear pores and kinetochores. *EMBO Rep* 19, e44742. <https://doi.org/10.15252/embr.201744742>
- Bhatia, V., Barroso, S.I., García-Rubio, M.L., Tumini, E., Herrera-Moyano, E., Aguilera, A., 2014. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* 511, 362–365. <https://doi.org/10.1038/nature13374>
- Blobel, G., 1985. Gene gating: a hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 82, 8527–8529. <https://doi.org/10.1073/pnas.82.24.8527>
- Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L., Carrico, R.J., 1986. Characterization of monoclonal antibody to DNA:RNA and its application to immunodetection of hybrids. *J. Immunol. Methods* 89, 123–130. [https://doi.org/10.1016/0022-1759\(86\)90040-2](https://doi.org/10.1016/0022-1759(86)90040-2)
- Bohnsack, M.T., Tollervey, D., Granneman, S., 2012. Identification of RNA helicase target sites by UV cross-linking and analysis of cDNA. *Methods Enzymol* 511, 275–288. <https://doi.org/10.1016/B978-0-12-396546-2.00013-9>

BIBLIOGRAPHY

- Bolger, T.A., Folkmann, A.W., Tran, E.J., Wenthe, S.R., 2008. The mRNA export factor Gle1 and inositol hexakisphosphate regulate distinct stages of translation. *Cell* 134, 624–633. <https://doi.org/10.1016/j.cell.2008.06.027>
- Bonnet, A., Grosso, A.R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S.C., Janbon, G., Géli, V., de Almeida, S.F., Palancade, B., 2017. Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability. *Molecular Cell* 67, 608–621.e6. <https://doi.org/10.1016/j.molcel.2017.07.002>
- Bou-Nader, C., Bothra, A., Garboczi, D.N., Leppla, S.H., Zhang, J., 2022. Structural basis of R-loop recognition by the S9.6 monoclonal antibody. *Nat Commun* 13, 1641. <https://doi.org/10.1038/s41467-022-29187-7>
- Bretes, H., Rouviere, J.O., Leger, T., Oeffinger, M., Devaux, F., Doye, V., Palancade, B., 2014. Sumoylation of the THO complex regulates the biogenesis of a subset of mRNPs. *Nucleic Acids Res.* 42, 5043–5058. <https://doi.org/10.1093/nar/gku124>
- Brickner, D.G., Ahmed, S., Meldi, L., Thompson, A., Light, W., Young, M., Hickman, T.L., Chu, F., Fabre, E., Brickner, J.H., 2012. Transcription factor binding to a DNA zip code controls interchromosomal clustering at the nuclear periphery. *Dev. Cell* 22, 1234–1246. <https://doi.org/10.1016/j.devcel.2012.03.012>
- Brickner, D.G., Brickner, J.H., 2010. Cdk phosphorylation of a nucleoporin controls localization of active genes through the cell cycle. *Mol. Biol. Cell* 21, 3421–3432. <https://doi.org/10.1091/mbc.E10-01-0065>
- Brickner, D.G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P.-C., Widom, J., Brickner, J.H., 2007. H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. *PLoS Biol.* 5, e81. <https://doi.org/10.1371/journal.pbio.0050081>
- Brickner, D.G., Randise-Hinchliff, C., Lebrun Corbin, M., Liang, J.M., Kim, S., Sump, B., D’Urso, A., Kim, S.H., Satomura, A., Schmit, H., Coukos, R., Hwang, S., Watson, R., Brickner, J.H., 2019. The Role of Transcription Factors and Nuclear Pore Proteins in Controlling the Spatial Organization of the Yeast Genome. *Dev. Cell* 49, 936–947.e4. <https://doi.org/10.1016/j.devcel.2019.05.023>
- Brickner, D.G., Sood, V., Tutucci, E., Coukos, R., Viets, K., Singer, R.H., Brickner, J.H., 2016. Subnuclear positioning and interchromosomal clustering of the GAL1-10 locus are controlled by separable, interdependent mechanisms. *Mol. Biol. Cell* 27, 2980–2993. <https://doi.org/10.1091/mbc.E16-03-0174>
- Brickner, J., 2017. Genetic and epigenetic control of the spatial organization of the genome. *Mol. Biol. Cell* 28, 364–369. <https://doi.org/10.1091/mbc.E16-03-0149>
- Brickner, J.H., Walter, P., 2004. Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS Biol.* 2, e342. <https://doi.org/10.1371/journal.pbio.0020342>
- Brosh, R.M., Li, J.L., Kenny, M.K., Karow, J.K., Cooper, M.P., Kureekattil, R.P., Hickson, I.D., Bohr, V.A., 2000. Replication protein A physically interacts with the Bloom’s syndrome protein and stimulates its helicase activity. *J. Biol. Chem.* 275, 23500–23508. <https://doi.org/10.1074/jbc.M001557200>
- Brown, R.E., Su, X.A., Fair, S., Wu, K., Verra, L., Jong, R., Andrykovich, K., Freudenreich, C.H., 2021. THO and TRAMP complexes prevent transcription-replication conflicts, DNA breaks, and CAG repeat contractions. <https://doi.org/10.1101/2021.12.06.471001>
- Cabal, G.G., Genovesio, A., Rodriguez-Navarro, S., Zimmer, C., Gadal, O., Lesne, A., Buc, H., Feuerbach-Fournier, F., Olivo-Marin, J.-C., Hurt, E.C., Nehrbass, U., 2006. SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature* 441, 770. <https://doi.org/10.1038/nature04752>
- Callan, H.G., Tomlin, S.G., 1950. Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. *Proc R Soc Lond B Biol Sci* 137, 367–378. <https://doi.org/10.1098/rspb.1950.0047>
- Cao, Y., Kogoma, T., 1993. Requirement for the polymerization and 5’→3’ exonuclease activities of DNA polymerase I in initiation of DNA replication at oriK sites in the absence of RecA in *Escherichia coli* rnhA mutants. *J Bacteriol* 175, 7254–7259. <https://doi.org/10.1128/jb.175.22.7254-7259.1993>
- Capelson, M., Doucet, C., Hetzer, M.W., 2010. Nuclear pore complexes: guardians of the nuclear genome. *Cold Spring Harb. Symp. Quant. Biol.* 75, 585–597. <https://doi.org/10.1101/sqb.2010.75.059>
- Capelson, Maya, Liang, Y., Schulte, R., Mair, W., Wagner, U., Hetzer, M.W., 2010. Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. *Cell* 140, 372–383. <https://doi.org/10.1016/j.cell.2009.12.054>

- Caridi, C.P., D'Agostino, C., Ryu, T., Zapotoczny, G., Delabaere, L., Li, X., Khodaverdian, V.Y., Amaral, N., Lin, E., Rau, A.R., Chiolo, I., 2018. Nuclear F-actin and myosins drive relocalization of heterochromatic breaks. *Nature* 559, 54–60. <https://doi.org/10.1038/s41586-018-0242-8>
- Carlborg, K.K., Kanno, T., Carter, S.D., Sjögren, C., 2015. Mec1-dependent phosphorylation of Mms21 modulates its SUMO ligase activity. *DNA Repair (Amst.)* 28, 83–92. <https://doi.org/10.1016/j.dnarep.2015.01.006>
- Carmody, S.R., Wenthe, S.R., 2009. mRNA nuclear export at a glance. *Journal of Cell Science* 122, 1933–1937. <https://doi.org/10.1242/jcs.041236>
- Carrasco-Salas, Y., Malapert, A., Sulthana, S., Molcrette, B., Chazot-Franguiadakis, L., Bernard, P., Chédin, F., Faivre-Moskalenko, C., Vanoosthuyse, V., 2019. The extruded non-template strand determines the architecture of R-loops. *Nucleic Acids Research* gkz341. <https://doi.org/10.1093/nar/gkz341>
- Casolari, J.M., Brown, C.R., Drubin, D.A., Rando, O.J., Silver, P.A., 2005. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. *Genes Dev.* 19, 1188–1198. <https://doi.org/10.1101/gad.1307205>
- Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., Silver, P.A., 2004. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* 117, 427–439.
- Castellano-Pozo, M., Santos-Pereira, J.M., Rondón, A.G., Barroso, S., Andújar, E., Pérez-Alegre, M., García-Muse, T., Aguilera, A., 2013. R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol. Cell* 52, 583–590. <https://doi.org/10.1016/j.molcel.2013.10.006>
- Castells-Roca, L., García-Martínez, J., Moreno, J., Herrero, E., Bellí, G., Pérez-Ortín, J.E., 2011. Heat shock response in yeast involves changes in both transcription rates and mRNA stabilities. *PLoS One* 6, e17272. <https://doi.org/10.1371/journal.pone.0017272>
- Castillo-Guzman, D., Chédin, F., 2021. Defining R-loop classes and their contributions to genome instability. *DNA Repair* 103182. <https://doi.org/10.1016/j.dnarep.2021.103182>
- Castillo-Guzman, D., Hartono, S.R., Sanz, L.A., Chédin, F., 2020. SF3B1-targeted Splicing Inhibition Triggers Global Alterations in Transcriptional Dynamics and R-Loop Metabolism. *bioRxiv* 2020.06.08.130583. <https://doi.org/10.1101/2020.06.08.130583>
- Ceccaldi, R., Sarangi, P., D'Andrea, A.D., 2016. The Fanconi anaemia pathway: new players and new functions. *Nat Rev Mol Cell Biol* 17, 337–349. <https://doi.org/10.1038/nrm.2016.48>
- Cerritelli, S.M., Crouch, R.J., 2019. RNase H2-RED carpets the path to eukaryotic RNase H2 functions. *DNA Repair (Amst.)* 102736. <https://doi.org/10.1016/j.dnarep.2019.102736>
- Cerritelli, S.M., Crouch, R.J., 2009. Ribonuclease H: the enzymes in eukaryotes. *FEBS J.* 276, 1494–1505. <https://doi.org/10.1111/j.1742-4658.2009.06908.x>
- Cerritelli, S.M., Frolova, E.G., Feng, C., Grinberg, A., Love, P.E., Crouch, R.J., 2003. Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell* 11, 807–815. [https://doi.org/10.1016/s1097-2765\(03\)00088-1](https://doi.org/10.1016/s1097-2765(03)00088-1)
- Chakraborty, P., 2020. New insight into the biology of R-loops. *Mutat. Res.* 821, 111711. <https://doi.org/10.1016/j.mrfmmm.2020.111711>
- Chakraborty, P., Grosse, F., 2011. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair* 10, 654–665. <https://doi.org/10.1016/j.dnarep.2011.04.013>
- Chakraborty, P., Huang, J.T.J., Hiom, K., 2018. DHX9 helicase promotes R-loop formation in cells with impaired RNA splicing. *Nature Communications* 9, 1–14. <https://doi.org/10.1038/s41467-018-06677-1>
- Chan, Y.A., Aristizabal, M.J., Lu, P.Y.T., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., Hieter, P., 2014. Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet.* 10, e1004288. <https://doi.org/10.1371/journal.pgen.1004288>
- Chang, E.Y.-C., Novoa, C.A., Aristizabal, M.J., Coulombe, Y., Segovia, R., Chaturvedi, R., Shen, Y., Keong, C., Tam, A.S., Jones, S.J.M., Masson, J.-Y., Kobor, M.S., Stirling, P.C., 2017. RECQ-like helicases Sgs1 and BLM regulate R-loop-associated genome instability. *J Cell Biol* 216, 3991–4005. <https://doi.org/10.1083/jcb.201703168>
- Chapman, K.B., Boeke, J.D., 1991. Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell* 65, 483–492. [https://doi.org/10.1016/0092-8674\(91\)90466-c](https://doi.org/10.1016/0092-8674(91)90466-c)

- Charifi, F., Churikov, D., Eckert-Boulet, N., Minguet, C., Jourquin, F., Hardy, J., Lisby, M., Simon, M.-N., Géli, V., 2021. Rad52 SUMOylation functions as a molecular switch that determines a balance between the Rad51- and Rad59-dependent survivors. *iScience* 24, 102231. <https://doi.org/10.1016/j.isci.2021.102231>
- Chaudhuri, J., Khuong, C., Alt, F.W., 2004. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* 430, 992–998. <https://doi.org/10.1038/nature02821>
- Chávez, S., Aguilera, A., 1997. The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes Dev.* 11, 3459–3470. <https://doi.org/10.1101/gad.11.24.3459>
- Chávez, S., Beilharz, T., Rondón, A.G., Erdjument-Bromage, H., Tempst, P., Sveistrup, J.Q., Lithgow, T., Aguilera, A., 2000. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae*. *EMBO J.* 19, 5824–5834. <https://doi.org/10.1093/emboj/19.21.5824>
- Chávez, S., García-Rubio, M., Prado, F., Aguilera, A., 2001. Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 21, 7054–7064. <https://doi.org/10.1128/MCB.21.20.7054-7064.2001>
- Chédin, F., 2016. Nascent Connections: R-Loops and Chromatin Patterning. *Trends Genet.* 32, 828–838. <https://doi.org/10.1016/j.tig.2016.10.002>
- Chedin, F., Benham, C.J., 2020. Emerging roles for R-loop structures in the management of topological stress. *J. Biol. Chem.* 295, 4684–4695. <https://doi.org/10.1074/jbc.REV119.006364>
- Chen, L., Chen, J.-Y., Zhang, X., Gu, Y., Xiao, R., Shao, C., Tang, P., Qian, H., Luo, D., Li, H., Zhou, Y., Zhang, D.-E., Fu, X.-D., 2017. R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol. Cell* 68, 745–757.e5. <https://doi.org/10.1016/j.molcel.2017.10.008>
- Chen, P.B., Chen, H.V., Acharya, D., Rando, O.J., Fazio, T.G., 2015. R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat. Struct. Mol. Biol.* 22, 999–1007. <https://doi.org/10.1038/nsmb.3122>
- Chiolo, I., Minoda, A., Colmenares, S.U., Polyzos, A., Costes, S.V., Karpen, G.H., 2011. Double-Strand Breaks in Heterochromatin Move Outside of a Dynamic HP1a Domain to Complete Recombinational Repair. *Cell* 144, 732–744. <https://doi.org/10.1016/j.cell.2011.02.012>
- Chiolo, I., Tang, J., Georgescu, W., Costes, S.V., 2013. Nuclear dynamics of radiation-induced foci in euchromatin and heterochromatin. *Mutat Res* 750, 10.1016/j.mrfmmm.2013.08.001. <https://doi.org/10.1016/j.mrfmmm.2013.08.001>
- Chouaib, R., Safieddine, A., Pichon, X., Imbert, A., Kwon, O.S., Samacoits, A., Traboulsi, A.-M., Robert, M.-C., Tsanov, N., Coleno, E., Poser, I., Zimmer, C., Hyman, A., Le Hir, H., Zibara, K., Peter, M., Mueller, F., Walter, T., Bertrand, E., 2020. A Dual Protein-mRNA Localization Screen Reveals Compartmentalized Translation and Widespread Co-translational RNA Targeting. *Dev Cell* 54, 773–791.e5. <https://doi.org/10.1016/j.devcel.2020.07.010>
- Chung, D.K.C., Chan, J.N.Y., Strecker, J., Zhang, W., Ebrahimi-Ardebili, S., Lu, T., Abraham, K.J., Durocher, D., Mekhail, K., 2015. Perinuclear tethers license telomeric DSBs for a broad kinesin- and NPC-dependent DNA repair process. *Nat Commun* 6, 7742. <https://doi.org/10.1038/ncomms8742>
- Churikov, D., Charifi, F., Eckert-Boulet, N., Silva, S., Simon, M.-N., Lisby, M., Géli, V., 2016. SUMO-Dependent Relocalization of Eroded Telomeres to Nuclear Pore Complexes Controls Telomere Recombination. *Cell Rep* 15, 1242–1253. <https://doi.org/10.1016/j.celrep.2016.04.008>
- Claussin, C., Chang, M., 2015. The many facets of homologous recombination at telomeres. *Microb Cell* 2, 308–321. <https://doi.org/10.15698/mic2015.09.224>
- Cloutier, S.C., Wang, S., Ma, W.K., Al Husini, N., Dhoondia, Z., Ansari, A., Pascuzzi, P.E., Tran, E.J., 2016. Regulated Formation of lncRNA-DNA Hybrids Enables Faster Transcriptional Induction and Environmental Adaptation. *Molecular Cell* 61, 393–404. <https://doi.org/10.1016/j.molcel.2015.12.024>
- Cockell, M., Gasser, S.M., 1999. Nuclear compartments and gene regulation. *Curr Opin Genet Dev* 9, 199–205. [https://doi.org/10.1016/S0959-437X\(99\)80030-6](https://doi.org/10.1016/S0959-437X(99)80030-6)
- Cohen, S., Puget, N., Lin, Y.-L., Clouaire, T., Aguirrebengoa, M., Rocher, V., Pasero, P., Canitrot, Y., Legube, G., 2018. Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nature Communications* 9, 533. <https://doi.org/10.1038/s41467-018-02894-w>

