THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biochimie et biologie moléculaire

École doctorale CBS2

Unité de recherche UMR 9002

Role of NF90 in the miRNA pathway

Présentée par Giuseppa Grasso

Le 08 novembre 2021

Sous la direction de Dr. Rosemary Kiernan

Devant le jury composé de

Dr. Florence Apparailly, Directrice de Recherche, INSERM U1183	Président du jury
Dr. Michele Trabucchi, Directeur de Recherche, INSERM U1065	Rapporteur
Dr. Dominique Weil, Directrice de Recherche, CNRS UMR 7622	Rapportrice
Dr. Rosemary Kiernan, Directrice de Recherche, CNRS UMR 9002	Directrice de Thèse

UNIVERSITÉ DE MONTPELLIER

Acknowledgements

First and foremost I wish to thank my PhD supervisor Dr. Rosemary Kiernan for her invaluable advice and support during these years. I'm grateful for our insightful scientific discussions that made this PhD experience successful.
I would like also to thank the members of my thesis defense jury for accepting to evaluate my work.
I'm grateful to all team members and especially Celine, Xavier, Kader, Marina and Matthieu for their support in the lab and outside. I wish to thank our collaborators for helping us improving the project and our publication.

I would like to express my gratitude to my family and friends, especially Xavier, Julia and Elena, for our chats, game nights and apèros on the beach that I will never forget.
Finally, I would like to thank Claudio for sharing with me this experience, for his tremendous understanding, encouragement and bioinformatic advice and support.

Contents

Acknowledgements **iii**

Abstracts **1**

French summary **4**

Preface **11**

1 Introduction **15**

1.1 Post-transcriptional gene regulation 15

1.1.1 Non-coding RNAs 16

small ncRNAs 18

1.2 microRNAs 19

1.2.1 Overview on miRNAs 19

1.2.2 Evolution of miRNAs 22

Transcriptional control of novel miRNAs 24

1.3 miRNA biogenesis 25

1.3.1 Overview on miRNA biogenesis 25

1.3.2 miRNAs transcription 27

1.3.3 Microprocessor cleavage 30

Microprocessor structure 31

Recognition and cleavage of the pri-miRNA stem loop by DROSHA and DGCR8 32

Microprocessor non-canonical substrates 35

pri-miRNA cleavage is co-transcriptional 36

Transcriptional and post-transcriptional regulation of DROSHA and DGCR8 39

RNA-binding proteins involved in Microprocessor activity . . . 40
1.3.4 Stability of pri-miRNA hairpins . . . 42
1.3.5 pre-miRNAs export . . . 43
1.3.6 DICER cleavage . . . 46
DICER structure . . . 47
Recognition and cleavage of the pre-miRNA by DICER . . . 49
Other proteins involved in DICER processing . . . 51
Non-canonical DICER substrates and functions . . . 53
DICER regulation . . . 54
1.4 RISC-mediated gene silencing . . . 55
1.4.1 AGO2 structure . . . 56
1.4.2 RISC loading and interplay between DICER and AGO2 . . . 58
1.4.3 Molecular mechanism of AGO2 . . . 60
Interplay between translation and RISC mediated silencing . . 60
Target mRNA recognition by AGO2 . . . 62
Interaction of miRNA and target mRNA . . . 64
RISC mediated translation inhibition, mRNA decapping, deadenylation and decay . . . 66
Non-canonical activities of Ago proteins . . . 68
1.4.4 AGO2 regulation . . . 69
1.4.5 Other proteins involved in RISC-mediated mRNA silencing . . 70
1.5 miRNAs stability and decay . . . 71
1.6 Non-canonical miRNAs . . . 73
miRtrons . . . 73
Simtrons . . . 75
snoRNA and tRNA-derived miRNAs . . . 76
Nuclear miRNAs . . . 77
1.7 miRNAs in gene regulation . . . 77
1.7.1 miRNAs in homeostatic gene silencing . . . 78
1.7.2 miRNAs in cancer . . . 78
1.8 NF90, double stranded RNA binding protein . . . 80
1.8.1 NF90 structure and binding mode . . . 81

Post translational modifications and regulation of NF90 83
1.8.2 NF90 functions 84
NF90 in transcription regulation 84
NF90 in the miRNA biogenesis pathway 85
NF90 in mRNA translation, stability and degradation 86
NF90 in viral replication 88

2 NF90 modulates processing of a subset of human pri-miRNAs 91

3 NF90 Interacts with Components of RISC and Modulates Association of Ago2 with mRNA 133

4 Discussion 165

A NF90 modulates processing of a subset of human pri-miRNAs 175

B Chromatin-associated MRN complex protects highly transcribing genes from genomic instability 203

List of Figures

1.1 Schematic representation of RNA classification 17
1.2 Overview of the miRNA biogenesis pathway 25
1.3 miRNA genome localization and their transcription 27
1.4 Microprocessor protein structure 31
1.5 Sequence motifs on pri-miRNAs 33
1.6 Structure of XPO5 complexed with a pre-miRNA 44
1.7 Schematic representation of DICER structure in complex with a pre-miRNA. 47
1.8 Model for DICER processing of pre-miRNA 50
1.9 Schematic representation of AGO2 structure 56
1.10 Schematic overview of miRtron biogenesis 75
1.11 Schematic overview of snoRNA and tRNA-derived miRNAs biogenesis 76
1.12 Schematic representation of NF90 and NF110 81

List of Abbreviations

ADAR	**A**denosine **D**eaminase **A**cting on **R**NA
AGO	**Ar**g**o**naute
AGPAT5	1-**A**cyl**g**lycerol-3-**P**hosphate O-**a**cyl**t**ransferase **5**
ARRE-2	**A**ntigen **R**ecognition **R**esponse **E**lement **2**
ASOs	**A**nti**s**ense **O**ligonucleotide**s**
AUF1	**AU**-rich binding **F**actor **1**
BAFF	**B**-cell **A**ctivating **F**actor
bp	**b**ase**p**air
C9ORF	**C**hromosome **9** **O**pen **R**eading **F**rame
CAGE-seq	**C**ap-**A**nalysis of **G**ene **E**xpression-**seq**uencing
CBC	**C**ap-**B**inding **C**omplex
CCR4	**C**-**C** motif chemokine **R**eceptor **4**
CDK	**C**yclin-**D**ependent **K**inase
CED	**CE**ntral **D**omain
ChIP-seq	**Ch**romatin **I** mmuno**P**recipitation-**seq**uencing
circRNA	**circ**ular **R**ibo**N**ucleic **A**cid
CLL	**C**hronic **L**ymphocytic **L**eukemia
CNOT	**C**CR4 **NOT** transcription complex
COL27A1	**COL**lagen alpha-1 **27** chain **A1**
cryo-EM	**cryo**genic-**E**lectron **M**icroscopy
CRM1	**C**h**R**omosomal **M**anteinance **1**
CTDSP1	**C**arboxyl-**T**erminal **D**omain **S**mall **P**hosphatase **1**
CTT	**C**-**T**erminal **T**ail
DCP2	**Dec**apping **P**rotein **2**
DDX5	**D**ea**D** bo**X** helicase **5**
DDR	**D**NA **D**amage **R**esponse