- Colombi, P., Webster, B.M., Fröhlich, F., Lusk, C.P., 2013. The transmission of nuclear pore complexes to daughter cells requires a cytoplasmic pool of Nsp1. *Journal of Cell Biology* 203, 215–232. <https://doi.org/10.1083/jcb.201305115>
- Coyle, J.H., Bor, Y.-C., Rekosh, D., Hammarskjöld, M.-L., 2011. The Tpr protein regulates export of mRNAs with retained introns that traffic through the Nxf1 pathway. *RNA* 17, 1344–1356. <https://doi.org/10.1261/rna.2616111>
- Cristini, A., Groh, M., Kristiansen, M.S., Gromak, N., 2018. RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. *Cell Reports* 23, 1891–1905. <https://doi.org/10.1016/j.celrep.2018.04.025>
- Cristini, A., Ricci, G., Britton, S., Salimbeni, S., Huang, S.N., Marinello, J., Calsou, P., Pommier, Y., Favre, G., Capranico, G., Gromak, N., Sordet, O., 2019. Dual Processing of R-Loops and Topoisomerase I Induces Transcription-Dependent DNA Double-Strand Breaks. *Cell Reports* 28, 3167–3181.e6. <https://doi.org/10.1016/j.celrep.2019.08.041>
- Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., Matunis, M.J., 2002. Proteomic analysis of the mammalian nuclear pore complex. *J Cell Biol* 158, 915–927. <https://doi.org/10.1083/jcb.200206106>
- Crossley, M.P., Bocek, M.J., Hamperl, S., Swigut, T., Cimprich, K.A., 2020. qDRIP: a method to quantitatively assess RNA-DNA hybrid formation genome-wide. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkaa500>
- Crossley, M.P., Brickner, J.R., Song, C., Zar, S.M.T., Maw, S.S., Chédin, F., Tsai, M.-S., Cimprich, K.A., 2021. Catalytically inactive, purified RNase H1: A specific and sensitive probe for RNA–DNA hybrid imaging. *Journal of Cell Biology* 220. <https://doi.org/10.1083/jcb.202101092>
- D’Alessandro, G., Whelan, D.R., Howard, S.M., Vitelli, V., Renaudin, X., Adamowicz, M., Iannelli, F., Jones-Weinert, C.W., Lee, M., Matti, V., Lee, W.T.C., Morten, M.J., Venkitaraman, A.R., Cejka, P., Rothenberg, E., d’Adda di Fagagna, F., 2018. BRCA2 controls DNA:RNA hybrid level at DSBs by mediating RNase H2 recruitment. *Nat Commun* 9, 5376. <https://doi.org/10.1038/s41467-018-07799-2>
- D’Angelo, M.A., Anderson, D.J., Richard, E., Hetzer, M.W., 2006. Nuclear pores form de novo from both sides of the nuclear envelope. *Science* 312, 440–443. <https://doi.org/10.1126/science.1124196>
- D’Angelo, M.A., Raices, M., Panowski, S.H., Hetzer, M.W., 2009. Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. *Cell* 136, 284–295. <https://doi.org/10.1016/j.cell.2008.11.037>
- Darst, R.P., Garcia, S.N., Koch, M.R., Pillus, L., 2008. Slx5 promotes transcriptional silencing and is required for robust growth in the absence of Sir2. *Mol. Cell. Biol.* 28, 1361–1372. <https://doi.org/10.1128/MCB.01291-07>
- Deaton, A.M., Bird, A., 2011. CpG islands and the regulation of transcription. *Genes Dev* 25, 1010–1022. <https://doi.org/10.1101/gad.2037511>
- Del Viso, F., Huang, F., Myers, J., Chalfant, M., Zhang, Y., Reza, N., Bewersdorf, J., Lusk, C.P., Khokha, M.K., 2016. Congenital Heart Disease Genetics Uncovers Context-Dependent Organization and Function of Nucleoporins at Cilia. *Dev Cell* 38, 478–492. <https://doi.org/10.1016/j.devcel.2016.08.002>
- Di Noia, J.M., Williams, G.T., Chan, D.T.Y., Buerstedde, J.-M., Baldwin, G.S., Neuberger, M.S., 2007. Dependence of antibody gene diversification on uracil excision. *Journal of Experimental Medicine* 204, 3209–3219. <https://doi.org/10.1084/jem.20071768>
- Diepinois, G., Iglesias, N., Stutz, F., 2006. Cotranscriptional Recruitment to the mRNA Export Receptor Mex67p Contributes to Nuclear Pore Anchoring of Activated Genes. *Molecular and Cellular Biology* 26, 7858–7870. <https://doi.org/10.1128/MCB.00870-06>
- Diepinois, G., Stutz, F., 2010. Connecting the transcription site to the nuclear pore: a multi-tether process that regulates gene expression. *Journal of Cell Science* 123, 1989–1999. <https://doi.org/10.1242/jcs.053694>
- Dilworth, D.J., Suprpto, A., Padovan, J.C., Chait, B.T., Wozniak, R.W., Rout, M.P., Aitchison, J.D., 2001. Nup2p Dynamically Associates with the Distal Regions of the Yeast Nuclear Pore Complex. *The Journal of Cell Biology* 153, 1465–1478. <https://doi.org/10.1083/jcb.153.7.1465>
- Dingwall, C., Lomonosoff, G.P., Laskey, R.A., 1981. High sequence specificity of micrococcal nuclease. *Nucleic Acids Res* 9, 2659–2673. <https://doi.org/10.1093/nar/9.12.2659>
- Dion, V., Kalck, V., Horigome, C., Towbin, B.D., Gasser, S.M., 2012. Increased mobility of double-strand breaks requires Mec1, Rad9 and the homologous recombination machinery. *Nat Cell Biol* 14, 502–509. <https://doi.org/10.1038/ncb2465>

BIBLIOGRAPHY

- Domínguez-Sánchez, M.S., Sáez, C., Japón, M.A., Aguilera, A., Luna, R., 2011. Differential expression of THOC1 and ALY mRNP biogenesis/export factors in human cancers. *BMC Cancer* 11, 77. <https://doi.org/10.1186/1471-2407-11-77>
- Doye, V., Wepf, R., Hurt, E.C., 1994. A novel nuclear pore protein Nup133p with distinct roles in poly(A)⁺ RNA transport and nuclear pore distribution. *The EMBO Journal* 13, 6062–6075. <https://doi.org/10.1002/j.1460-2075.1994.tb06953.x>
- Drolet, M., 2006. Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Mol. Microbiol.* 59, 723–730. <https://doi.org/10.1111/j.1365-2958.2005.05006.x>
- Drolet, M., Brochu, J., 2019. R-loop-dependent replication and genomic instability in bacteria. *DNA Repair (Amst.)* 84, 102693. <https://doi.org/10.1016/j.dnarep.2019.102693>
- Drolet, M., Phoenix, P., Menzel, R., Massé, E., Liu, L.F., Crouch, R.J., 1995. Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3526–3530. <https://doi.org/10.1073/pnas.92.8.3526>
- Dubarry, M., Loïdice, I., Chen, C.L., Thermes, C., Taddei, A., 2011. Tight protein-DNA interactions favor gene silencing. *Genes Dev* 25, 1365–1370. <https://doi.org/10.1101/gad.611011>
- Dueva, R., Iliakis, G., 2020. Replication protein A: a multifunctional protein with roles in DNA replication, repair and beyond. *NAR Cancer* 2, zcaa022. <https://doi.org/10.1093/narcan/zcaa022>
- Dultz, E., Ellenberg, J., 2010. Live imaging of single nuclear pores reveals unique assembly kinetics and mechanism in interphase. *Journal of Cell Biology* 191, 15–22. <https://doi.org/10.1083/jcb.201007076>
- Dultz, E., Wojtynek, M., Medalia, O., Onischenko, E., 2022. The Nuclear Pore Complex: Birth, Life, and Death of a Cellular Behemoth. *Cells* 11, 1456. <https://doi.org/10.3390/cells11091456>
- Dumelie, J.G., Jaffrey, S.R., 2017. Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. *Elife* 6. <https://doi.org/10.7554/eLife.28306>
- Dunn, K., Griffith, J.D., 1980. The presence of RNA in a double helix inhibits its interaction with histone protein. *Nucleic Acids Res* 8, 555–566. <https://doi.org/10.1093/nar/8.3.555>
- Duquette, M.L., Handa, P., Vincent, J.A., Taylor, A.F., Maizels, N., 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* 18, 1618–1629. <https://doi.org/10.1101/gad.1200804>
- El Hage, A., French, S.L., Beyer, A.L., Tollervey, D., 2010. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* 24, 1546–1558. <https://doi.org/10.1101/gad.573310>
- El Hage, A., Webb, S., Kerr, A., Tollervey, D., 2014. Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet.* 10, e1004716. <https://doi.org/10.1371/journal.pgen.1004716>
- Elmore, Z.C., Donaher, M., Matson, B.C., Murphy, H., Westerbeck, J.W., Kerscher, O., 2011. Sumo-dependent substrate targeting of the SUMO protease Ulp1. *BMC Biol* 9, 74. <https://doi.org/10.1186/1741-7007-9-74>
- Enenkel, C., Lehmann, A., Kloetzel, P.M., 1998. Subcellular distribution of proteasomes implicates a major location of protein degradation in the nuclear envelope-ER network in yeast. *EMBO J* 17, 6144–6154. <https://doi.org/10.1093/emboj/17.21.6144>
- Feng, C.-M., Qiu, Y., Van Buskirk, E.K., Yang, E.J., Chen, M., 2014. Light-regulated gene repositioning in Arabidopsis. *Nat Commun* 5, 3027. <https://doi.org/10.1038/ncomms4027>
- Feretzi, M., Pospisilova, M., Valador Fernandes, R., Lunardi, T., Krejci, L., Lingner, J., 2020. RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature*. <https://doi.org/10.1038/s41586-020-2815-6>
- Fischer, T., Strässer, K., Rácz, A., Rodríguez-Navarro, S., Oppizzi, M., Ihrig, P., Lechner, J., Hurt, E., 2002. The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *EMBO J.* 21, 5843–5852. <https://doi.org/10.1093/emboj/cdf590>
- Fiserova, J., Kiseleva, E., Goldberg, M.W., 2009. Nuclear envelope and nuclear pore complex structure and organization in tobacco BY-2 cells. *Plant J* 59, 243–255. <https://doi.org/10.1111/j.1365-313X.2009.03865.x>

- Folz, H., Niño, C.A., Taranum, S., Caesar, S., Latta, L., Waharte, F., Salamero, J., Schlenstedt, G., Dargemont, C., 2019. SUMOylation of the nuclear pore complex basket is involved in sensing cellular stresses. *Journal of Cell Science* 132, jcs224279. <https://doi.org/10.1242/jcs.224279>
- Fontoura, B.M., Blobel, G., Matunis, M.J., 1999. A conserved biogenesis pathway for nucleoporins: proteolytic processing of a 186-kilodalton precursor generates Nup98 and the novel nucleoporin, Nup96. *J Cell Biol* 144, 1097–1112. <https://doi.org/10.1083/jcb.144.6.1097>
- Forey, R., Barthe, A., Tittel-Elmer, M., Wery, M., Barrault, M.-B., Ducrot, C., Seeber, A., Krietenstein, N., Szachnowski, U., Skrzypczak, M., Ginalski, K., Rowicka, M., Cobb, J.A., Rando, O.J., Soutourina, J., Werner, M., Dubrana, K., Gasser, S.M., Morillon, A., Pasero, P., Lengronne, A., Poli, J., 2021. A Role for the Mre11-Rad50-Xrs2 Complex in Gene Expression and Chromosome Organization. *Molecular Cell* 81, 183-197.e6. <https://doi.org/10.1016/j.molcel.2020.11.010>
- Freudenreich, C.H., Su, X.A., 2016. Relocalization of DNA lesions to the nuclear pore complex. *FEMS Yeast Res.* 16. <https://doi.org/10.1093/femsyr/fow095>
- Gall, J.G., 1967. Octagonal nuclear pores. *J Cell Biol* 32, 391–399. <https://doi.org/10.1083/jcb.32.2.391>
- Gallardo, M., Aguilera, A., 2001. A new hyperrecombination mutation identifies a novel yeast gene, THP1, connecting transcription elongation with mitotic recombination. *Genetics* 157, 79–89.
- Gallardo, M., Luna, R., Erdjument-Bromage, H., Tempst, P., Aguilera, A., 2003. Nab2p and the Thp1p-Sac3p complex functionally interact at the interface between transcription and mRNA metabolism. *J. Biol. Chem.* 278, 24225–24232. <https://doi.org/10.1074/jbc.M302900200>
- Galy, V., Gadai, O., Fromont-Racine, M., Romano, A., Jacquier, A., Nehrbass, U., 2004. Nuclear Retention of Unspliced mRNAs in Yeast Is Mediated by Perinuclear Mlp1. *Cell* 116, 63–73. [https://doi.org/10.1016/S0092-8674\(03\)01026-2](https://doi.org/10.1016/S0092-8674(03)01026-2)
- Galy, V., Olivo-Marin, J.C., Scherthan, H., Doye, V., Rascalou, N., Nehrbass, U., 2000. Nuclear pore complexes in the organization of silent telomeric chromatin. *Nature* 403, 108–112. <https://doi.org/10.1038/47528>
- Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J.L., Li, X., 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* 25, 2041–2056. <https://doi.org/10.1101/gad.17010011>
- García Fernández, F., Lemos, B., Khalil, Y., Batrin, R., Haber, J.E., Fabre, E., 2021. Modified chromosome structure caused by phosphomimetic H2A modulates the DNA damage response by increasing chromatin mobility in yeast. *J Cell Sci* 134, jcs258500. <https://doi.org/10.1242/jcs.258500>
- García-Benítez, F., Gaillard, H., Aguilera, A., 2017. Physical proximity of chromatin to nuclear pores prevents harmful R loop accumulation contributing to maintain genome stability. *Proc. Natl. Acad. Sci. U.S.A.* 114, 10942–10947. <https://doi.org/10.1073/pnas.1707845114>
- García-Pichardo, D., Cañas, J.C., García-Rubio, M.L., Gómez-González, B., Rondón, A.G., Aguilera, A., 2017. Histone Mutants Separate R Loop Formation from Genome Instability Induction. *Mol. Cell* 66, 597-609.e5. <https://doi.org/10.1016/j.molcel.2017.05.014>
- García-Rubio, M.L., Pérez-Calero, C., Barroso, S.I., Tumini, E., Herrera-Moyano, E., Rosado, I.V., Aguilera, A., 2015. The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLoS Genet.* 11, e1005674. <https://doi.org/10.1371/journal.pgen.1005674>
- Gavaldá, S., Gallardo, M., Luna, R., Aguilera, A., 2013. R-loop mediated transcription-associated recombination in trf4Δ mutants reveals new links between RNA surveillance and genome integrity. *PLoS ONE* 8, e65541. <https://doi.org/10.1371/journal.pone.0065541>
- Gillespie, P.J., Khoudoli, G.A., Stewart, G., Swedlow, J.R., Blow, J.J., 2007. ELYS/MEL-28 chromatin association coordinates nuclear pore complex assembly and replication licensing. *Curr Biol* 17, 1657–1662. <https://doi.org/10.1016/j.cub.2007.08.041>
- Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., Chédin, F., 2012. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell* 45, 814–825. <https://doi.org/10.1016/j.molcel.2012.01.017>
- Glover, D.M., Hogness, D.S., 1977. A novel arrangement of the 18S and 28S sequences in a repeating unit of *Drosophila melanogaster* rDNA. *Cell* 10, 167–176. [https://doi.org/10.1016/0092-8674\(77\)90212-4](https://doi.org/10.1016/0092-8674(77)90212-4)
- Gomar-Alba, M., Mendoza, M., 2019. Modulation of Cell Identity by Modification of Nuclear Pore Complexes. *Front Genet* 10, 1301. <https://doi.org/10.3389/fgene.2019.01301>

- Gomar-Alba, M., Pozharskaia, V., Cichocki, B., Schaal, C., Kumar, A., Jacquet, B., Charvin, G., Igual, J.C., Mendoza, M., 2022. Nuclear pore complex acetylation regulates mRNA export and cell cycle commitment in budding yeast. *EMBO J* e2021110271. <https://doi.org/10.15252/embj.2021110271>
- Gómez-González, B., García-Rubio, M., Bermejo, R., Gaillard, H., Shirahige, K., Marín, A., Foiani, M., Aguilera, A., 2011. Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *EMBO J* 30, 3106–3119. <https://doi.org/10.1038/emboj.2011.206>
- Gonzalez-Estevez, A., Verrico, A., Orniacki, C., Reina-San-Martin, B., Doye, V., 2021. Integrity of the short arm of the nuclear pore Y-complex is required for mouse embryonic stem cell growth and differentiation. *J Cell Sci* 134, jcs258340. <https://doi.org/10.1242/jcs.258340>
- Gorthi, A., Romero, J.C., Loranc, E., Cao, L., Lawrence, L.A., Goodale, E., Iniguez, A.B., Bernard, X., Masamsetti, V.P., Roston, S., Lawlor, E.R., Toretsky, J.A., Stegmaier, K., Lessnick, S.L., Chen, Y., Bishop, A.J.R., 2018. EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature* 555, 387–391. <https://doi.org/10.1038/nature25748>
- Green, E.M., Jiang, Y., Joyner, R., Weis, K., 2012. A negative feedback loop at the nuclear periphery regulates GAL gene expression. *Mol Biol Cell* 23, 1367–1375. <https://doi.org/10.1091/mbc.E11-06-0547>
- Gruber, S., 2018. SMC complexes sweeping through the chromosome: going with the flow and against the tide. *Curr Opin Microbiol* 42, 96–103. <https://doi.org/10.1016/j.mib.2017.10.004>
- Grunseich, C., Wang, I.X., Watts, J.A., Burdick, J.T., Guber, R.D., Zhu, Z., Bruzel, A., Lanman, T., Chen, K., Schindler, A.B., Edwards, N., Ray-Chaudhury, A., Yao, J., Lehky, T., Piszczek, G., Crain, B., Fischbeck, K.H., Cheung, V.G., 2018. Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. *Molecular Cell* 69, 426–437.e7. <https://doi.org/10.1016/j.molcel.2017.12.030>
- Gwizdek, C., Iglesias, N., Rodriguez, M.S., Ossareh-Nazari, B., Hobeika, M., Divita, G., Stutz, F., Dargemont, C., 2006. Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription. *Proc Natl Acad Sci U S A* 103, 16376–16381. <https://doi.org/10.1073/pnas.0607941103>
- Hahn, J.-S., Hu, Z., Thiele, D.J., Iyer, V.R., 2004. Genome-wide analysis of the biology of stress responses through heat shock transcription factor. *Mol Cell Biol* 24, 5249–5256. <https://doi.org/10.1128/MCB.24.12.5249-5256.2004>
- Hakhverdyan, Z., Molloy, K.R., Keegan, S., Herricks, T., Lepore, D.M., Munson, M., Subbotin, R.I., Fenyő, D., Aitchison, J.D., Fernandez-Martinez, J., Chait, B.T., Rout, M.P., 2021. Dissecting the Structural Dynamics of the Nuclear Pore Complex. *Molecular Cell* 81, 153–165.e7. <https://doi.org/10.1016/j.molcel.2020.11.032>
- Halász, L., Karányi, Z., Boros-Oláh, B., Kuik-Rózsa, T., Sipos, É., Nagy, É., Mosolygó-L, Á., Mázló, A., Rajnavölgyi, É., Halmos, G., Székvölgyi, L., 2017. RNA-DNA hybrid (R-loop) immunoprecipitation mapping: an analytical workflow to evaluate inherent biases. *Genome Res* 27, 1063–1073. <https://doi.org/10.1101/gr.219394.116>
- Hamperl, S., Bocek, M.J., Saldivar, J.C., Swigut, T., Cimprich, K.A., 2017. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* 170, 774–786.e19. <https://doi.org/10.1016/j.cell.2017.07.043>
- Hamperl, S., Cimprich, K.A., 2014. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair* 19, 84–94. <https://doi.org/10.1016/j.dnarep.2014.03.023>
- Hampelz, B., Andres-Pons, A., Kastiris, P., Beck, M., 2019. Structure and Assembly of the Nuclear Pore Complex. *Annu. Rev. Biophys.* 48, 515–536. <https://doi.org/10.1146/annurev-biophys-052118-115308>
- Hartono, S.R., Malapert, A., Legros, P., Bernard, P., Chédin, F., Vanoosthuyse, V., 2018. The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast. *J. Mol. Biol.* 430, 272–284. <https://doi.org/10.1016/j.jmb.2017.12.016>
- Hatchi, E., Skourti-Stathaki, K., Ventz, S., Pinello, L., Yen, A., Kamieniarz-Gdula, K., Dimitrov, S., Pathania, S., McKinney, K.M., Eaton, M.L., Kellis, M., Hill, S.J., Parmigiani, G., Proudfoot, N.J., Livingston, D.M., 2015. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol Cell* 57, 636–647. <https://doi.org/10.1016/j.molcel.2015.01.011>
- Hayakawa, A., Babour, A., Sengmanivong, L., Dargemont, C., 2012. Ubiquitylation of the nuclear pore complex controls nuclear migration during mitosis in *S. cerevisiae*. *Journal of Cell Biology* 196, 19–27. <https://doi.org/10.1083/jcb.201108124>

- Herbert, S., Brion, A., Arbona, J.-M., Lelek, M., Veillet, A., Lelandais, B., Parmar, J., Fernández, F.G., Almayrac, E., Khalil, Y., Birgy, E., Fabre, E., Zimmer, C., 2017. Chromatin stiffening underlies enhanced locus mobility after DNA damage in budding yeast. *EMBO J* 36, 2595–2608. <https://doi.org/10.15252/embj.201695842>
- Heun, P., Laroche, T., Shimada, K., Furrer, P., Gasser, S.M., 2001. Chromosome dynamics in the yeast interphase nucleus. *Science* 294, 2181–2186. <https://doi.org/10.1126/science.1065366>
- Hodroj, D., Serhal, K., Maiorano, D., 2017. Ddx19 links mRNA nuclear export with progression of transcription and replication and suppresses genomic instability upon DNA damage in proliferating cells. *Nucleus* 8, 489–495. <https://doi.org/10.1080/19491034.2017.1348448>
- Hodson, C., O'Rourke, J.J., Twest, S. van, Murphy, V.J., Dunn, E., Deans, A.J., 2018. FANCM-family branchpoint translocases remove co-transcriptional R-loops. <https://doi.org/10.1101/248161>
- Holt, I.J., 2019. The mitochondrial R-loop. *Nucleic Acids Res* 47, 5480–5489. <https://doi.org/10.1093/nar/gkz277>
- Hong, X., Cadwell, G.W., Kogoma, T., 1995. Escherichia coli RecG and RecA proteins in R-loop formation. *EMBO J* 14, 2385–2392. <https://doi.org/10.1002/j.1460-2075.1995.tb07233.x>
- Horigome, C., Bustard, D.E., Marcomini, I., Delgosaie, N., Tsai-Pflugfelder, M., Cobb, J.A., Gasser, S.M., 2016. PolySUMOylation by Siz2 and Mms21 triggers relocation of DNA breaks to nuclear pores through the Slx5/Slx8 STUbL. *Genes Dev.* 30, 931–945. <https://doi.org/10.1101/gad.277665.116>
- Horigome, C., Oma, Y., Konishi, T., Schmid, R., Marcomini, I., Hauer, M.H., Dion, V., Harata, M., Gasser, S.M., 2014. SWR1 and INO80 chromatin remodelers contribute to DNA double-strand break perinuclear anchorage site choice. *Mol Cell* 55, 626–639. <https://doi.org/10.1016/j.molcel.2014.06.027>
- Horigome, C., Unozawa, E., Ooki, T., Kobayashi, T., 2019. Ribosomal RNA gene repeats associate with the nuclear pore complex for maintenance after DNA damage. *PLoS Genet* 15, e1008103. <https://doi.org/10.1371/journal.pgen.1008103>
- Huertas, P., Aguilera, A., 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* 12, 711–721. <https://doi.org/10.1016/j.molcel.2003.08.010>
- li, T., Fung, J., Mullen, J.R., Brill, S.J., 2007. The yeast Slx5-Slx8 DNA integrity complex displays ubiquitin ligase activity. *Cell Cycle* 6, 2800–2809. <https://doi.org/10.4161/cc.6.22.4882>
- Ishii, K., Arib, G., Lin, C., Van Houwe, G., Laemmli, U.K., 2002. Chromatin Boundaries in Budding Yeast: The Nuclear Pore Connection. *Cell* 109, 551–562. [https://doi.org/10.1016/S0092-8674\(02\)00756-0](https://doi.org/10.1016/S0092-8674(02)00756-0)
- Itoh, T., Tomizawa, J., 1980. Formation of an RNA primer for initiation of replication of Cole1 DNA by ribonuclease H. *PNAS* 77, 2450–2454. <https://doi.org/10.1073/pnas.77.5.2450>
- Jakob, B., Splinter, J., Conrad, S., Voss, K.-O., Zink, D., Durante, M., Löbrich, M., Taucher-Scholz, G., 2011. DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. *Nucleic Acids Research* 39, 6489–6499. <https://doi.org/10.1093/nar/gkr230>
- Jiang, F., Doudna, J.A., 2017. CRISPR-Cas9 Structures and Mechanisms. *Annu Rev Biophys* 46, 505–529. <https://doi.org/10.1146/annurev-biophys-062215-010822>
- Jimeno, S., Rondón, A.G., Luna, R., Aguilera, A., 2002. The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *EMBO J.* 21, 3526–3535. <https://doi.org/10.1093/emboj/cdf335>
- Joseph, J., Dasso, M., 2008. The nucleoporin Nup358 associates with and regulates interphase microtubules. *FEBS Letters* 582, 190–196. <https://doi.org/10.1016/j.febslet.2007.11.087>
- Joseph, J., Liu, S.-T., Jablonski, S.A., Yen, T.J., Dasso, M., 2004. The RanGAP1-RanBP2 complex is essential for microtubule-kinetochore interactions in vivo. *Curr Biol* 14, 611–617. <https://doi.org/10.1016/j.cub.2004.03.031>
- Kadota, S., Ou, J., Shi, Y., Lee, J.T., Sun, J., Yildirim, E., 2020. Nucleoporin 153 links nuclear pore complex to chromatin architecture by mediating CTCF and cohesin binding. *Nat Commun* 11, 2606. <https://doi.org/10.1038/s41467-020-16394-3>
- Kalverda, B., Pickersgill, H., Shloma, V.V., Fornerod, M., 2010. Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. *Cell* 140, 360–371. <https://doi.org/10.1016/j.cell.2010.01.011>