DGCR8	**D**i**G**eorge syndrome **C**ritical **R**egion **8**
DHS	**D**Nase I **H**ypersensitive **S**ite
DNA	**D**eoxyribo**N**ucleic **A**cid
dsRBD	**d**ouble **s**tranded **R**NA **B**inding **D**omain
dsRNA	**d**ouble **s**tranded **R**ibo**N**ucleic **A**cid
DUF283	**D**omain of **U**nknown **F**unction **283**
DV	**D**engue **V**irus
DZF	**D**sRBM and **Z**inc **F**inger associated
EBOV	**EBO**la **V**irus
eCLIP	**e**nhanced **C**rosslinking **I**mmuno**P**recipitation
EDC3	**E**nhancer of **Dec**apping **3**
eIF4F	**e**ukaryotic translation **I**nitiation **F**actor **4F**
EMSA	**E**lectrophoretic **M**obility **S**hift **A**ssay
EMT	**E**pithelial-to-**M**esenchymal **T**ransition
ENCODE	**ENC**yclopedia **O**f **D**NA **E**lements
ERK	**E**xtracellular signal **R**egulated **K**inase
fCLIP	formaldehyde **C**rosslinking **I**mmuno**P**recipitation
FMRP	**F**ragile X **M**ental **R**etardation **P**rotein
GDF15	**G**rowth **D**ifferentiation **F**actor **15**
GO	**G**ene **O**ntology
GSK3β	**G**lycogen **S**ynthase **K**inase **3**β
HCC	**H**epato**C**ellular **C**arcinoma
HCV	**H**epatitis **C** **V**irus
HDAC1	**H**istone **D**eacetylase **1**
HEK-293T	**H**uman **E**mbryonic **K**idney **293** (SV40 large) **T** (antigen)
HIV-1	**H**uman **I**mmunodeficiency **V**irus-**1**
hnRNP	**H**eterogeneous **n**uclear **R**ibo**N**ucleo**P**rotein
HP1BP3	**H**eterochromatin **P**rotein **1** **B**inding **P**rotein **3**
HPV	**H**uman **P**apilloma **V**irus
HTLV1	**H**uman **T** **L**ymphoma **V**irus **1**
HuR	**Hu** antigen **R**

IAV	**I**nfluenza **A** **V**irus
iCLIP	**i**ndividual-nucleotide-resolution UV **C**rosslinking **I**mmuno**P**recipitation
IL2	**I**nter**l**eukin **2**
ILF3	**I**nter**l**eukin enhancer binding **F**actor **3**
IP	**I**mmuno**P**recipitation
kb	**k**ilo**b**ase
KD	**K**nock-**D**own
KO	**K**nock-**O**ut
KSRP	**K**H-type **S**plicing **R**egulatory **P**rotein
LET	**L**ow **E**xpressed in **T**umors
lncRNA	**l**ong **n**on-**c**oding **R**ibo**N**ucleic **A**cid
MAPK	**M**itogen-**A**ctivated **P**rotein **k**inase
miPDC	**mi**RNA **P**recursor **D**eposit **C**omplex
miRNA	**mi**cro **R**ibo**N**ucleic **A**cid
miRISC	**mi**RNA **RISC**
miRLC	**miR**NA **L**oading **C**omplex
MOV10	**MO**loney leukemia **V**irus **10**
MRE	**M**iRNA **R**ecognition **E**lement
mRNA	**m**essenger **R**ibo**N**ucleic **A**cid
MSI2	**Mus**ash**i** homolog **2**
ncRNA	**n**on-**c**oding **R**ibo**N**ucleic **A**cid
NF90	**N**uclear **F**actor **90**
Ngn2	**N**eurog**en**in **2**
NLS	**N**uclear **L**ocalization **S**ignal
NSCLC	**N**on-**S**mall-**C**ell **L**ung **C**arcinoma
nt	**n**ucleo**t**ide
PABP	**P**oly**(A)**-**B**inding **P**rotein
PACT	**P**KR **ACT**ivating (protein)
PAZ	**P**iwi-**A**rgonaute-**Z**wille
PB	**P**rocessing **B**odies
PHAX	**P**hosp**h**orylated **A**dapter RNA e**x**port protein

PK	**P**rotein **K**inase
PKR	**P**rotein activator of the **K**inase **R**
pre-miRNA	**pre**cursor **mi**cro **R**ibo**N**ucleic **A**cid
pre-RISC	**pre**cursor **RISC**
P-rich	**P**roline-**rich**
pri-miRNA	**pri**mary **mi**cro **R**ibo**N**ucleic **A**cid
RACE	**R**apid **A**mplification of **C**DNA **E**nds
Ran-GTP	**Ra**s-related **N**uclear protein- **G**uanosine **T**ri**P**hosphate
RBP	**R**NA **B**inding **P**rotein
Rhed	**R**NA-binding **he**me **D**omain
RIP	**R**NA **I**mmuno**P**recipitation
RISC	**R**NA-**I**nduced **S**ilencing **C**omplex
RNA	**R**ibo**N**ucleic **A**cid
RNAi	**R**ibo**N**ucleic **A**cid **i**nterference
RNA Pol II	**R**ibo**N**ucleic **A**cid **Pol**ymerase **II**
rRNAs	**r**ibosomal **R**ibo**N**ucleic **A**cid
SF1	UPF-like helicase **S**uper**f**amily **1**
simtrons	**S**plicing-**i**ndependent **mi**Rtron-like miRNAs
siRNA	**s**hort-**i**nterfering **R**ibo**N**ucleic **A**cid
sncRNA	**s**mall **n**on-**c**oding **R**ibo**N**ucleic **A**cid
snRNA	**s**mall **n**uclear **R**ibo**N**ucleic **A**cid
ssRNA	**s**ingle **s**tranded **R**ibo**N**ucleic **A**cid
TAR	**T**rans**a**ctive **R**esponse (binding protein)
TIAM2	**T**-cell lymphoma **I**nvasion and **M**etastasis **2**
TDMD	**T**arget-**D**irected **M**icroRNA **D**egradation
TE	**T**ransposable **E**lement
TNRC6A	**Tri**nucleotide **R**epeat-**C**ontaining gene **6A** (protein)
TRBP	**T**AR **R**NA-**B**inding **P**rotein
tRFs	**tR**NA-derived **F**ragment**s**
tRNAs	**t**ransfer **R**ibo**N**ucleic **A**cid
TSS	**T**ranscription **S**tarting **S**ite

UTR	**U**n**T**ranslated **R**egion
XPO5	E**XPO**rtin **5**
XRN1	5′-to-3′ nuclease e**X**ori**bo**n**u**clease **1**
ZRANB2	**Z**inc finger **RAN-b**ingind domain containing **2**
ZFR	**Z**inc **F**inger **R**NA-binding (protein)

Abstract in english

Nuclear Factor 90 (NF90) is a double-stranded RNA-binding protein (RBP) found in both the nucleus and the cytoplasm that is involved in a plethora of different cellular processes and pathways, such as transcription, splicing, translation and mRNA stability or degradation. For this reason, NF90 compartmentalization and shuttling from the nucleus to the cytoplasm is a very important and strictly controlled process, guided by several stimuli such as viral infection or hypoxia. With this PhD we analyzed the role of NF90 in the microRNA (miRNA) biogenesis, in the nucleus, and its role in translational control and messenger RNA (mRNA) stability, in the cytoplasm.