- Kasahara, M., Clikeman, J.A., Bates, D.B., Kogoma, T., 2000. RecA protein-dependent R-loop formation in vitro. *Genes Dev.* 14, 360–365. <https://doi.org/10.1101/gad.14.3.360>
- Kasper, L.H., Brindle, P.K., Schnabel, C.A., Pritchard, C.E., Cleary, M.L., van Deursen, J.M., 1999. CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98-HOXA9 oncogenicity. *Mol Cell Biol* 19, 764–776. <https://doi.org/10.1128/MCB.19.1.764>
- Kee, H.L., Dishinger, J.F., Blasius, T.L., Liu, C.-J., Margolis, B., Verhey, K.J., 2012. A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia. *Nat Cell Biol* 14, 431–437. <https://doi.org/10.1038/ncb2450>
- Khadaroo, B., Teixeira, M.T., Luciano, P., Eckert-Boulet, N., Germann, S.M., Simon, M.N., Gallina, I., Abdallah, P., Gilson, E., Géli, V., Lisby, M., 2009. The DNA damage response at eroded telomeres and tethering to the nuclear pore complex. *Nat Cell Biol* 11, 980–987. <https://doi.org/10.1038/ncb1910>
- Kim, S.J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B.D., Hogan, J.A., Upla, P., Chemmama, I.E., Pellarin, R., Echeverria, I., Shivaraju, M., Chaudhury, A.S., Wang, J., Williams, R., Unruh, J.R., Greenberg, C.H., Jacobs, E.Y., Yu, Z., de la Cruz, M.J., Mironska, R., Stokes, D.L., Aitchison, J.D., Jarrold, M.F., Gerton, J.L., Ludtke, S.J., Akey, C.W., Chait, B.T., Sali, A., Rout, M.P., 2018. Integrative structure and functional anatomy of a nuclear pore complex. *Nature* 555, 475–482. <https://doi.org/10.1038/nature26003>
- König, F., Schubert, T., Längst, G., 2017. The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences. *PLoS ONE* 12, e0178875. <https://doi.org/10.1371/journal.pone.0178875>
- Kosar, M., Giannattasio, M., Piccini, D., Maya-Mendoza, A., García-Benítez, F., Bartkova, J., Barroso, S.I., Gaillard, H., Martini, E., Restuccia, U., Ramirez-Otero, M.A., Garre, M., Verga, E., Andújar-Sánchez, M., Maynard, S., Hodny, Z., Costanzo, V., Kumar, A., Bachi, A., Aguilera, A., Bartek, J., Foiani, M., 2021. The human nucleoporin Tpr protects cells from RNA-mediated replication stress. *Nat Commun* 12, 3937. <https://doi.org/10.1038/s41467-021-24224-3>
- Kotnis, A., Du, L., Liu, C., Popov, S.W., Pan-Hammarström, Q., 2009. Non-homologous end joining in class switch recombination: the beginning of the end. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 364, 653–665. <https://doi.org/10.1098/rstb.2008.0196>
- Kramarz, K., Schirmeisen, K., Boucherit, V., Ait Saada, A., Lovo, C., Palancade, B., Freudenreich, C., Lambert, S.A.E., 2020. The nuclear pore primes recombination-dependent DNA synthesis at arrested forks by promoting SUMO removal. *Nat Commun* 11, 5643. <https://doi.org/10.1038/s41467-020-19516-z>
- Kuhn, T.M., Pascual-Garcia, P., Gozalo, A., Little, S.C., Capelson, M., 2019. Chromatin targeting of nuclear pore proteins induces chromatin decondensation. *Journal of Cell Biology* 218, 2945–2961. <https://doi.org/10.1083/jcb.201807139>
- Kumar, A., Fournier, L.-A., Stirling, P.C., 2022. Integrative analysis and prediction of human R-loop binding proteins. G3 (Bethesda) *jkac142*. <https://doi.org/10.1093/g3journal/jkac142>
- Kumar, A., Sharma, P., Gomar-Alba, M., Shcheprova, Z., Daulny, A., Sanmartín, T., Matucci, I., Funaya, C., Beato, M., Mendoza, M., 2018. Daughter-cell-specific modulation of nuclear pore complexes controls cell cycle entry during asymmetric division. *Nat Cell Biol* 20, 432–442. <https://doi.org/10.1038/s41556-018-0056-9>
- Kumar, C., Batra, S., Griffith, J.D., Remus, D., 2021. The interplay of RNA:DNA hybrid structure and G-quadruplexes determines the outcome of R-loop-replisome collisions. *Elife* 10, e72286. <https://doi.org/10.7554/eLife.72286>
- Kutay, U., Jühlen, R., Antonin, W., 2021. Mitotic disassembly and reassembly of nuclear pore complexes. *Trends in Cell Biology* 31, 1019–1033. <https://doi.org/10.1016/j.tcb.2021.06.011>
- Lafuente-Barquero, J., Luke-Glaser, S., Graf, M., Silva, S., Gómez-González, B., Lockhart, A., Lisby, M., Aguilera, A., Luke, B., 2017. The Smc5/6 complex regulates the yeast Mph1 helicase at RNA-DNA hybrid-mediated DNA damage. *PLoS Genet.* 13, e1007136. <https://doi.org/10.1371/journal.pgen.1007136>
- Lambert, S., Watson, A., Sheedy, D.M., Martin, B., Carr, A.M., 2005. Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier. *Cell* 121, 689–702. <https://doi.org/10.1016/j.cell.2005.03.022>
- Lamm, N., Read, M.N., Nobis, M., Van Ly, D., Page, S.G., Masamsetti, V.P., Timpson, P., Biro, M., Cesare, A.J., 2020. Nuclear F-actin counteracts nuclear deformation and promotes fork repair during replication stress. *Nat Cell Biol* 22, 1460–1470. <https://doi.org/10.1038/s41556-020-00605-6>

- Lang, K.S., Hall, A.N., Merrikh, C.N., Ragheb, M., Tabakh, H., Pollock, A.J., Woodward, J.J., Dreifus, J.E., Merrikh, H., 2017. Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. *Cell* 170, 787-799.e18. <https://doi.org/10.1016/j.cell.2017.07.044>
- Lautier, O., Penzo, A., Rouvière, J.O., Chevreux, G., Collet, L., Loïdice, I., Taddei, A., Devaux, F., Collart, M.A., Palancade, B., 2021. Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis. *Mol Cell* 81, 2417-2427.e5. <https://doi.org/10.1016/j.molcel.2021.03.030>
- Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., Krause, H.M., 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131, 174–187. <https://doi.org/10.1016/j.cell.2007.08.003>
- Lee, C.-W., Wilfling, F., Ronchi, P., Allegretti, M., Mosalaganti, S., Jentsch, S., Beck, M., Pfander, B., 2020. Selective autophagy degrades nuclear pore complexes. *Nat Cell Biol* 22, 159–166. <https://doi.org/10.1038/s41556-019-0459-2>
- Lee, S.E., Moore, J.K., Holmes, A., Umezu, K., Kolodner, R.D., Haber, J.E., 1998. *Saccharomyces* Ku70, mre11/rad50 and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* 94, 399–409. [https://doi.org/10.1016/s0092-8674\(00\)81482-8](https://doi.org/10.1016/s0092-8674(00)81482-8)
- Lee, S.H., Sterling, H., Burlingame, A., McCormick, F., 2008. Tpr directly binds to Mad1 and Mad2 and is important for the Mad1-Mad2-mediated mitotic spindle checkpoint. *Genes Dev* 22, 2926–2931. <https://doi.org/10.1101/gad.1677208>
- Legros, P., Malapert, A., Niinuma, S., Bernard, P., Vanoosthuysse, V., 2014. RNA processing factors Swd2.2 and Sen1 antagonize RNA Pol III-dependent transcription and the localization of condensin at Pol III genes. *PLoS Genet.* 10, e1004794. <https://doi.org/10.1371/journal.pgen.1004794>
- Lemaître, C., Fischer, B., Kalousi, A., Hoffbeck, A.-S., Guirouilh-Barbat, J., Shahar, O.D., Genet, D., Goldberg, M., Bertrand, P., Lopez, B., Brino, L., Soutoglou, E., 2012. The nucleoporin 153, a novel factor in double-strand break repair and DNA damage response. *Oncogene* 31, 4803–4809. <https://doi.org/10.1038/onc.2011.638>
- Li, B., Kohler, J.J., 2014. Glycosylation of the nuclear pore. *Traffic* 15, 347–361. <https://doi.org/10.1111/tra.12150>
- Li, S.-J., Hochstrasser, M., 2003. The Ulp1 SUMO isopeptidase: distinct domains required for viability, nuclear envelope localization, and substrate specificity. *J Cell Biol* 160, 1069–1081. <https://doi.org/10.1083/jcb.200212052>
- Li, X., Manley, J.L., 2005. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* 122, 365–378. <https://doi.org/10.1016/j.cell.2005.06.008>
- Liang, Z., Liang, F., Teng, Y., Chen, X., Liu, J., Longerich, S., Rao, T., Green, A.M., Collins, N.B., Xiong, Y., Lan, L., Sung, P., Kupfer, G.M., 2019. Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination. *Cell Rep* 26, 564-572.e5. <https://doi.org/10.1016/j.celrep.2018.12.084>
- Light, W.H., Brickner, D.G., Brand, V.R., Brickner, J.H., 2010. Interaction of a DNA zip code with the nuclear pore complex promotes H2A.Z incorporation and INO1 transcriptional memory. *Mol. Cell* 40, 112–125. <https://doi.org/10.1016/j.molcel.2010.09.007>
- Light, W.H., Brickner, J.H., 2013. Nuclear pore proteins regulate chromatin structure and transcriptional memory by a conserved mechanism. *Nucleus* 4, 357–360. <https://doi.org/10.4161/nucl.26209>
- Lim, Y.W., Sanz, L.A., Xu, X., Hartono, S.R., Chédin, F., 2015. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi–Goutières syndrome. *eLife* 4, e08007. <https://doi.org/10.7554/eLife.08007>
- Lisby, M., Mortensen, U.H., Rothstein, R., 2003. Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. *Nat Cell Biol* 5, 572–577. <https://doi.org/10.1038/ncb997>
- Lockhart, A., Pires, V.B., Bento, F., Kellner, V., Luke-Glaser, S., Yakoub, G., Ulrich, H.D., Luke, B., 2019. RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids. *Cell Rep* 29, 2890-2900.e5. <https://doi.org/10.1016/j.celrep.2019.10.108>
- Loeillet, S., Palancade, B., Cartron, M., Thierry, A., Richard, G.-F., Dujon, B., Doye, V., Nicolas, A., 2005. Genetic network interactions among replication, repair and nuclear pore deficiencies in yeast. *DNA Repair (Amst.)* 4, 459–468. <https://doi.org/10.1016/j.dnarep.2004.11.010>
- Long, R.M., Singer, R.H., Meng, X., Gonzalez, I., Nasmyth, K., Jansen, R.P., 1997. Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* 277, 383–387. <https://doi.org/10.1126/science.277.5324.383>

BIBLIOGRAPHY

- Lottersberger, F., Karssemeijer, R.A., Dimitrova, N., de Lange, T., 2015. 53BP1 and the LINC Complex Promote Microtubule-Dependent DSB Mobility and DNA Repair. *Cell* 163, 880–893. <https://doi.org/10.1016/j.cell.2015.09.057>
- Lu, B., Dong, L., Yi, D., Zhang, M., Zhu, C., Li, X., Yi, C., 2020. Transposase-assisted tagmentation of RNA/DNA hybrid duplexes. *eLife* 9, e54919. <https://doi.org/10.7554/eLife.54919>
- Lu, J., Wu, T., Zhang, B., Liu, S., Song, W., Qiao, J., Ruan, H., 2021. Types of nuclear localization signals and mechanisms of protein import into the nucleus. *Cell Communication and Signaling* 19, 60. <https://doi.org/10.1186/s12964-021-00741-y>
- Luna, R., Jimeno, S., Marín, M., Huertas, P., García-Rubio, M., Aguilera, A., 2005. Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. *Mol. Cell* 18, 711–722. <https://doi.org/10.1016/j.molcel.2005.05.001>
- Lupu, F., Alves, A., Anderson, K., Doye, V., Lacy, E., 2008. Nuclear Pore Composition Regulates Neural Stem/Progenitor Cell Differentiation in the Mouse Embryo. *Developmental Cell* 14, 831–842. <https://doi.org/10.1016/j.devcel.2008.03.011>
- Luthra, R., Kerr, S.C., Harreman, M.T., Apponi, L.H., Fasken, M.B., Ramineni, S., Chaurasia, S., Valentini, S.R., Corbett, A.H., 2007. Actively Transcribed GAL Genes Can Be Physically Linked to the Nuclear Pore by the SAGA Chromatin Modifying Complex. *J. Biol. Chem.* 282, 3042–3049. <https://doi.org/10.1074/jbc.M608741200>
- Mackay, D.R., Makise, M., Ullman, K.S., 2010. Defects in nuclear pore assembly lead to activation of an Aurora B-mediated abscission checkpoint. *J Cell Biol* 191, 923–931. <https://doi.org/10.1083/jcb.201007124>
- Mackay, R.P., Xu, Q., Weinberger, P.M., 2020. R-Loop Physiology and Pathology: A Brief Review. *DNA Cell Biol.* <https://doi.org/10.1089/dna.2020.5906>
- Maestroni, L., Audry, J., Luciano, P., Coulon, S., Géli, V., Corda, Y., 2020. RPA and Pif1 cooperate to remove G-rich structures at both leading and lagging strand. *Cell Stress* 4, 48–63. <https://doi.org/10.15698/cst2020.03.214>
- Mahadevan, K., Zhang, H., Akef, A., Cui, X.A., Gueroussov, S., Cenik, C., Roth, F.P., Palazzo, A.F., 2013. RanBP2/Nup358 Potentiates the Translation of a Subset of mRNAs Encoding Secretory Proteins. *PLOS Biology* 11, e1001545. <https://doi.org/10.1371/journal.pbio.1001545>
- Makharashvili, N., Arora, S., Yin, Y., Fu, Q., Wen, X., Lee, J.-H., Kao, C.-H., Leung, J.W., Miller, K.M., Paull, T.T., 2018. Sae2/CtIP prevents R-loop accumulation in eukaryotic cells. *eLife* 7, e42733. <https://doi.org/10.7554/eLife.42733>
- Makio, T., Lapetina, D.L., Wozniak, R.W., 2013. Inheritance of yeast nuclear pore complexes requires the Nsp1p subcomplex. *Journal of Cell Biology* 203, 187–196. <https://doi.org/10.1083/jcb.201304047>
- Malig, M., Hartono, S.R., Giafaglione, J.M., Sanz, L.A., Chedin, F., 2020. Ultra-deep Coverage Single-molecule R-loop Footprinting Reveals Principles of R-loop Formation. *J. Mol. Biol.* <https://doi.org/10.1016/j.jmb.2020.02.014>
- Manfrini, N., Trovesi, C., Wery, M., Martina, M., Cesena, D., Describes, M., Morillon, A., d’Adda di Fagagna, F., Longhese, M.P., 2015. RNA-processing proteins regulate Mec1/ATR activation by promoting generation of RPA-coated ssDNA. *EMBO Rep* 16, 221–231. <https://doi.org/10.15252/embr.201439458>
- Manzo, S.G., Hartono, S.R., Sanz, L.A., Marinello, J., De Biasi, S., Cossarizza, A., Capranico, G., Chedin, F., 2018. DNA Topoisomerase I differentially modulates R-loops across the human genome. *Genome Biol.* 19, 100. <https://doi.org/10.1186/s13059-018-1478-1>
- Marabitti, V., Lillo, G., Malacaria, E., Palermo, V., Sanchez, M., Pichierri, P., Franchitto, A., 2019. ATM pathway activation limits R-loop-associated genomic instability in Werner syndrome cells. *Nucleic Acids Research* 47, 3485–3502. <https://doi.org/10.1093/nar/gkz025>
- Marcand, S., Pardo, B., Gratias, A., Cahun, S., Callebaut, I., 2008. Multiple pathways inhibit NHEJ at telomeres. *Genes Dev* 22, 1153–1158. <https://doi.org/10.1101/gad.455108>
- Marnef, A., Finoux, A.-L., Arnould, C., Guillou, E., Daburon, V., Rocher, V., Mangeat, T., Mangeot, P.E., Ricci, E.P., Legube, G., 2019. A cohesin/HUSH- and LINC-dependent pathway controls ribosomal DNA double-strand break repair. *Genes Dev* 33, 1175–1190. <https://doi.org/10.1101/gad.324012.119>
- Marnef, A., Legube, G., 2021. R-loops as Janus-faced modulators of DNA repair. *Nat Cell Biol* 23, 305–313. <https://doi.org/10.1038/s41556-021-00663-4>