In the nucleus, NF90 was recently shown to be involved in miRNA biogenesis regulation by negatively affecting the Microprocessor activity. However, this mechanism, which our lab previously linked to ovarian carcinoma progression and metastasis, is not fully understood. Here, we show the extent of NF90-mediated pri-miRNA regulation in hepatocellular carcinoma (HCC), for which NF90 is associated with poor prognosis. Genome-wide approaches revealed that NF90 increases the abundance of 286 miRNAs in HepG2 cell line. Of these, 22 pri-miRNAs are directly associated with NF90 through their stem region, in a manner that is largely exclusive of Microprocessor. NF90-targeted pri-miRNAs are mainly intronic, highly stable and have lower free energy and fewer mismatches compared to all human pri-miRNAs. A group of mRNAs hosting NF90-bound and modulated pri-miRNAs were significantly downregulated after loss of NF90 or showed splicing defects in the introns containing the pri-miRNAs. These findings suggest that NF90 is involved in the biogenesis of a subset of highly stable, intronic miRNAs.

In the cytoplasm, NF90 was shown to bind mRNAs, which increases their stability or influences their translation. However, the exact mechanism is still largely unknown. Here, we show that NF90 interacts with RBPs involved in RNA-induced silencing complex (RISC)-mediated silencing, such as Moloney leukemia virus 10 (MOV10) and Argonaute 2 (Ago2), in an RNA-dependent manner. Upon glycerol gradient sedimentation, we found that NF90, MOV10 and Ago2 can be found in the same complex in HEK293T cell line. Using published data of NF90 and MOV10 enhanced crosslinking immunoprecipitation (eCLIP) and individual-nucleotide-resolution UV crosslinking (iCLIP), respectively, we identified a subset of mRNAs that can be bound by both NF90 and MOV10 in the 3'UTR. RNA immunoprecipitation (RIP) analyses suggest that the binding of MOV10 to target mRNAs might prevent or reduce the binding of NF90 on the same targets, and vice versa. Moreover, loss of NF90 increased association of Ago2 to the target mRNAs while reducing their abundance. These findings suggest that NF90 might have a role in RISC-mediated silencing by modulating Ago2 association with mRNAs and thereby enhancing their stability.

Abstract in french

Le facteur nucléaire 90 (NF90) est une protéine de liaison à l'ARN (RBP) double brin qui est impliquée dans une pléthore de différents processus et voies cellulaires à la fois dans le noyau et le cytoplasme, tels que la transcription, l'épissage, la traduction et la stabilité ou la dégradation de l'ARNm. Pour cette raison, la compartimentation de NF90 dans le noyau ou le cytoplasme est un processus très important contrôlé par plusieurs stimuli tels que l'infection virale ou l'hypoxie. Au cours de cette thèse, nous avons analysé le rôle de NF90 dans la biogenèse des microARN (miARN), dans le noyau, et son rôle dans le contrôle de la traduction et la stabilité des ARN messagers (ARNm), dans le cytoplasme.

Al a récemment été montré que la fraction nucléaire de NF90 est impliquée dans la régulation de la biogenèse des miARN en réprimant l'activité du microprocesseur. Cependant, ce mécanisme, que notre laboratoire avait précédemment lié à la progression et aux métastases du carcinome ovarien, n'est pas entièrement compris. Ici, nous montrons l'étendue de la régulation des pri-miARN médiés par NF90 dans le carcinome hépatocellulaire (CHC), pour lequel NF90 est un facteur de mauvais pronostic connu. Des approches à l'échelle du génome ont révélé que NF90 augmente potentiellement l'abondance de 286 miARN dans la lignée cellulaire HepG2. Parmi ceux-ci, 22 pri-miARN sont directement associés à NF90 via leur région en tige, alors qu'il ne sont pas liés au microprocesseur. Les pri-miARN ciblés par NF90 sont principalement introniques, très stables et ont une énergie libre plus faible et moins de mésappariements par rapport à tous les pri-miARN humains. Un groupe d'ARNm hébergeant des pri-miARN modulés et liés à NF90 était significativement diminué après la perte de NF90 ou présentait un défaut d'épissage dans les introns contenant les pri-miARN. Ces résultats suggèrent que NF90 est impliqué dans la biogenèse d'un sous-ensemble de miARN

introniques hautement stables.

Dans le cytoplasme, NF90 s'est avéré se lier aux ARNm, augmentant leur stabilité ou influençant leur traduction. Cependant, le mécanisme exact est encore largement inconnu. Ici, nous montrons que NF90 interagit avec des RBP impliquées dans le silençage induit par l'ARN (RISC), telles que le virus de la leucémie Moloney 10 (MOV10) et Argonaute 2 (Ago2), d'une manière dépendante de l'ARN. Lors de la sédimentation en gradient de glycérol, nous avons constaté que NF90, MOV10 et Ago2 peuvent être trouvés dans le même complexe dans la lignée cellulaire HEK293T. En utilisant les données publiées d'immunoprécipitation après cross-linking de NF90 (eCLIP) et de cross-linking UV à résolution nucléotidique de MOV10 (iCLIP), nous avons constaté que NF90 et MOV10 peuvent se lier aux 3'UTR des mêmes ARNm. Les analyses d'immunoprécipitation d'ARN (RIP) suggèrent que la liaison de MOV10 aux ARNm cibles pourrait empêcher ou réduire la liaison de NF90 sur les mêmes cibles, et vice versa. De plus, la perte de NF90 augmente l'association d'Ago2 aux ARNm cibles tout en réduisant leur abondance. Ces résultats suggèrent que NF90 pourrait avoir un rôle dans le silencage induit par RISC en inhibant sa fonction et en stabilisant les ARNm.

French summary

Introduction

Les microARN (miARN) sont une classe d'ARN courts non codants d'une longueur d'environ 20 à 22 nucléotides de longueur. Ils régulent l'expression de plus de 60% des gènes de mammifères grâce à un mécanisme de silençage génique post-transcriptionnel appelé interférence ARN (ARNi). Par conséquent, les miARN sont impliqués dans plusieurs processus biologiques et ils sont fréquemment dérégulés dans de nombreux types de cancers(Peng & Croce, 2016).

Les miARN sont transcrits par une ARN polymérase II sous la forme de longs miARN primaires (pri-miARN). Leur maturation commence dans le noyau avec le clivage en miARN précurseurs (pre-miARN) d'environ 65 nucléotides, par le microprocesseur. Après le clivage, le pre-miARN est exporté vers le cytoplasme par Exportin5 (XPO5) afin de subir un deuxième clivage par DICER pour obtenir le miARN. Le miARN est ensuite pris en charge par les protéines Argonautes (Ago) qui permettent la formation du complexe pré-miRISC (RNA-Induced Silencing Complex). Enfin, la maturation des miARN se conclut avec la formation du complexe miRISC, après l'élimination du brin passager du miARN et sa dégradation. Le complexe miRISC mature, par le biais du brin guide , permet de réguler l'expression génique, par sa fonction de répression. Le brin guide permet d'associer RISC à l'ARNm auquel il est complémentaire. La liaison imparfaite du miARN sur sa séquence cible a pour conséquence de réprimer l'expression des gènes en détruisant l'ARNm par l'activité des exoribonucléases ou par blocage de la traduction (Ha & Kim, 2014). De nombreuses protéines, en particulier des protéines de liaison à l'ARN, sont impliquée es dans ce processus et, parmi elles, NF90 (Michlewski & Cáceres, 2019).

Le facteur nucléaire 90 (NF90) est une protéine de liaison à l'ARN double brin qui

se lie à des ARN structurés, tels que les ARN viraux (Y. Li & Belshan, 2016) et les miARN (Barbier et al., 2018) et qui a, récemment, été impliqué dans la régulation de la biogenèse des miARN (Barbier et al., 2018; Higuchi et al., 2016). Il a été montré que NF90 est capable de lier les pri-miARN en compétition avec la première étape de la biogenèse des miARN catalysée par le microprocesseur. Nous avons précédemment montré que, lorsque NF90 est déplété, le clivage du pri-miARN par le microprocesseur interfère avec l'épissage du gène hôte, DICER, ce qui diminue son niveau d'expression de protéine. Enfin, nous avons montré que dans les cellules de carcinome ovarien primaire, la surexpression de NF90 favorise l'expression de DICER et réduit la prolifération; tandis que la surexpression de NF90 dans les modèles de souris nues avec xénogreffe conduit à une réduction significative de la taille de la tumeur et des métastases. Par conséquent, nous avons identifié NF90 comme un bon facteur de pronostic pour le cancer de l'ovaire. Cependant, la littérature sur le rôle de NF90 dans la biogenèse des miARN est limitée (Barbier et al., 2018; Higuchi et al., 2016). Pour cette raison, dans le cadre de mon projet de doctorat, nous avons décidé de détecter tous les miARN potentiellement modulés par NF90 dans les HepG2, une lignée de carcinome hépatocellulaire (CHC); pour lequel NF90 est un facteur de mauvais pronostic connu.

En plus de contrôler et de moduler la maturation des miARN, NF90 peut réguler directement la stabilité des ARNm spécifiques dans le cytoplasme. En effect, dans le cytoplasme, NF90 a déjà été associé à la stabilité des ARNm humains et viraux, étant impliqué dans leur traduction et leur dégradation (Castella et al., 2015; Vumbaca et al., 2008). De plus, il a été montré qu'à la suite d'une infection virale, la NF90 est phosphorylée ce qui conduit à une dissociation de NF45 et l'export nucléaire de NF90. Le NF90 phosphorylé s'accumule sur les ribosomes où il s'associe aux ARN viraux inhibant leur traduction (Harashima et al., 2010). Cependant, le mécanisme de régulation des mARN par NF90 reste encore inconnu. Au cours de ma thèse, j'ai essayé de comprendre le rôle de la NF90 dans le cytoplasme et en particulier comment son effet sur les ARNm est médié.

Matériel et méthodes

Pour mieux comprendre l'effet de NF90 sur la biogenèse des miARN dans les HepG2, j'ai effectué un profilage de miARN et un RNA-seq, après une diminution de NF90 par des siARN, pour détecter les miARN potentiellement régulés par NF90 et le niveau de leurs gènes hôtes. Pendant la deuxième année de ma thèse, j'ai eu la possibilité d'apprendre à analyser ces données complexes de séquençage sous différents aspects tels que l'expression différentielle, l'analyse d'épissage et leur représentation graphique. Ces connaissances ont également été appliquées à l'extrapolation de nouveaux résultats par l'analyse des données d'immunoprécipitation par CrossLinking UV (eCLIP) de NF90, réalisée dans HepG2 et récemment publiée par une autre équipe (Nussbacher & Yeo, 2018).

Pour comprendre pourquoi NF90 ne semble se lier et moduler qu'un sous-ensemble de miARN nous avons collaboré avec un groupe de bioinformaticiens pour identifier les caractéristiques de ce sous-groupe de miARN. Ensuite, nous avons testé ces résultats en effectuant des mutations au sein des pri-miARN liés par NF90 afin de former des mésappariements et réduire leur stabilité et nos collaborateurs ont mesuré leur association à NF90 par Electrophoretic Mobility Shift Assay (EMSA). J'ai aussi réalisé des analyses d'épissage en utilisant des PCR quantitatives (qPCR) avec des amorces projetées pour amplifier des régions spécifiques du mARN hôte.

Enfin, l'une des techniques les plus importantes et laborieuses de cet article est l'immunoprécipitation de l'ARN (RIP) que j'ai utilisé pendant tout ce travail pour valider et renforcer les données décrites dans mon article.

Concernant le rôle de NF90 dans le cytoplasme, nous avons identifié son interactome en réalisant une spectrométrie de masse après une immunopurification en tandem de NF90 exogène dans les cellules HEK293T. Afin de valider les résultats de la spectrométrie de masse, et de comprendre si la liaison de NF90 à ses interactants est dépendante de l'ARN, nous avons réalisé des Co-IPs de NF90 en présence de cocktail RNAse A/T1. De plus, un gradient de glycérol a été réalisé pour comprendre si ces protéines se trouvaient dans le même complexe ou dans des complexes NF90 différents. Pour comprendre l'effet de la liaison de NF90 aux

ARNm et son interaction avec MOV10, nous avons effectué des RIPs sur une sélection d'ARNm cibles liés par MOV10 et NF90, après dérégulation de l'un ou l'autre facteur. Enfin, puisque MOV10 est une hélicase impliquée dans le silençage génique médié par RISC, nous avons effectué des RIPs d'Ago2 sur le même ARNm cibles liés par NF90 et MOV10.

Résultats

Afin de comprendre l'ampleur de l'effet de NF90 dans la biogenèse des miARN, nous avons effectué un séquençage de miARN dans HepG2. L'analyse de l'expression différentielle suggère que, après la perte de NF90, 268 miARN sur 1661 miARN humains exprimés dans HepG2 ont été régulés à la hausse. En utilisant un jeu de données de eCLIP, nous avons pu identifier 38 pri-miARN dont la région double brin est directement liée par NF90. Puisque des études précédentes ont montrées que NF90 est en concurrence avec le microprocesseur pour la liaison des pri-miARN, nous avons également analysé les eCLIP de DGCR8 et DROSHA. Comme prévu, environ 60% des pri-miARN liés par DGCR8 étaient également liés par DROSHA. Au contraire, seulement environ 20% des pri-miARN liés à NF90 étaient significativement liés par DGCR8 ou DROSHA, ce qui indique que les pri-miARN liés à NF90 ne sont pas fortement associés au microprocesseur.

Nous nous sommes également demandé si la liaison de NF90 aux pri-miARN interférait avec le clivage par le microprocesseur. Si tel est le cas, nous détecterions une augmentation du niveau de miARN matures après la perte de NF90. Comme prévu, le séquençage des miARN a identifié 22 miARN parmi les 38 pri-miARN associés à NF90, qui ont montré une augmentation de leur produit mature. Nous avons nommé ces miARNs «double-positifs»: ils sont associés à NF90 et leur abondance est augmentée suite à la perte de NF90.

Un autre objectif de mon projet était de comprendre pourquoi NF90 semble se lier et ne moduler qu'un sous-ensemble de miARN. En analysant la structure des 22 miARN doubles positifs, nous avons constaté que, comparés à tous les pri-miARN humains, ils ont moins de mésappariements et une énergie libre plus faible; ce qui signifie qu'ils sont très stables. Pour tester cette hypothèse, nous avons conçu des

mutations au sein des pri-miARN liés par NF90 afin de former des mésappariements et réduire leur stabilité et nous avons testé leur association à NF90 par EMSA.

Les résultats montrent que si les structures sauvages étaient fortement liées par NF90, leurs structures mutées n'étaient que faiblement associées ou pas du tout associées à NF90. Il est intéressant de noter que les caractéristiques structurelles trouvées pour les pri-miARN double positifs étaient également partagées entre les pri-miARN modulés par NF90 mais pas positifs par eCLIP. Il est en effet possible que certains pri-miARN associés à NF90 n'aient pas été détectés par l'analyse eCLIP. En effet, l'analyse par EMSA a confirmé la liaison de NF90 à 2 pri-miARN, très stables et avec peu de mésappariements, sélectionnés dans le groupe régulé à la hausse mais négatif par eCLIP.