- Maruyama, E.O., Hori, T., Tanabe, H., Kitamura, H., Matsuda, R., Tone, S., Hozak, P., Habermann, F.A., von Hase, J., Cremer, C., Fukagawa, T., Harata, M., 2012. The actin family member Arp6 and the histone variant H2A.Z are required for spatial positioning of chromatin in chicken cell nuclei. *J Cell Sci* 125, 3739–3743. <https://doi.org/10.1242/jcs.103903>
- Matthews, A.J., Zheng, S., DiMenna, L.J., Chaudhuri, J., 2014. Chapter One - Regulation of Immunoglobulin Class-Switch Recombination: Choreography of Noncoding Transcription, Targeted DNA Deamination, and Long-Range DNA Repair, in: Alt, F.W. (Ed.), *Advances in Immunology*. Academic Press, pp. 1–57. <https://doi.org/10.1016/B978-0-12-800267-4.00001-8>
- Maul, G.G., 1971. On the octagonality of the nuclear pore complex. *J Cell Biol* 51, 558–563. <https://doi.org/10.1083/jcb.51.2.558>
- Mazina, O.M., Somarowthu, S., Kadyrova, L.Y., Baranovskiy, A.G., Tahirov, T.H., Kadyrov, F.A., Mazin, A.V., 2020. Replication protein A binds RNA and promotes R-loop formation. *Journal of Biological Chemistry* 295, 14203–14213. <https://doi.org/10.1074/jbc.RA120.013812>
- Meinel, D.M., Burkert-Kautzsch, C., Kieser, A., O’Duibhir, E., Siebert, M., Mayer, A., Cramer, P., Söding, J., Holstege, F.C.P., Strässer, K., 2013. Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II. *PLoS Genet* 9, e1003914. <https://doi.org/10.1371/journal.pgen.1003914>
- Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., Wilm, M., Stunnenberg, H.G., Saumweber, H., Akhtar, A., 2006. Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol Cell* 21, 811–823. <https://doi.org/10.1016/j.molcel.2006.02.007>
- Merrikh, C.N., Merrikh, H., 2018. Gene inversion potentiates bacterial evolvability and virulence. *Nat Commun* 9, 4662. <https://doi.org/10.1038/s41467-018-07110-3>
- Miller, H.E., Montemayor, D., Abdul, J., Vines, A., Levy, S.A., Hartono, S.R., Sharma, K., Frost, B., Chédin, F., Bishop, A.J.R., 2022. Quality-controlled R-loop meta-analysis reveals the characteristics of R-loop consensus regions. *Nucleic Acids Res* gkac537. <https://doi.org/10.1093/nar/gkac537>
- Milman, G., Langridge, R., Chamberlin, M.J., 1967. The structure of a DNA-RNA hybrid. *PNAS* 57, 1804–1810. <https://doi.org/10.1073/pnas.57.6.1804>
- Miné-Hattab, J., Recamier, V., Izeddin, I., Rothstein, R., Darzacq, X., 2017. Multi-scale tracking reveals scale-dependent chromatin dynamics after DNA damage. *Mol Biol Cell* mbc.E17-05-0317. <https://doi.org/10.1091/mbc.E17-05-0317>
- Miné-Hattab, J., Rothstein, R., 2012. Increased chromosome mobility facilitates homology search during recombination. *Nat Cell Biol* 14, 510–517. <https://doi.org/10.1038/ncb2472>
- Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., Proudfoot, N.J., 2011. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell* 41, 21–32. <https://doi.org/10.1016/j.molcel.2010.12.007>
- Moradi-Fard, S., Sarthi, J., Tittel-Elmer, M., Lalonde, M., Cusanelli, E., Chartrand, P., Cobb, J.A., 2016. Smc5/6 Is a Telomere-Associated Complex that Regulates Sir4 Binding and TPE. *PLoS Genet*. 12, e1006268. <https://doi.org/10.1371/journal.pgen.1006268>
- Mosler, T., Conte, F., Longo, G.M.C., Mikicic, I., Kreim, N., Möckel, M.M., Petrosino, G., Flach, J., Barau, J., Luke, B., Roukos, V., Beli, P., 2021. R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nat Commun* 12, 7314. <https://doi.org/10.1038/s41467-021-27530-y>
- Mouaikel, J., Causse, S.Z., Rougemaille, M., Daubenton-Carafa, Y., Blugeon, C., Lemoine, S., Devaux, F., Darzacq, X., Libri, D., 2013. High-frequency promoter firing links THO complex function to heavy chromatin formation. *Cell Rep* 5, 1082–1094. <https://doi.org/10.1016/j.celrep.2013.10.013>
- Moudry, P., Lukas, C., Macurek, L., Neumann, B., Heriche, J.-K., Pepperkok, R., Ellenberg, J., Hodny, Z., Lukas, J., Bartek, J., 2012. Nucleoporin NUP153 guards genome integrity by promoting nuclear import of 53BP1. *Cell Death Differ* 19, 798–807. <https://doi.org/10.1038/cdd.2011.150>
- Mullen, J.R., Brill, S.J., 2008. Activation of the Slx5-Slx8 ubiquitin ligase by poly-small ubiquitin-like modifier conjugates. *J. Biol. Chem.* 283, 19912–19921. <https://doi.org/10.1074/jbc.M802690200>
- Mullen, J.R., Kaliraman, V., Ibrahim, S.S., Brill, S.J., 2001. Requirement for three novel protein complexes in the absence of the Sgs1 DNA helicase in *Saccharomyces cerevisiae*. *Genetics* 157, 103–118.

- Munden, A., Lauren Benton, M., Capra, J.A., Nordman, J., 2022. R-loop mapping and characterization during *Drosophila* embryogenesis reveals developmental plasticity in R-loop signatures. *J Mol Biol* 167645. <https://doi.org/10.1016/j.jmb.2022.167645>
- Murakami, Y., Eki, T., Hurwitz, J., 1992. Studies on the initiation of simian virus 40 replication in vitro: RNA primer synthesis and its elongation. *Proc Natl Acad Sci U S A* 89, 952–956. <https://doi.org/10.1073/pnas.89.3.952>
- Musiálek, M.W., Rybaczek, D., 2021. Hydroxyurea-The Good, the Bad and the Ugly. *Genes (Basel)* 12, 1096. <https://doi.org/10.3390/genes12071096>
- Nadel, J., Athanasiadou, R., Lemetre, C., Wijetunga, N.A., Ó Broin, P., Sato, H., Zhang, Z., Jeddelloh, J., Montagna, C., Golden, A., Seoighe, C., Grealley, J.M., 2015. RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. *Epigenetics Chromatin* 8, 46. <https://doi.org/10.1186/s13072-015-0040-6>
- Nader, G.P. de F., Agüera-Gonzalez, S., Routet, F., Gratia, M., Maurin, M., Cancila, V., Cadart, C., Palamidessi, A., Ramos, R.N., San Roman, M., Gentili, M., Yamada, A., Williart, A., Lodillinsky, C., Lagoutte, E., Villard, C., Viovy, J.-L., Tripodo, C., Galon, J., Scita, G., Manel, N., Chavrier, P., Piel, M., 2021. Compromised nuclear envelope integrity drives TREX1-dependent DNA damage and tumor cell invasion. *Cell* 184, 5230-5246.e22. <https://doi.org/10.1016/j.cell.2021.08.035>
- Nagai, S., Dubrana, K., Tsai-Pflugfelder, M., Davidson, M.B., Roberts, T.M., Brown, G.W., Varela, E., Hediger, F., Gasser, S.M., Krogan, N.J., 2008. Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science* 322, 597–602. <https://doi.org/10.1126/science.1162790>
- Neaves, K.J., Huppert, J.L., Henderson, R.M., Edwardson, J.M., 2009. Direct visualization of G-quadruplexes in DNA using atomic force microscopy. *Nucleic Acids Res* 37, 6269–6275. <https://doi.org/10.1093/nar/gkp679>
- Nguyen, H.D., Yadav, T., Giri, S., Saez, B., Graubert, T.A., Zou, L., 2017. Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Mol. Cell* 65, 832-847.e4. <https://doi.org/10.1016/j.molcel.2017.01.029>
- Niepel, M., Molloy, K.R., Williams, R., Farr, J.C., Meinema, A.C., Vecchietti, N., Cristea, I.M., Chait, B.T., Rout, M.P., Strambio-De-Castillia, C., 2013. The nuclear basket proteins Mlp1p and Mlp2p are part of a dynamic interactome including Esc1p and the proteasome. *Mol Biol Cell* 24, 3920–3938. <https://doi.org/10.1091/mbc.E13-07-0412>
- Niño, C.A., Guet, D., Gay, A., Brutus, S., Jourquin, F., Mendiratta, S., Salamero, J., Géli, V., Dargemont, C., 2016. Posttranslational marks control architectural and functional plasticity of the nuclear pore complex basket. *Journal of Cell Biology* 212, 167–180. <https://doi.org/10.1083/jcb.201506130>
- Nowotny, M., Cerritelli, S.M., Ghirlando, R., Gaidamakov, S.A., Crouch, R.J., Yang, W., 2008. Specific recognition of RNA/DNA hybrid and enhancement of human RNase H1 activity by HBD. *EMBO J.* 27, 1172–1181. <https://doi.org/10.1038/emboj.2008.44>
- O’Connell, K., Jinks-Robertson, S., Petes, T.D., 2015. Elevated Genome-Wide Instability in Yeast Mutants Lacking RNase H Activity. *Genetics* 201, 963–975. <https://doi.org/10.1534/genetics.115.182725>
- Okamura, M., Inose, H., Masuda, S., 2015. RNA Export through the NPC in Eukaryotes. *Genes* 6, 124–149. <https://doi.org/10.3390/genes6010124>
- Onischenko, E., Noor, E., Fischer, J.S., Gillet, L., Wojtynek, M., Vallotton, P., Weis, K., 2020. Maturation Kinetics of a Multiprotein Complex Revealed by Metabolic Labeling. *Cell* 183, 1785-1800.e26. <https://doi.org/10.1016/j.cell.2020.11.001>
- Onischenko, E., Tang, J.H., Andersen, K.R., Knockenhauer, K.E., Vallotton, P., Derrer, C.P., Kralt, A., Mugler, C.F., Chan, L.Y., Schwartz, T.U., Weis, K., 2017. Natively Unfolded FG Repeats Stabilize the Structure of the Nuclear Pore Complex. *Cell* 171, 904-917.e19. <https://doi.org/10.1016/j.cell.2017.09.033>
- Oshidari, R., Strecker, J., Chung, D.K.C., Abraham, K.J., Chan, J.N.Y., Damaren, C.J., Mekhail, K., 2018. Nuclear microtubule filaments mediate non-linear directional motion of chromatin and promote DNA repair. *Nat Commun* 9, 2567. <https://doi.org/10.1038/s41467-018-05009-7>
- O’Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., Proudfoot, N.J., 2004. Gene loops juxtapose promoters and terminators in yeast. *Nat. Genet.* 36, 1014–1018. <https://doi.org/10.1038/ng1411>
- Oza, P., Jaspersen, S.L., Miele, A., Dekker, J., Peterson, C.L., 2009. Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. *Genes Dev.* 23, 912–927. <https://doi.org/10.1101/gad.1782209>

- Palancade, B., Doye, V., 2008. Sumoylating and desumoylating enzymes at nuclear pores: underpinning their unexpected duties? *Trends Cell Biol.* 18, 174–183. <https://doi.org/10.1016/j.tcb.2008.02.001>
- Palancade, B., Liu, X., Garcia-Rubio, M., Aguilera, A., Zhao, X., Doye, V., 2007. Nucleoporins prevent DNA damage accumulation by modulating Ulp1-dependent sumoylation processes. *Mol. Biol. Cell* 18, 2912–2923. <https://doi.org/10.1091/mbc.e07-02-0123>
- Palancade, B., Rothstein, R., 2021. The Ultimate (Mis)match: When DNA Meets RNA. *Cells* 10, 1433. <https://doi.org/10.3390/cells10061433>
- Palancade, B., Zuccolo, M., Loeillet, S., Nicolas, A., Doye, V., 2005. Pml39, a novel protein of the nuclear periphery required for nuclear retention of improper messenger ribonucleoparticles. *Mol. Biol. Cell* 16, 5258–5268. <https://doi.org/10.1091/mbc.e05-06-0527>
- Panse, V.G., Küster, B., Gerstberger, T., Hurt, E., 2003. Unconventional tethering of Ulp1 to the transport channel of the nuclear pore complex by karyopherins. *Nat Cell Biol* 5, 21–27. <https://doi.org/10.1038/ncb893>
- Parry, G., 2013. Assessing the function of the plant nuclear pore complex and the search for specificity. *J Exp Bot* 64, 833–845. <https://doi.org/10.1093/jxb/ers289>
- Pascual-Garcia, P., Debo, B., Aleman, J.R., Talamas, J.A., Lan, Y., Nguyen, N.H., Won, K.J., Capelson, M., 2017. Metazoan Nuclear Pores Provide a Scaffold for Poised Genes and Mediate Induced Enhancer-Promoter Contacts. *Mol Cell* 66, 63–76.e6. <https://doi.org/10.1016/j.molcel.2017.02.020>
- Patel, P.S., Krishnan, R., Hakem, R., 2022. Emerging roles of DNA topoisomerases in the regulation of R-loops. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* 876–877, 503450. <https://doi.org/10.1016/j.mrgentox.2022.503450>
- Paul, S., Million-Weaver, S., Chattopadhyay, S., Sokurenko, E., Merrikh, H., 2013. Accelerated gene evolution through replication–transcription conflicts. *Nature* 495, 512–515. <https://doi.org/10.1038/nature11989>
- Perego, M.G.L., Taiana, M., Bresolin, N., Comi, G.P., Corti, S., 2019. R-Loops in Motor Neuron Diseases. *Mol Neurobiol* 56, 2579–2589. <https://doi.org/10.1007/s12035-018-1246-y>
- Pérez-Calero, C., Bayona-Feliu, A., Xue, X., Barroso, S.I., Muñoz, S., González-Basallote, V.M., Sung, P., Aguilera, A., 2020. UAP56/DDX39B is a major cotranscriptional RNA-DNA helicase that unwinds harmful R loops genome-wide. *Genes Dev.* <https://doi.org/10.1101/gad.336024.119>
- Pessina, F., Giavazzi, F., Yin, Y., Gioia, U., Vitelli, V., Galbiati, A., Barozzi, S., Garre, M., Oldani, A., Flaus, A., Cerbino, R., Parazzoli, D., Rothenberg, E., d’Adda di Fagagna, F., 2019. Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. *Nat Cell Biol* 21, 1286–1299. <https://doi.org/10.1038/s41556-019-0392-4>
- Petermann, E., Lan, L., Zou, L., 2022. Sources, resolution and physiological relevance of R-loops and RNA-DNA hybrids. *Nat Rev Mol Cell Biol.* <https://doi.org/10.1038/s41580-022-00474-x>
- Petersen-Mahrt, S.K., Harris, R.S., Neuberger, M.S., 2002. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* 418, 99–104. <https://doi.org/10.1038/nature00862>
- Phillips, D.D., Garboczi, D.N., Singh, K., Hu, Z., Leppla, S.H., Leysath, C.E., 2013. The sub-nanomolar binding of DNA-RNA hybrids by the single-chain Fv fragment of antibody S9.6. *J. Mol. Recognit.* 26, 376–381. <https://doi.org/10.1002/jmr.2284>
- Pichler, A., Gast, A., Seeler, J.S., Dejean, A., Melchior, F., 2002. The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell* 108, 109–120. [https://doi.org/10.1016/s0092-8674\(01\)00633-x](https://doi.org/10.1016/s0092-8674(01)00633-x)
- Pinzaru, A.M., Kareh, M., Lamm, N., Lazzarini-Denchi, E., Cesare, A.J., Sfeir, A., 2020. Replication stress conferred by POT1 dysfunction promotes telomere relocalization to the nuclear pore. *Genes Dev* 34, 1619–1636. <https://doi.org/10.1101/gad.337287.120>
- Piruat, J.I., Aguilera, A., 1998. A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *EMBO J.* 17, 4859–4872. <https://doi.org/10.1093/emboj/17.16.4859>
- Prado, F., Piruat, J.I., Aguilera, A., 1997. Recombination between DNA repeats in yeast hpr1delta cells is linked to transcription elongation. *EMBO J.* 16, 2826–2835. <https://doi.org/10.1093/emboj/16.10.2826>
- Promonet, A., Padioleau, I., Liu, Y., Sanz, L., Biernacka, A., Schmitz, A.-L., Skrzypczak, M., Sarrazin, A., Mettling, C., Rowicka, M., Ginalska, K., Chedin, F., Chen, C.-L., Lin, Y.-L., Pasero, P., 2020. Topoisomerase 1 prevents

- replication stress at R-loop-enriched transcription termination sites. *Nat Commun* 11, 3940. <https://doi.org/10.1038/s41467-020-17858-2>
- Psakhye, I., Jentsch, S., 2012. Protein Group Modification and Synergy in the SUMO Pathway as Exemplified in DNA Repair. *Cell* 151, 807–820. <https://doi.org/10.1016/j.cell.2012.10.021>
- Ptak, C., Saik, N.O., Premashankar, A., Lapetina, D.L., Aitchison, J.D., Montpetit, B., Wozniak, R.W., 2021. Phosphorylation-dependent mitotic SUMOylation drives nuclear envelope–chromatin interactions. *Journal of Cell Biology* 220, e202103036. <https://doi.org/10.1083/jcb.202103036>
- Qiao, Q., Wang, L., Meng, F.-L., Hwang, J.K., Alt, F.W., Wu, H., 2017. AID Recognizes Structured DNA for Class Switch Recombination. *Mol. Cell* 67, 361–373.e4. <https://doi.org/10.1016/j.molcel.2017.06.034>
- Qiu, J., Qian, Y., Frank, P., Wintersberger, U., Shen, B., 1999. *Saccharomyces cerevisiae* RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. *Mol Cell Biol* 19, 8361–8371. <https://doi.org/10.1128/MCB.19.12.8361>
- Rabut, G., Doye, V., Ellenberg, J., 2004. Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nat Cell Biol* 6, 1114–1121. <https://doi.org/10.1038/ncb1184>
- Randise-Hinchliff, C., Coukos, R., Sood, V., Sumner, M.C., Zdraljevic, S., Meldi Sholl, L., Garvey Brickner, D., Ahmed, S., Watchmaker, L., Brickner, J.H., 2016. Strategies to regulate transcription factor-mediated gene positioning and interchromosomal clustering at the nuclear periphery. *J. Cell Biol.* 212, 633–646. <https://doi.org/10.1083/jcb.201508068>
- Ribbeck, K., Görlich, D., 2001. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J* 20, 1320–1330. <https://doi.org/10.1093/emboj/20.6.1320>
- Ribeiro de Almeida, C., Dhir, S., Dhir, A., Moghaddam, A.E., Sattentau, Q., Meinhart, A., Proudfoot, N.J., 2018. RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol. Cell* 70, 650–662.e8. <https://doi.org/10.1016/j.molcel.2018.04.001>
- Richard, P., Manley, J.L., 2017. R Loops and Links to Human Disease. *J. Mol. Biol.* 429, 3168–3180. <https://doi.org/10.1016/j.jmb.2016.08.031>
- Roberts, R.W., Crothers, D.M., 1992. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science* 258, 1463–1466. <https://doi.org/10.1126/science.1279808>
- Robinett, C.C., Straight, A., Li, G., Wilhelm, C., Sudlow, G., Murray, A., Belmont, A.S., 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *Journal of Cell Biology* 135, 1685–1700. <https://doi.org/10.1083/jcb.135.6.1685>
- Rocha, E.P.C., Danchin, A., 2003. Gene essentiality determines chromosome organisation in bacteria. *Nucleic Acids Res* 31, 6570–6577. <https://doi.org/10.1093/nar/gkg859>
- Rodríguez-Navarro, S., Fischer, T., Luo, M.-J., Antúnez, O., Brettschneider, S., Lechner, J., Pérez-Ortín, J.E., Reed, R., Hurt, E., 2004. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* 116, 75–86. [https://doi.org/10.1016/s0092-8674\(03\)01025-0](https://doi.org/10.1016/s0092-8674(03)01025-0)
- Rougemaille, M., Dieppois, G., Kisseleva-Romanova, E., Gudipati, R.K., Lemoine, S., Blugeon, C., Boulay, J., Jensen, T.H., Stutz, F., Devaux, F., Libri, D., 2008a. THO/Sub2p Functions to Coordinate 3'-End Processing with Gene-Nuclear Pore Association. *Cell* 135, 308–321. <https://doi.org/10.1016/j.cell.2008.08.005>
- Rougemaille, M., Villa, T., Gudipati, R.K., Libri, D., 2008b. mRNA journey to the cytoplasm: attire required. *Biol Cell* 100, 327–342. <https://doi.org/10.1042/BC20070143>
- Rout, M.P., Aitchison, J.D., Suprpto, A., Hjertaas, K., Zhao, Y., Chait, B.T., 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol* 148, 635–651. <https://doi.org/10.1083/jcb.148.4.635>
- Rout, M.P., Blobel, G., 1993. Isolation of the yeast nuclear pore complex. *J Cell Biol* 123, 771–783. <https://doi.org/10.1083/jcb.123.4.771>
- Rouvière, J.O., Bulfoni, M., Tuck, A., Cosson, B., Devaux, F., Palancade, B., 2018. A SUMO-dependent feedback loop senses and controls the biogenesis of nuclear pore subunits. *Nat Commun* 9, 1665. <https://doi.org/10.1038/s41467-018-03673-3>
- Roy, D., Yu, K., Lieber, M.R., 2008. Mechanism of R-Loop Formation at Immunoglobulin Class Switch Sequences. *Molecular and Cellular Biology* 28, 50–60. <https://doi.org/10.1128/MCB.01251-07>

- Roy, D., Zhang, Z., Lu, Z., Hsieh, C.-L., Lieber, M.R., 2010. Competition between the RNA Transcript and the Nontemplate DNA Strand during R-Loop Formation In Vitro: a Nick Can Serve as a Strong R-Loop Initiation Site. *Molecular and Cellular Biology* 30, 146–159. <https://doi.org/10.1128/MCB.00897-09>
- Ruiz, J.F., Gómez-González, B., Aguilera, A., 2011. AID induces double-strand breaks at immunoglobulin switch regions and c-MYC causing chromosomal translocations in yeast THO mutants. *PLoS Genet.* 7, e1002009. <https://doi.org/10.1371/journal.pgen.1002009>
- Rydberg, B., Game, J., 2002. Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc Natl Acad Sci U S A* 99, 16654–16659. <https://doi.org/10.1073/pnas.262591699>
- Ryu, T., Bonner, M.R., Chiolo, I., 2016. Cervantes and Quijote protect heterochromatin from aberrant recombination and lead the way to the nuclear periphery. *Nucleus* 7, 485–497. <https://doi.org/10.1080/19491034.2016.1239683>
- Ryu, T., Spatola, B., Delabaere, L., Bowlin, K., Hopp, H., Kunitake, R., Karpen, G.H., Chiolo, I., 2015. Heterochromatic breaks move to the nuclear periphery to continue recombinational repair. *Nat. Cell Biol.* 17, 1401–1411. <https://doi.org/10.1038/ncb3258>
- Saik, N.O., Park, N., Ptak, C., Adames, N., Aitchison, J.D., Wozniak, R.W., 2020. Recruitment of an Activated Gene to the Yeast Nuclear Pore Complex Requires Sumoylation. *Front Genet* 11, 174. <https://doi.org/10.3389/fgene.2020.00174>
- Saitoh, H., Pizzi, M.D., Wang, J., 2002. Perturbation of SUMOylation enzyme Ubc9 by distinct domain within nucleoporin RanBP2/Nup358. *J Biol Chem* 277, 4755–4763. <https://doi.org/10.1074/jbc.M104453200>
- Salas-Armenteros, I., Pérez-Calero, C., Bayona-Feliu, A., Tumini, E., Luna, R., Aguilera, A., 2017. Human THO-Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability. *EMBO J.* 36, 3532–3547. <https://doi.org/10.15252/embj.201797208>
- Saldi, T., Cortazar, M.A., Sheridan, R.M., Bentley, D.L., 2016. Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *J Mol Biol* 428, 2623–2635. <https://doi.org/10.1016/j.jmb.2016.04.017>
- Salina, D., Enarson, P., Rattner, J.B., Burke, B., 2003. Nup358 integrates nuclear envelope breakdown with kinetochore assembly. *Journal of Cell Biology* 162, 991–1001. <https://doi.org/10.1083/jcb.200304080>
- San Martín-Alonso, M., Soler-Oliva, M.E., García-Rubio, M., García-Muse, T., Aguilera, A., 2021. Harmful R-loops are prevented via different cell cycle-specific mechanisms. *Nat Commun* 12, 4451. <https://doi.org/10.1038/s41467-021-24737-x>
- Santos-Pereira, J.M., Aguilera, A., 2015. R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* 16, 583–597. <https://doi.org/10.1038/nrg3961>
- Sanz, L.A., Hartono, S.R., Lim, Y.W., Steyaert, S., Rajpurkar, A., Ginno, P.A., Xu, X., Chédin, F., 2016. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Mol. Cell* 63, 167–178. <https://doi.org/10.1016/j.molcel.2016.05.032>
- Schmid, M., Arib, G., Laemmli, C., Nishikawa, J., Durussel, T., Laemmli, U.K., 2006. Nup-PI: the nucleopore-promoter interaction of genes in yeast. *Mol. Cell* 21, 379–391. <https://doi.org/10.1016/j.molcel.2005.12.012>
- Schneider, M., Hellerschmied, D., Schubert, T., Amlacher, S., Vinayachandran, V., Reja, R., Pugh, B.F., Clausen, T., Köhler, A., 2015. The Nuclear Pore-Associated TREX-2 Complex Employs Mediator to Regulate Gene Expression. *Cell* 162, 1016–1028. <https://doi.org/10.1016/j.cell.2015.07.059>
- Scholz, B.A., Sumida, N., de Lima, C.D.M., Chachoua, I., Martino, M., Tzelepis, I., Nikoshkov, A., Zhao, H., Mehmood, R., Sifakis, E.G., Bhartiya, D., Göndör, A., Ohlsson, R., 2019. WNT signaling and AHCTF1 promote oncogenic MYC expression through super-enhancer-mediated gene gating. *Nat Genet* 51, 1723–1731. <https://doi.org/10.1038/s41588-019-0535-3>
- Schuller, A.P., Wojtynek, M., Mankus, D., Tatli, M., Kronenberg-Tenga, R., Regmi, S.G., Dip, P.V., Lytton-Jean, A.K.R., Brignole, E.J., Dasso, M., Weis, K., Medalia, O., Schwartz, T.U., 2021. The cellular environment shapes the nuclear pore complex architecture. *Nature* 598, 667–671. <https://doi.org/10.1038/s41586-021-03985-3>
- Schwab, R.A., Nieminuszczy, J., Shah, F., Langton, J., Lopez Martinez, D., Liang, C.-C., Cohn, M.A., Gibbons, R.J., Deans, A.J., Niedzwiedz, W., 2015. The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Molecular Cell* 60, 351–361. <https://doi.org/10.1016/j.molcel.2015.09.012>