Comme expliqué précédemment, dans le cas de miR-3173, NF90 semble moduler l'efficacité d'épissage du gène hôte pri-miARN, DICER. Pour comprendre si cela s'appliquait également à nos nouvelles cibles (20 pri-miARNs sont introniques), nous avons effectué un RNA-seq et nous avons montré que la régulation à la baisse de NF90 diminue l'expression de 3 gènes contenant des pri-miARNs associés et modulés par NF90. De plus, en effectuant des qPCR, nous avons identifié un sous-ensemble de gènes hôtes dont l'efficacité d'épissage était diminuée pour les introns contenant les pri-miARN, comme par exemple TIAM2 hébergeant pri-miR-1273c. Les niveaux de protéines diminuées ont été mesurés par Western Blot soutenant la régulation négative et les défauts d'épissage des gènes hôtes observés par RNA-seq et qPCR.

Pour ce qui concerne le rôle de NF90 dans le cytoplasme, grâce à la spectrométrie de masse, nous avons montré que NF90 interagit avec de nombreuses protéines faisant partie du complexe RISC ou associées à la répression traductionnelle et au métabolisme des ARNm, telles que AGO2 et MOV10. De plus, nous avons montré que NF90 est dans le même complexe protéique avec AGO2 et MOV10 et que leur interaction se fait par l'ARN. En analysant des donne de eCLIP et iCLIP déjà publiés (Kenny et al., 2014; Nussbacher & Yeo, 2018), nous avons montré que les 41% des ARNm liés à MOV10 sont également liés par NF90, suggérant que leur

interaction se produit souvent dans la cellule. Afin de comprendre la nature de l'interaction entre NF90 et MOV10, nous avons effectué des RIP sur certaines cibles sélectionnées, après dérégulation de l'une ou l'autre protéines. Les résultats ont montré qu'après une dérégulation de NF90, nous pouvons détecter plus d'ARNm lié à MOV10 et vice versa, suggérant un mécanisme de compétition entre ces deux protéines. MOV10 est connu pour promouvoir l'association d'Ago2 aux ARNm, améliorant ainsi la repression médié par RISC. Étant donné que NF90 et MOV10 s'influencent mutuellement sur la liaison aux ARNm cibles communs, nous nous sommes demandé si l'altération du complexe d'association NF90/NF45 pouvait avoir des effets similaires sur la liaison d'Ago2. À cette fin, nous avons effectué un RIP d'Ago2 ou IgG après repression de NF90/NF45. La perte de NF90/NF45 a augmenté de manière significative la liaison d'Ago2 aux ARNm cibles testés.

Discussion

Nous avons montré que NF90 a un effet plus important sur la biogenèse de pri-miARN qu'on ne le pensait auparavant. Avec nos travaux sur NF90 et les miARN, nous avons identifié un sous-ensemble de pri-miARN humains qui sont liés par NF90 et qui partagent une structure similaire et très stable. Ces données étendent nos connaissances sur la façon dont les pri-miARN peuvent être modulé par les protéines de liaison a l'ARN et la compréhension des perturbations des miARN dans des conditions pathologiques comme le CHC. Le carcinome hépatocellulaire est la tumeur du foie la plus courante et la deuxième cause de décès par cancer. Le développement et la progression du CHC sont des processus extrêmement compliqués et, à ce jour, les mécanismes moléculaires sous-jacents sont encore largement inconnus. Des études plus approfondies sont donc nécessaires pour clarifier les mécanismes menant à la progression du cancer et aux métastases et cela apporterait de la lumière sur de nouvelles cibles thérapeutiques et / ou des biomarqueurs efficaces pour un diagnostic précoce.

Pour ce qui concerne le rôle de NF90 dans le cytoplasme, nous avons montré que NF90 et l'hélicase MOV10 interfèrent négativement pour la liaison des ARNm

cibles. Cette compétition, probablement indirecte, a un effet sur le recrutement d'AGO2 vers l'ARNm qui conduit à la modulation du sort de l'ARNm. Ces nouvelles découvertes pourraient avoir des implications importantes lors d'une infection virale ou d'un cancer, car il est connu que le niveau de NF90 dans le cytoplasme est modifié dans le cadre de ces cas. L' interfèrence entre la liason de NF90 et MOV10, que nous avons montré, pourrait donc affecter le recrutement de Ago2 aux ARN cibles et par conséquent diminuer l'efficacité du silençage génique induit par le complexe RISC. La compréhension de ces mécanismes pourra nous permettre, à l'avenir, d'aborder des pathologies dans lesquelles le niveau de NF90 est modulé, comme le cancer ou des infections virales et peut être proposer des thérapies ciblées et donc plus efficaces.

Preface

RNA and, in particular, non-coding and regulatory RNAs have been my main research and study interest for the last ten years.

Since its discovery in 1868 by Friedrich Miescher, RNA has engaged and puzzled scientists in the world with its complexity and plethora of different fundamental functions, today still in expansion. Such a simple and versatile molecule is able to organize to generate extreme complexity and variety, with so much more to unveil.

With this thesis, I hope to convey my passion and eagerness towards this remarkable and surprising field and science, in general.

A PhD is never easy, but I believe that my daily dedication and excitement to contribute to the current knowledge has made this last four years invaluable both personally and professionally.

Chapter 1

Introduction

1.1 Post-transcriptional gene regulation

Appropriate and controlled gene expression is essential for every living organism, as the production of a required amount of a specific protein at the right moment and location is critical for homeostasis, differentiation and development (Christou-Kent et al., 2020; Willimott & Wagner, 2010). As a consequence, alteration of gene expression can have profound effects and it can often be responsible for diseases such as cancer and genetic disorders (Audic & Hartley, 2004). Therefore, given the importance of controlled genes expression, many mechanisms have evolved at this purpose.

While specific gene expression can be achieved at the gene transcription level, numerous regulatory mechanisms contribute to gene expression pattern post-transcriptionally. Post-transcriptional gene regulation is involved in a wide range of processes and acts at different levels from splicing to export, translation initiation and RNA degradation (Corbett, 2018). In a RNA molecule, the 5′ untranslated region (UTR) and the 3′UTR are often targets of a variety of mechanisms that ultimately lead to regulation of RNA degradation, stability or mRNA translational efficiency. Across the animal kingdom, several conserved sequence elements, which are targets of gene regulation mechanisms, can be found in the UTRs of protein-coding RNAs or in non-coding RNAs (ncRNAs) (Duret et al., 1993). Other crucial elements for RNA regulation are the 5′ mRNA capping (5′ cap) and the 3′ terminal poly(A). The 5′ cap is a 7-methylguanylate (m^7G)

modification of the first guanine nucleotide at the 5' end of a mRNA molecule and is involved in nuclear export regulation, stability of the mRNA and translation initiation (Shuman, 2015). On the other hand, the poly(A) tail consists of multiple adenosine monophosphates at the 3' end of several RNA classes. The length of the poly(A) tail dynamically controls translation efficiency and stability (Nicholson & Pasquinelli, 2019).

All the above sequence and structural regulatory elements in the RNA molecules can be bound by a multitude of RNA-binding proteins (RBP) with a large variety of functions (Gerstberger et al., 2014). By binding to specific RNAs, RBP form ribonucleoproteins (RNP) complexes that are able to coordinate transport, stability, maturation and degradation of all classes of RNAs. As mentioned before, the UTRs of an mRNA usually contain specific RNA sequence motifs that can be target of RBP or RNA molecules themselves. In fact, mRNA regulation is a very complex mechanism that is guided not only by RBP but also by the highly versatile structural or sequence directed ncRNAs.