- Scott, R.J., Lusk, C.P., Dilworth, D.J., Aitchison, J.D., Wozniak, R.W., 2005. Interactions between Mad1p and the nuclear transport machinery in the yeast *Saccharomyces cerevisiae*. *Mol Biol Cell* 16, 4362–4374. <https://doi.org/10.1091/mbc.e05-01-0011>
- Seidel, M., Becker, A., Pereira, F., Landry, J.J.M., de Azevedo, N.T.D., Fusco, C.M., Kaindl, E., Romanov, N., Baumbach, J., Langer, J.D., Schuman, E.M., Patil, K.R., Hummer, G., Benes, V., Beck, M., 2022. Co-translational assembly orchestrates competing biogenesis pathways. *Nat Commun* 13, 1224. <https://doi.org/10.1038/s41467-022-28878-5>
- Sessa, G., Gómez-González, B., Silva, S., Pérez-Calero, C., Beaupere, R., Barroso, S., Martineau, S., Martin, C., Ehlén, Å., Martínez, J.S., Lombard, B., Loew, D., Vagner, S., Aguilera, A., Carreira, A., 2021. BRCA2 promotes R-loop resolution by DDX5 helicase at DNA breaks to facilitate their repair by homologous recombination. *EMBO J* e106018. <https://doi.org/10.15252/embj.2020106018>
- Shafiq, S., Chen, C., Yang, J., Cheng, L., Ma, F., Widemann, E., Sun, Q., 2017. DNA Topoisomerase 1 Prevents R-loop Accumulation to Modulate Auxin-Regulated Root Development in Rice. *Mol Plant* 10, 821–833. <https://doi.org/10.1016/j.molp.2017.04.001>
- Sharma, S., Anand, R., Zhang, X., Francia, S., Michelini, F., Galbiati, A., Williams, H., Ronato, D.A., Masson, J.-Y., Rothenberg, E., Cejka, P., d'Adda di Fagagna, F., 2021. MRE11-RAD50-NBS1 Complex Is Sufficient to Promote Transcription by RNA Polymerase II at Double-Strand Breaks by Melting DNA Ends. *Cell Rep* 34, 108565. <https://doi.org/10.1016/j.celrep.2020.108565>
- Shcheprova, Z., Baldi, S., Frei, S.B., Gonnet, G., Barral, Y., 2008. A mechanism for asymmetric segregation of age during yeast budding. *Nature* 454, 728–734. <https://doi.org/10.1038/nature07212>
- Shen, Q., Wang, Y.E., Truong, M., Mahadevan, K., Wu, J.J., Zhang, H., Li, J., Smith, H.W., Smibert, C.A., Palazzo, A.F., 2021. RanBP2/Nup358 enhances miRNA activity by sumoylating Argonautes. *PLoS Genet* 17, e1009378. <https://doi.org/10.1371/journal.pgen.1009378>
- Shen, W., Sun, H., De Hoyos, C.L., Bailey, J.K., Liang, X.-H., Crooke, S.T., 2017. Dynamic nucleoplasmic and nucleolar localization of mammalian RNase H1 in response to RNAP I transcriptional R-loops. *Nucleic Acids Res* 45, 10672–10692. <https://doi.org/10.1093/nar/gkx710>
- Sikorski, T.W., Ficarro, S.B., Holik, J., Kim, T., Rando, O.J., Marto, J.A., Buratowski, S., 2011. Sub1 and RPA associate with RNA polymerase II at different stages of transcription. *Mol. Cell* 44, 397–409. <https://doi.org/10.1016/j.molcel.2011.09.013>
- Skourti-Stathaki, K., Proudfoot, N.J., Gromak, N., 2011. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* 42, 794–805. <https://doi.org/10.1016/j.molcel.2011.04.026>
- Skourti-Stathaki, K., Torlai Triglia, E., Warburton, M., Voigt, P., Bird, A., Pombo, A., 2019. R-Loops Enhance Polycomb Repression at a Subset of Developmental Regulator Genes. *Mol Cell* 73, 930–945.e4. <https://doi.org/10.1016/j.molcel.2018.12.016>
- Slater, M.L., 1973. Effect of reversible inhibition of deoxyribonucleic acid synthesis on the yeast cell cycle. *J Bacteriol* 113, 263–270. <https://doi.org/10.1128/jb.113.1.263-270.1973>
- Smolka, J.A., Sanz, L.A., Hartono, S.R., Chédin, F., 2021. Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA:DNA hybrids. *J Cell Biol* 220, e202004079. <https://doi.org/10.1083/jcb.202004079>
- Smythe, C., Jenkins, H.E., Hutchison, C.J., 2000. Incorporation of the nuclear pore basket protein nup153 into nuclear pore structures is dependent upon lamina assembly: evidence from cell-free extracts of *Xenopus* eggs. *EMBO J* 19, 3918–3931. <https://doi.org/10.1093/emboj/19.15.3918>
- Sollier, J., Stork, C.T., García-Rubio, M.L., Paulsen, R.D., Aguilera, A., Cimprich, K.A., 2014. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol. Cell* 56, 777–785. <https://doi.org/10.1016/j.molcel.2014.10.020>
- Song, C., Hotz-Wagenblatt, A., Voit, R., Grummt, I., 2017. SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev* 31, 1370–1381. <https://doi.org/10.1101/gad.300624.117>
- Stewart, M., 2010. Nuclear export of mRNA. *Trends Biochem Sci* 35, 609–617. <https://doi.org/10.1016/j.tibs.2010.07.001>

- Stirling, P.C., Chan, Y.A., Minaker, S.W., Aristizabal, M.J., Barrett, I., Sipahimalani, P., Kobor, M.S., Hieter, P., 2012. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev.* 26, 163–175. <https://doi.org/10.1101/gad.179721.111>
- Strässer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondón, A.G., Aguilera, A., Struhl, K., Reed, R., Hurt, E., 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* 417, 304–308. <https://doi.org/10.1038/nature746>
- Strawn, L.A., Shen, T., Shulga, N., Goldfarb, D.S., Wentz, S.R., 2004. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat Cell Biol* 6, 197–206. <https://doi.org/10.1038/ncb1097>
- Stuckey, R., García-Rodríguez, N., Aguilera, A., Wellinger, R.E., 2015. Role for RNA:DNA hybrids in origin-independent replication priming in a eukaryotic system. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5779–5784. <https://doi.org/10.1073/pnas.1501769112>
- Su, X.A., Dion, V., Gasser, S.M., Freudenreich, C.H., 2015. Regulation of recombination at yeast nuclear pores controls repair and triplet repeat stability. *Genes Dev.* 29, 1006–1017. <https://doi.org/10.1101/gad.256404.114>
- Su, X.A., Freudenreich, C.H., 2017. Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 114, E8392–E8401. <https://doi.org/10.1073/pnas.1711283114>
- Sumner, M.C., Torrisi, S.B., Brickner, D.G., Brickner, J.H., 2021. Random sub-diffusion and capture of genes by the nuclear pore reduces dynamics and coordinates inter-chromosomal movement. *Elife* 10, e66238. <https://doi.org/10.7554/eLife.66238>
- Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N.J., Dean, C., 2013. R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. *Science* 340, 619–621. <https://doi.org/10.1126/science.1234848>
- Swartz, R.K., Rodriguez, E.C., King, M.C., 2014. A role for nuclear envelope-bridging complexes in homology-directed repair. *Mol Biol Cell* 25, 2461–2471. <https://doi.org/10.1091/mbc.E13-10-0569>
- Taddei, A., Van Houwe, G., Hediger, F., Kalck, V., Cubizolles, F., Schober, H., Gasser, S.M., 2006. Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* 441, 774–778. <https://doi.org/10.1038/nature04845>
- Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T., Hara-Nishimura, I., 2010. Identification and characterization of nuclear pore complex components in *Arabidopsis thaliana*. *Plant Cell* 22, 4084–4097. <https://doi.org/10.1105/tpc.110.079947>
- Tan-Wong, S.M., Dhir, S., Proudfoot, N.J., 2019. R-Loops Promote Antisense Transcription across the Mammalian Genome. *Mol. Cell* 76, 600-616.e6. <https://doi.org/10.1016/j.molcel.2019.10.002>
- Tan-Wong, S.M., Wijayatilake, H.D., Proudfoot, N.J., 2009. Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes Dev.* 23, 2610–2624. <https://doi.org/10.1101/gad.1823209>
- Teixeira, M.T., Fabre, E., Dujon, B., 1999. Self-catalyzed cleavage of the yeast nucleoporin Nup145p precursor. *J Biol Chem* 274, 32439–32444. <https://doi.org/10.1074/jbc.274.45.32439>
- Teixeira, M.T., Siniosoglou, S., Podtelejnikov, S., Bénichou, J.C., Mann, M., Dujon, B., Hurt, E., Fabre, E., 1997. Two functionally distinct domains generated by in vivo cleavage of Nup145p: a novel biogenesis pathway for nucleoporins. *EMBO J* 16, 5086–5097. <https://doi.org/10.1093/emboj/16.16.5086>
- Teloni, F., Michelena, J., Lezaja, A., Kilic, S., Ambrosi, C., Menon, S., Dobrovolna, J., Imhof, R., Janscak, P., Baubec, T., Altmeyer, M., 2019. Efficient Pre-mRNA Cleavage Prevents Replication-Stress-Associated Genome Instability. *Molecular Cell* 73, 670-683.e12. <https://doi.org/10.1016/j.molcel.2018.11.036>
- Teng, Y., Yadav, T., Duan, M., Tan, J., Xiang, Y., Gao, B., Xu, J., Liang, Z., Liu, Y., Nakajima, S., Shi, Y., Levine, A.S., Zou, L., Lan, L., 2018. ROS-induced R loops trigger a transcription-coupled but BRCA1/2-independent homologous recombination pathway through CSB. *Nat Commun* 9, 4115. <https://doi.org/10.1038/s41467-018-06586-3>
- Terlecki-Zaniewicz, S., Humer, T., Eder, T., Schmoellerl, J., Heyes, E., Manhart, G., Kuchynka, N., Parapatics, K., Liberante, F.G., Müller, A.C., Tomazou, E.M., Grebien, F., 2021. Biomolecular condensation of NUP98 fusion proteins drives leukemogenic gene expression. *Nat Struct Mol Biol* 28, 190–201. <https://doi.org/10.1038/s41594-020-00550-w>
- Terry, L.J., Wentz, S.R., 2007. Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. *Journal of Cell Biology* 178, 1121–1132. <https://doi.org/10.1083/jcb.200704174>

- Texari, L., Dieppo, G., Vinciguerra, P., Contreras, M.P., Groner, A., Letourneau, A., Stutz, F., 2013. The Nuclear Pore Regulates GAL1 Gene Transcription by Controlling the Localization of the SUMO Protease Ulp1. *Molecular Cell* 51, 807–818. <https://doi.org/10.1016/j.molcel.2013.08.047>
- Teytelman, L., Thurtle, D.M., Rine, J., van Oudenaarden, A., 2013. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *PNAS* 110, 18602–18607.
- Therizols, P., Fairhead, C., Cabal, G.G., Genovesio, A., Olivo-Marin, J.-C., Dujon, B., Fabre, E., 2006. Telomere tethering at the nuclear periphery is essential for efficient DNA double strand break repair in subtelomeric region. *Journal of Cell Biology* 172, 189–199. <https://doi.org/10.1083/jcb.200505159>
- Thomas, M., White, R.L., Davis, R.W., 1976. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc. Natl. Acad. Sci. U.S.A.* 73, 2294–2298. <https://doi.org/10.1073/pnas.73.7.2294>
- Tian, M., Alt, F.W., 2000. Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision repair nucleases in vitro. *J Biol Chem* 275, 24163–24172. <https://doi.org/10.1074/jbc.M003343200>
- Tingey, M., Li, Y., Yu, W., Young, A., Yang, W., 2022. Spelling out the roles of individual nucleoporins in nuclear export of mRNA. *Nucleus* 13, 170–193. <https://doi.org/10.1080/19491034.2022.2076965>
- Toledo, L.I., Altmeyer, M., Rask, M.-B., Lukas, C., Larsen, D.H., Povlsen, L.K., Bekker-Jensen, S., Mailand, N., Bartek, J., Lukas, J., 2013. ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell* 155, 1088–1103. <https://doi.org/10.1016/j.cell.2013.10.043>
- Tomioka, Y., Kotani, T., Kirisako, H., Oikawa, Y., Kimura, Y., Hirano, H., Ohsumi, Y., Nakatogawa, H., 2020. TORC1 inactivation stimulates autophagy of nucleoporin and nuclear pore complexes. *J. Cell Biol.* 219. <https://doi.org/10.1083/jcb.201910063>
- Torres-Rosell, J., Machin, F., Aragón, L., 2005. Smc5-Smc6 complex preserves nucleolar integrity in *S. cerevisiae*. *Cell Cycle* 4, 868–872. <https://doi.org/10.4161/cc.4.7.1825>
- Torres-Rosell, J., Sunjevaric, I., De Piccoli, G., Sacher, M., Eckert-Boulet, N., Reid, R., Jentsch, S., Rothstein, R., Aragón, L., Lisby, M., 2007. The Smc5-Smc6 complex and SUMO modification of Rad52 regulates recombinational repair at the ribosomal gene locus. *Nat. Cell Biol.* 9, 923–931. <https://doi.org/10.1038/ncb1619>
- Tran, E.J., Zhou, Y., Corbett, A.H., Wentz, S.R., 2007. The DEAD-Box Protein Dbp5 Controls mRNA Export by Triggering Specific RNA:Protein Remodeling Events. *Molecular Cell* 28, 850–859. <https://doi.org/10.1016/j.molcel.2007.09.019>
- Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantescio, V., De Vos, J., Thomas, A., Theillet, C., Pommier, Y., Tazi, J., Coquelle, A., Pasero, P., 2009. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat Cell Biol* 11, 1315–1324. <https://doi.org/10.1038/ncb1984>
- Ulrich, H.D., Davies, A.A., 2009. In vivo detection and characterization of sumoylation targets in *Saccharomyces cerevisiae*. *Methods Mol Biol* 497, 81–103. https://doi.org/10.1007/978-1-59745-566-4_6
- Vaquerez, J.M., Suyama, R., Kind, J., Miura, K., Luscombe, N.M., Akhtar, A., 2010. Nuclear pore proteins nup153 and megator define transcriptionally active regions in the *Drosophila* genome. *PLoS Genet.* 6, e1000846. <https://doi.org/10.1371/journal.pgen.1000846>
- Vazquez, J., Belmont, A.S., Sedat, J.W., 2001. Multiple regimes of constrained chromosome motion are regulated in the interphase *Drosophila* nucleus. *Curr Biol* 11, 1227–1239. [https://doi.org/10.1016/s0960-9822\(01\)00390-6](https://doi.org/10.1016/s0960-9822(01)00390-6)
- Villa, T., Barucco, M., Martin-Niclos, M.-J., Jacquier, A., Libri, D., 2020. Degradation of Non-coding RNAs Promotes Recycling of Termination Factors at Sites of Transcription. *Cell Rep* 32, 107942. <https://doi.org/10.1016/j.celrep.2020.107942>
- Vitale, J., Khan, A., Neuner, A., Schiebel, E., 2022. A perinuclear α -helix with amphipathic features in Brl1 promotes NPC assembly. *Mol Biol Cell* 33, ar35. <https://doi.org/10.1091/mbc.E21-12-0616>
- Vodala, S., Abruzzi, K.C., Rosbash, M., 2008. The nuclear exosome and adenylation regulate posttranscriptional tethering of yeast GAL genes to the nuclear periphery. *Mol Cell* 31, 104–113. <https://doi.org/10.1016/j.molcel.2008.05.015>
- Wahba, L., Costantino, L., Tan, F.J., Zimmer, A., Koshland, D., 2016. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* 30, 1327–1338. <https://doi.org/10.1101/gad.280834.116>