1.1.1 Non-coding RNAs

One of the long-standing principles of molecular biology is that the genome sequence provides all the information needed for messenger RNAs transcription, which act as a template for protein translation. The genome indeed contains all the information but most of the transcripts are in fact not translated into proteins. Thanks to the ENCODE project (Encyclopedia of DNA Elements) and the growing Next Generation Sequencing technologies, we know that the genome is pervasively transcribed, but only around 1.5% generates protein coding RNA molecules (Alexander et al., 2010).

A much larger portion of transcribed molecules does not code for proteins but they nevertheless play an important role in all cellular pathways. These RNAs are the ncRNAs and they can be classified in different categories based on their length and function.

Based on their function, ncRNAs can be classified in structural and regulatory

ncRNAs (Fig. 1.1).

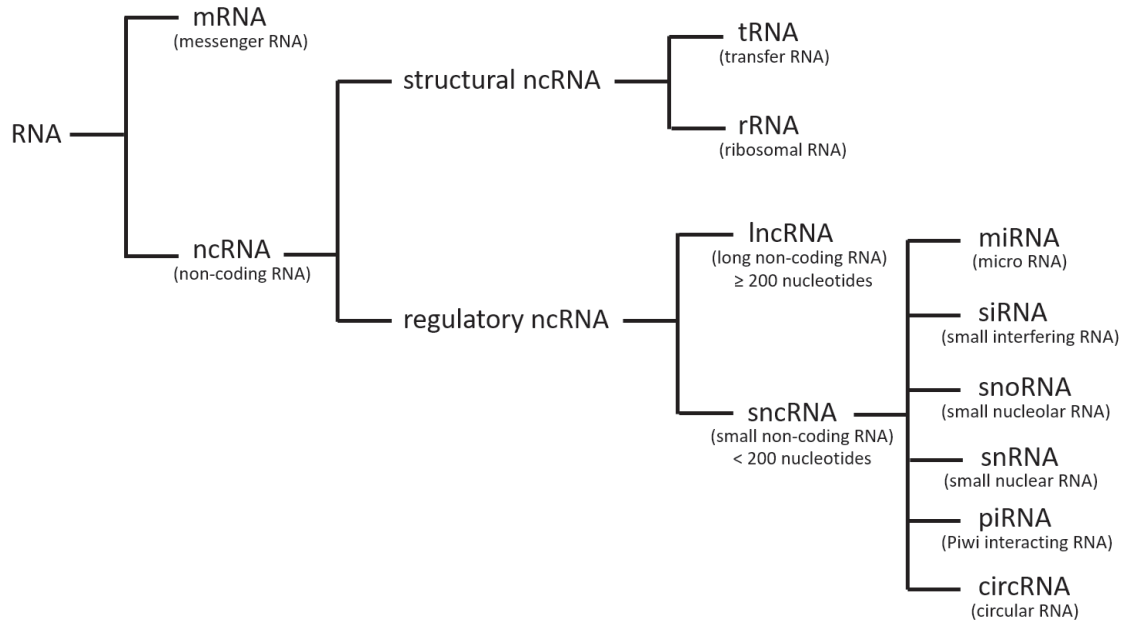


FIGURE 1.1: Schematic representation of RNA classification

ncRNAs with a structural function include transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) involved in translation. Regulatory ncRNAs can be divided in small ncRNAs (sncRNAs) or long ncRNAs (lncRNAs) (Hombach & Kretz, 2016). LncRNAs are >200 nt in length and they often function as a scaffold playing crucial roles in a variety of biological and pathological processes such as gene regulation, mRNA stability, cancer progression, X-chromosome inactivation and differentiation (Geisler & Coller, 2013; Pagani et al., 2013). On the other hand, sncRNAs are between 18 nt and 200 nt in length and they can be further classified into: endogenous short-interfering RNAs (siRNAs, 20–25 nt), microRNAs (miRNAs, 20-22 nt) and Piwi-interacting RNAs (piRNAs, 26-31 nt) which act as key regulators of gene expression in many different cellular pathways, and small nuclear RNAs (snRNAs) or small nucleolar RNAs (snoRNAs) involved in splicing events or nucleolytic processing of rRNAs (Ambros, 2004; Y. Liu et al., 2019; C. Yan et al., 2019). More recently, circular RNAs (circRNAs) have been identified as a family of stable long or short ncRNAs that can regulate gene expression by sponging RBPs and miRNAs, therefore affecting transcription, mRNA tournover and translation (X. Zhang et al., 2018). In recent years, regulatory ncRNAs such as lncRNAs and miRNAs gained attention for their ability to modulate gene expression in

physiological conditions and for most biological processes such as homeostasis, development, viral response, apoptosis and cancer.

small ncRNAs

The two major classes of small regulatory RNAs are miRNAs and siRNAs. Both are around 22nts long but they differ in their biogenesis and target recognition. These sncRNAs are implicated in gene expression regulation and they play a role in virtually all biological pathways (Q. Liu & Paroo, 2010; Romano et al., 2017). Both siRNAs and miRNAs exert their function by guiding effectors RBPs to their target RNAs via base pairing, a mechanism called RNA interference (RNAi) (Saw & Song, 2020).

siRNAs originate from long dsRNA molecules and they are typically highly complementary to their targets, while miRNAs derive from RNA hairpins that usually contain several mismatched bases. miRNAs are not highly complementary to their targets but they require a 6-8 nts "seed" region to be perfectly matched, while other positions might contribute weakly to target specificity (V. N. Kim et al., 2009).
siRNAs were first described in plants and invertebrates as part of an antiviral defense mechanism (Hamilton & Baulcombe, 1999). In particular, the double stranded RNA (dsRNA) intermediate from viral replication or from transposable elements is recognized and cleaved by DICER endonucleases to form siRNAs that are subsequently loaded onto an RNA-induced silencing complex (RISC). In this complex, one strand of the siRNA is used as a fully complementary guide to bind and cleave the complementary viral RNA genome. However, there are relatively few reports of endogenous mammalian siRNAs and their implication in antiviral immunity (Cullen, 2006).

On the other hand, the complete RNAi machinery in mammals is mostly dedicated to the production of miRNAs that, upon association with the RISC, target endogenous mRNAs for repression. RISC-mediated gene silencing guided by miRNAs is one of the most extensive mechanism for gene regulation in mammals,

being involved in most biological pathways. In fact, deregulation in the miRNA pathway has been widely linked to a variety of diseases such as cancer or genetic disorders (Ha & Kim, 2014). Therefore, it's very important to deeply understand the biogenesis of miRNAs, its regulation and miRNAs function.

In the next sections of this introduction the state of the art of the miRNA biogenesis pathway and their regulatory function will be thoroughly detailed.

1.2 microRNAs

1.2.1 Overview on miRNAs

microRNAs are a class of short regulatory ncRNAs of around 20-22 nt in length (Ha & Kim, 2014). miRNA genes represent one of the most abundant gene families and they can be widely found in animals, plants, viruses and protists (Pasquinelli et al., 2000). According to the latest version of miRNA database (miRbase V22), 2883 mature miRNAs were annotated in *homo sapiens*, 495 mature miRNAs in *Drosophila melanogaster* and 450 mature miRNAs in *Caenorhabditis elegans* (Kozomara et al., 2019).