- Wahba, L., Gore, S.K., Koshland, D., 2013. The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability. *eLife* 2, e00505. <https://doi.org/10.7554/eLife.00505>
- Wang, A., Kolhe, J., Gioacchini, N., Baade, I., Briehner, W.M., Peterson, C.L., Freeman, B.C., 2020. Mechanism of Long-Range Chromosome Motion Triggered by Gene Activation. *Dev Cell* 52, 309–320.e5. <https://doi.org/10.1016/j.devcel.2019.12.007>
- Wang, I.X., Grunseich, C., Fox, J., Burdick, J., Zhu, Z., Ravazian, N., Hafner, M., Cheung, V.G., 2018. Human proteins that interact with RNA/DNA hybrids. *Genome Res.* 28, 1405–1414. <https://doi.org/10.1101/gr.237362.118>
- Wang, K., Wang, H., Li, C., Yin, Z., Xiao, R., Li, Q., Xiang, Y., Wang, W., Huang, J., Chen, L., Fang, P., Liang, K., 2021. Genomic profiling of native R loops with a DNA-RNA hybrid recognition sensor. *Sci Adv* 7. <https://doi.org/10.1126/sciadv.abe3516>
- Wang, Z., Jones, G.M., Prelich, G., 2006. Genetic analysis connects SLX5 and SLX8 to the SUMO pathway in *Saccharomyces cerevisiae*. *Genetics* 172, 1499–1509. <https://doi.org/10.1534/genetics.105.052811>
- Webster, B.M., Colombi, P., Jäger, J., Lusk, C.P., 2014. Surveillance of Nuclear Pore Complex Assembly by ESCRT-III/Vps4. *Cell* 159, 388–401. <https://doi.org/10.1016/j.cell.2014.09.012>
- Westover, K.D., Bushnell, D.A., Kornberg, R.D., 2004. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* 303, 1014–1016. <https://doi.org/10.1126/science.1090839>
- Whalen, J.M., Dhingra, N., Wei, L., Zhao, X., Freudenreich, C.H., 2020. Relocation of Collapsed Forks to the Nuclear Pore Complex Depends on Sumoylation of DNA Repair Proteins and Permits Rad51 Association. *Cell Rep* 31, 107635. <https://doi.org/10.1016/j.celrep.2020.107635>
- White, R.L., Hogness, D.S., 1977. R loop mapping of the 18S and 28S sequences in the long and short repeating units of *Drosophila melanogaster* rDNA. *Cell* 10, 177–192. [https://doi.org/10.1016/0092-8674\(77\)90213-6](https://doi.org/10.1016/0092-8674(77)90213-6)
- Wiermer, M., Palma, K., Zhang, Y., Li, X., 2007. Should I stay or should I go? Nucleocytoplasmic trafficking in plant innate immunity. *Cell Microbiol* 9, 1880–1890. <https://doi.org/10.1111/j.1462-5822.2007.00962.x>
- Wimberly, H., Shee, C., Thornton, P.C., Sivaramakrishnan, P., Rosenberg, S.M., Hastings, P.J., 2013. R-loops and nicks initiate DNA breakage and genome instability in non-growing *Escherichia coli*. *Nature Communications* 4. <https://doi.org/10.1038/ncomms3115>
- Wohlschlegel, J.A., Johnson, E.S., Reed, S.I., Yates, J.R., 2004. Global analysis of protein sumoylation in *Saccharomyces cerevisiae*. *J Biol Chem* 279, 45662–45668. <https://doi.org/10.1074/jbc.M409203200>
- Wolf, C., Rapp, A., Berndt, N., Staroske, W., Schuster, M., Dobrick-Mattheuer, M., Kretschmer, S., König, N., Kurth, T., Wiczorek, D., Kast, K., Cardoso, M.C., Günther, C., Lee-Kirsch, M.A., 2016. RPA and Rad51 constitute a cell intrinsic mechanism to protect the cytosol from self DNA. *Nat Commun* 7, 11752. <https://doi.org/10.1038/ncomms11752>
- Wu, T., Nance, J., Chu, F., Fazio, T.G., 2021. Characterization of R-loop-interacting proteins in embryonic stem cells reveals roles in ribosomal RNA processing and gene expression. *Mol Cell Proteomics* 100142. <https://doi.org/10.1016/j.mcpro.2021.100142>
- Wu, X., Han, J., Guo, C., 2022. Function of Nuclear Pore Complexes in Regulation of Plant Defense Signaling. *Int J Mol Sci* 23, 3031. <https://doi.org/10.3390/ijms23063031>
- Wulfridge, P., Sarma, K., 2021. A nuclease- and bisulfite-based strategy captures strand-specific R-loops genome-wide. *Elife* 10. <https://doi.org/10.7554/eLife.65146>
- Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.-C., Dufour, M.-E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., Libri, D., Jacquier, A., 2005. Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase. *Cell* 121, 725–737. <https://doi.org/10.1016/j.cell.2005.04.030>
- Xie, Y., Kerscher, O., Kroetz, M.B., McConchie, H.F., Sung, P., Hochstrasser, M., 2007. The yeast Hex3.Slx8 heterodimer is a ubiquitin ligase stimulated by substrate sumoylation. *J. Biol. Chem.* 282, 34176–34184. <https://doi.org/10.1074/jbc.M706025200>
- Xu, W., Li, K., Li, S., Hou, Q., Zhang, Y., Liu, K., Sun, Q., 2020. The R-Loop Atlas of Arabidopsis Development and Responses to Environmental Stimuli. *Plant Cell* 32, 888–903. <https://doi.org/10.1105/tpc.19.00802>
- Xu, W., Xu, H., Li, K., Fan, Y., Liu, Y., Yang, X., Sun, Q., 2017. The R-loop is a common chromatin feature of the Arabidopsis genome. *Nat Plants* 3, 704–714. <https://doi.org/10.1038/s41477-017-0004-x>

- Xu, X.M., Rose, A., Muthuswamy, S., Jeong, S.Y., Venkatakrishnan, S., Zhao, Q., Meier, I., 2007. NUCLEAR PORE ANCHOR, the Arabidopsis homolog of Tpr/Mlp1/Mlp2/megator, is involved in mRNA export and SUMO homeostasis and affects diverse aspects of plant development. *Plant Cell* 19, 1537–1548. <https://doi.org/10.1105/tpc.106.049239>
- Yan, P., Liu, Z., Song, M., Wu, Z., Xu, W., Li, K., Ji, Q., Wang, S., Liu, X., Yan, K., Esteban, C.R., Ci, W., Belmonte, J.C.I., Xie, W., Ren, J., Zhang, W., Sun, Q., Qu, J., Liu, G.-H., 2020. Genome-wide R-loop Landscapes during Cell Differentiation and Reprogramming. *Cell Rep* 32, 107870. <https://doi.org/10.1016/j.celrep.2020.107870>
- Yan, Q., Shields, E.J., Bonasio, R., Sarma, K., 2019. Mapping Native R-Loops Genome-wide Using a Targeted Nuclease Approach. *Cell Reports* 29, 1369–1380.e5. <https://doi.org/10.1016/j.celrep.2019.09.052>
- Yan, Q., Wulfridge, P., Doherty, J., Fernandez-Luna, J.L., Real, P.J., Tang, H.-Y., Sarma, K., 2022. Proximity labeling identifies a repertoire of site-specific R-loop modulators. *Nat Commun* 13, 53. <https://doi.org/10.1038/s41467-021-27722-6>
- Yang, X., Liu, Q.-L., Xu, W., Zhang, Y.-C., Yang, Y., Ju, L.-F., Chen, J., Chen, Y.-S., Li, K., Ren, J., Sun, Q., Yang, Y.-G., 2019. m6A promotes R-loop formation to facilitate transcription termination. *Cell Res* 29, 1035–1038. <https://doi.org/10.1038/s41422-019-0235-7>
- Yang, Y., McBride, K.M., Hensley, S., Lu, Y., Chedin, F., Bedford, M.T., 2014. Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. *Mol Cell* 53, 484–497. <https://doi.org/10.1016/j.molcel.2014.01.011>
- Yang, Y., Wang, W., Chu, Z., Zhu, J.-K., Zhang, H., 2017. Roles of Nuclear Pores and Nucleo-cytoplasmic Trafficking in Plant Stress Responses. *Front Plant Sci* 8, 574. <https://doi.org/10.3389/fpls.2017.00574>
- Yasuhara, T., Kato, R., Hagiwara, Y., Shiotani, B., Yamauchi, M., Nakada, S., Shibata, A., Miyagawa, K., 2018. Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. *Cell* 175, 558–570.e11. <https://doi.org/10.1016/j.cell.2018.08.056>
- Yeeles, J.T.P., Poli, J., Marians, K.J., Pasero, P., 2013. Rescuing Stalled or Damaged Replication Forks. *Cold Spring Harb Perspect Biol* 5, a012815. <https://doi.org/10.1101/cshperspect.a012815>
- Yoo, T.Y., Mitchison, T.J., 2021. O-GlcNAc modification of nuclear pore complexes accelerates bidirectional transport. *Journal of Cell Biology* 220, e202010141. <https://doi.org/10.1083/jcb.202010141>
- Yoshida, T., Shimada, K., Oma, Y., Kalck, V., Akimura, K., Taddei, A., Iwahashi, H., Kugou, K., Ohta, K., Gasser, S.M., Harata, M., 2010. Actin-related protein Arp6 influences H2A.Z-dependent and -independent gene expression and links ribosomal protein genes to nuclear pores. *PLoS Genet.* 6, e1000910. <https://doi.org/10.1371/journal.pgen.1000910>
- Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T.E., Lieber, M.R., 2003. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* 4, 442–451. <https://doi.org/10.1038/ni919>
- Yu, K., Lieber, M.R., 2019. Current insights into the mechanism of mammalian immunoglobulin class switch recombination. *Crit. Rev. Biochem. Mol. Biol.* 54, 333–351. <https://doi.org/10.1080/10409238.2019.1659227>
- Yuan, W., Zhou, J., Tong, J., Zhuo, W., Wang, L., Li, Y., Sun, Q., Qian, W., 2019. ALBA protein complex reads genic R-loops to maintain genome stability in Arabidopsis. *Sci Adv* 5, eaav9040. <https://doi.org/10.1126/sciadv.aav9040>
- Zaitsev, E.N., Kowalczykowski, S.C., 2000. A novel pairing process promoted by Escherichia coli RecA protein: inverse DNA and RNA strand exchange. *Genes Dev.* 14, 740–749. <https://doi.org/10.1101/gad.14.6.740>
- Zander, G., Hackmann, A., Bender, L., Becker, D., Lingner, T., Salinas, G., Krebber, H., 2016. mRNA quality control is bypassed for immediate export of stress-responsive transcripts. *Nature* 540, 593–596. <https://doi.org/10.1038/nature20572>
- Zhang, C., Roberts, T.M., Yang, J., Desai, R., Brown, G.W., 2006. Suppression of genomic instability by SLX5 and SLX8 in *Saccharomyces cerevisiae*. *DNA Repair (Amst.)* 5, 336–346. <https://doi.org/10.1016/j.dnarep.2005.10.010>
- Zhang, H., Saitoh, H., Matunis, M.J., 2002. Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. *Mol Cell Biol* 22, 6498–6508. <https://doi.org/10.1128/MCB.22.18.6498-6508.2002>
- Zhang, Z.Z., Pannunzio, N.R., Hsieh, C.-L., Yu, K., Lieber, M.R., 2015. Complexities due to single-stranded RNA during antibody detection of genomic rna:dna hybrids. *BMC Res Notes* 8, 127. <https://doi.org/10.1186/s13104-015-1092-1>

- Zhao, X., Muller, E.G., Rothstein, R., 1998. A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. *Mol Cell* 2, 329–340. [https://doi.org/10.1016/s1097-2765\(00\)80277-4](https://doi.org/10.1016/s1097-2765(00)80277-4)
- Zhao, X., Wu, C.-Y., Blobel, G., 2004. Mlp-dependent anchorage and stabilization of a desumoylating enzyme is required to prevent clonal lethality. *J. Cell Biol.* 167, 605–611. <https://doi.org/10.1083/jcb.200405168>
- Zhou, L., Panté, N., 2010. The nucleoporin Nup153 maintains nuclear envelope architecture and is required for cell migration in tumor cells. *FEBS Lett* 584, 3013–3020. <https://doi.org/10.1016/j.febslet.2010.05.038>
- Zhu, Y., Liu, T.-W., Madden, Z., Yuzwa, S.A., Murray, K., Cecioni, S., Zachara, N., Vocadlo, D.J., 2016. Post-translational O-GlcNAcylation is essential for nuclear pore integrity and maintenance of the pore selectivity filter. *Journal of Molecular Cell Biology* 8, 2–16. <https://doi.org/10.1093/jmcb/mjv033>
- Zimmerli, C.E., Allegretti, M., Rantos, V., Goetz, S.K., Obarska-Kosinska, A., Zagoriy, I., Halavatyi, A., Hummer, G., Mahamid, J., Kosinski, J., Beck, M., 2021. Nuclear pores dilate and constrict in cellulo. *Science* 374, eabd9776. <https://doi.org/10.1126/science.abd9776>
- Zuccolo, M., Alves, A., Galy, V., Bolhy, S., Formstecher, E., Racine, V., Sibarita, J.-B., Fukagawa, T., Shiekhattar, R., Yen, T., Doye, V., 2007. The human Nup107-160 nuclear pore subcomplex contributes to proper kinetochore functions. *EMBO J* 26, 1853–1864. <https://doi.org/10.1038/sj.emboj.7601642>