The first miRNA characterized in *Caenorhabditis elegans*, *lin-4*, was linked to the control of developmental timinig, mediating post-transcriptional control of *lin-14* gene by sequence complementarity with its 3' UTR (R. C. Lee et al., 1993). Following the identification of lin-4, many more miRNAs have been characterized in different species and, interestingly, multiple miRNA *loci* with related sequences were found throughout the genomes analyzed. For instance one of the first miRNAs indentified in *Caenorhabditis elegans*, *let-7* (Reinhart et al., 2000), was subsequently found to be conserved across a wide range of animal phyla, including *homo sapiens* (Pasquinelli et al., 2000). This highlighted the potential impact of miRNAs as a conserved gene regulation mechanism and suggested their evolution by genome duplication (Berezikov, 2011). Genome duplication could also explain the hairpin secondary structure common to all miRNAs in all species. In fact, the central requirement for a miRNA is a primary transcript harboring a hairpin-like

secondary structure, essential for its processing and maturation to a functional miRNA (see next section for more details).

Canonical miRNA genes are transcribed to produce primary miRNAs (pri-miRNAs) by RNA PolII to produce m^7G-capped pri-miRNA containing one or more mismatched hairpin structures. In *homo sapiens*, around half of canonical miRNAs are expressed from introns of non coding or protein coding genes and they are called intronic miRNAs. Some other miRNAs can localize in intergenic regions (intergenic miRNAs) and a small group of miRNAs can localize in exons (exonic miRNAs). According to their localization, miRNAs expression could be dependent on the transcription of the host gene or they could have their own promoter (Monteys et al., 2010; Ozsolak et al., 2008) (see section 1.3.2 for details).

The pri-miRNA hairpin can be recognized and processed to form a mature miRNA by sequential cleavage by two RNAse III proteins, DROSHA in the nucleus and DICER in the cytoplasm. After their processing by RNAse III proteins, miRNAs maturation concludes with the formation of a short RNA duplex which is loaded into Argonaute family protein (AGO)- containing RISC and it is able to induce RNA silencing by binding to specific mRNAs. Contrary to siRNAs, miRNAs do not fully perfectly base-pair with their target mRNAs. Instead, only a short sequence of around 6-8 nts, called "seed region", is perfectly complementary to the target RNA (Ha & Kim, 2014).

However, in the last decades, studies have identified several non-canonical miRNAs that undergo alternative biogenesis pathways. For instance, some miRNAs, transcribed by RNA PolIII, can result from the processing of transfer RNAs (tRNAs) or small nuclear RNAs (snRNAs) processing. Another class of intronic miRNAs, that can be found in very small introns (miRtrons), can bypass DROSHA processing exploiting intron splicing mechanisms to produce a mature miRNA (Ruby et al., 2007). Similarly, some pre-miRNAs can be directly transcribed by RNA PolII which generates the pre-miRNA m^7G-capped 5′ end and 3′ end following transcription termination. Therefore 5′m^7G-capped pre-miRNAs bypass

nuclear processing and are directly exported to the cytoplasm for DICER-mediated cleavage (Xie et al., 2013). On the other hand, the biogenesis of another class of non-canonical miRNAs, the simtrons, does not require processing by DICER but is able to perform gene silencing by binding to different AGO proteins (Havens et al., 2012) (see section 1.6 for details).

Independently of their biogenesis, it has been shown that miRNAs can regulate gene expression at different levels, for example influencing transcription, mRNA stability and translation (Di Leva et al., 2014; Long et al., 2018; S. Yan & Jiao, 2016). To exploit this function, miRNAs act as a guide by base-pairing with its target RNA, precisely with the miRNA recognition elements (MRE), usually at the 3' UTR of target mRNAs. According to the degree of base pairing, a miRNA binding to a target coding transcript promotes its deadenylation, degradation or translation inhibition, in any case leading to post-transcriptional gene silencing.

By binding to their target mRNAs, miRNAs are predicted to regulate the expression of more than 60% of mammalian genes, thus contributing to a plethora of physiological and pathological events. For instance, miR-122, the most abundant tissue-specific miRNA expressed in the liver, is involved in the reduction of plasma cholesterol levels by regulating genes responsible for cholesterol biosynthesis (Hsu et al., 2012). In addition, miR-122 has been shown to be essential for Hepatitis C virus replication interacting with the 5' untranslated region of the viral genome (Jopling et al., 2005). Finally, low levels of miR-122 have been strongly linked to liver cancer where the overexpression of miR-122 promotes apoptosis of hepatocellular carcinoma cells (Wei et al., 2019).

The regulation of target mRNAs by miRNAs is an interesting and well-studied mechanism that is drawing attention because of its evident implication in homeostasis, development, cancer, apoptosis and many more physiological and pathological events. However, the regulation of miRNAs themselves is still poorly understood.

1.2.2 Evolution of miRNAs

Small RNAs emerged early in the evolutionary chain as an ancient system in bacteria, archea and eukarya. They evolved primarily as a defense mechanism against foreign nucleic acids and transposons, involving key RBPs whose helicase and RNAse III domains are conserved from prokaryotes to eukaryotes (Obbard et al., 2009; Shabalina & Koonin, 2008). The coordination of sncRNAs and effector RBPs led to the establishment of the RNAi pathway acting as a defense mechanism against viruses, transposable elements (Shabalina & Koonin, 2008). However, this system seems to have been substituted with innate immunity and interferon response in vertebrates (Cullen, 2006). In fact, in animals, there are very few examples of siRNA-mediated RNAi mechanism. One exception to this observation is represented by virus-infected mammalian embryonic stem cells that do not have a functional interferon system (Berkhout, 2018; Cullen, 2006). On the other hand, miRNA-mediated mRNA silencing in mammals represents a key feature in the control of gene expression.

Interestingly, the miRNA repertoire drastically expanded simultaneously with animal evolution and its complexity (Heimberg et al., 2008). Moreover, it seems that more recent and conserved miRNAs arose in bursts coinciding with major organism developments, while organisms that diverged early in the evolutionary line tend to have few and non-conserved miRNAs (Cock et al., 2017). For instance, it's estimated that around 46% of human miRNAs are primate-specific and 14% are human-specific (V. D. Patel & Capra, 2017).

Although thousands of miRNAs have been identified, their origin and their expansion mechanism is still largely unknown. However, it is likely that highly conserved secondary structures play an important role in miRNA evolution by selecting miRNAs transcripts that form mismatched hairpins. Three different models are proposed for the origin and expansion of miRNAs:

- transposable elements (TEs) insertion and derivatization. There are many examples of miRNAs derived from different types of TEs. For instance, it was

shown that the miR-548 family derived from *Made1* TEs, while miR-28 and miR-95 represent an example of LINE-2 derived miRNAs (Piriyapongsa et al., 2007). Two types of TEs derived miRNAs were identified: type I, which derived from inverted TE sequences, and type II that only partially overlap with TEs. However, how miRNAs originate from TEs remains largely unknown. Recently, it was speculated that type II TE-derived miRNAs could originate from the head or tail of TEs or they could evolve by partially losing TE sequence from the whole TE (Qin et al., 2015).

- spontaneous evolution of random sequences. This model is based on the understanding that the human genome encodes for millions of potential hairpins, mostly found in transcripts (Bentwich et al., 2005). Therefore, functional targeting miRNAs could emerge from a pool of hairpin transcripts via random mutations (Felippes et al., 2008).

- inverted gene duplication. This model suggests that miRNAs with new binding and targeting specificities could evolve by duplication of an existing functional miRNA followed by random mutation in the target recognition site. This theory could explain the fact that approximately 37% of human miRNA genes are found in paralogous clusters (Shabalina & Koonin, 2008).

The most supported mechanism for miRNA expansion is miRNA gene duplication followed by sequence diversification, which leads to the organization of miRNAs into clusters, often transcribed together (O'Brien et al., 2018). Sequence diversification can occur by mutations in the mature sequence region, which could lead to recognition of different target mRNAs, or mutation elsewhere in the miRNA transcript, which leads to change in the hairpin structure and potentially modifies its binding to the effector RBPs (Berezikov, 2011). However, mutations in the mature sequence region and, in particular, in the "seed region" at the 5' end of mature miRNAs, are less frequent since they determine the specific mRNA target repertoire and they are therefore under strong selective pressure. (Berezikov, 2011). Interestingly, between 20 and 40% of human miRNAs are organized into clusters (Chaulk et al., 2015). miRNA clusters are usually divided into two groups: homologous and heterologous clusters. Homologous clusters are composed by

miRNAs of the same family, recognized by the conservation of similar seed regions and often targeting mRNAs with similar biological functions (Abbott et al., 2005; Luo et al., 2018). On the other hand, heterologous clusters are composed by miRNAs that are not part of the same family and can be involved in the regulation of mRNAs in different biological pathways.

Introns might also contribute to miRNA expansion. Indeed, studies suggests that introns, just like miRNA genes, were evolutionary gained in bursts, together with organism complexity, thanks to the insertion of TEs between preexisting nucleosomes (Worden et al., 2009). This mechanism would support the evidence that almost half of human miRNAs, are located in intronic regions and have the same orientation as their host-gene. Furthermore, more recent and conserved miRNAs are more often localized in introns, compared to ancient miRNAs (Campo-Paysaa et al., 2011). Therefore, introns seems to provide support for the evolution of hairpin structures.
Importantly, given the extent and the significance of miRNA-mediated gene regulation, the expansion of novel miRNAs has to emerge concurrently with mechanisms regulating their expression to avoid undesired effects.

Transcriptional control of novel miRNAs

Novel miRNAs, especially if spontaneously evolved from random sequences, have the potential to non-specifically target many mRNAs. It's thought that the transcription of novel miRNAs is placed under a strict control mechanism to repress its expression. Because novel miRNAs, without target specificity, could have deleterious off-target effects, this mechanism would avoid adverse influence on the fitness of the organism (Shabalina & Koonin, 2008). With time and the appearance of random mutations within the miRNA transcript, the off-target sites would be purged and the expression of the miRNA could increase without deleterious consequences, eventually acquiring physiologically important regulatory functions.
This is in agreement with the observation that the estimated evolutionary age of a miRNA corralates with its expression level (Berezikov et al., 2006).

1.3 miRNA biogenesis

1.3.1 Overview on miRNA biogenesis

The majority of miRNA genes are transcribed by RNA polymerase II (Pol II) in the nucleus which produces long primary transcripts (pri-miRNA) containing hairpin structures where the miRNAs sequences are embedded (Fig. 1.2).

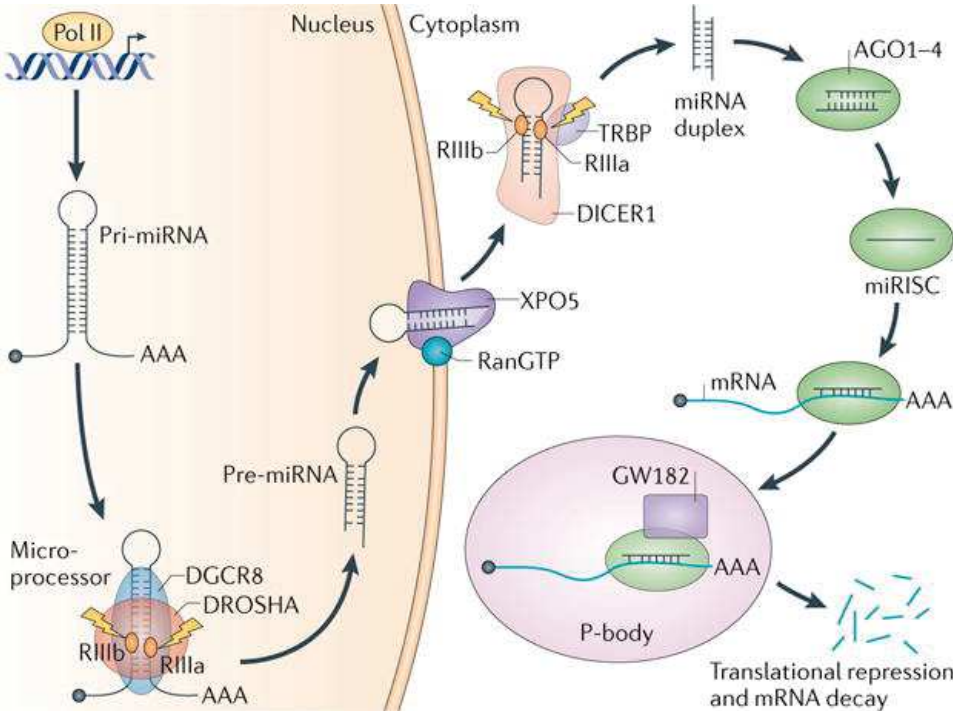


FIGURE 1.2: Overview of the miRNA biogenesis pathway, from S. Lin and Gregory, 2015

Pri-miRNAs are capped, spliced and polyadenylated and often they can produce a single miRNA or contain clusters of two or more miRNAs that are processed from a common primary transcript. The canonical pri-miRNA transcript is typically over 1 kb and it includes an hairpin structure composed of a 33-35 bps long stem with a terminal loop that is flanked by single stranded RNA fragments at both 5′ and 3′ sides.

Following transcription, the pri-miRNA is cleaved by the nuclear RNAse III-type endonuclease DROSHA which crops the stem-loop releasing a small hairpin of around 65 nt in length, the precursor miRNA (pre-miRNA). DROSHA acts in synergy with factors, such as DiGeorge syndrome critical region 8 (DGCR8), which

provides additional RNA-binding activity, and together they form a functional complex called Microprocessor. The Microprocessor cleaves the hairpin 13 bps before and 11 bps after the "basal junction" (the boundary between single strand RNAs (ssRNAs) and the stem), producing a characteristic 2 nucleotide 3′ overhang that is essential for further processing (Han et al., 2006).

After this nuclear process, the pre-miRNA is exported to the cytoplasm by Exportin-5. In the cytoplasm, the RNAse III-type endonuclease DICER binds the pre-miRNA with a preference for the 2-nt 3′ overhang and cleaves the dsRNA stem at a fixed distance from the 3′ end (3′-counting rule) or from the 5′ end (5′-counting rule), producing a 22 nt miRNA duplex.

Once the small RNA duplex is generated by DICER, it is loaded onto an AGO protein, forming the RISC. Only one strand of the duplex, the guide strand, will be part of the mature RISC; the other strand, the passenger strand, will be removed and degraded.
The mature RISC is able to bind target mRNA through the "seed" region of the miRNA guide strand. This region is located at the 5′ end of the miRNA and it is 6-8 base pairs long, therefore each microRNA has the potential to target multiple different mRNAs (Lewis et al., 2005). The fate of the target mRNA depends on the degree of base-pairing complementarity between the mRNA molecule and the "seed" region, which can lead mRNA decay or translational repression (Fig. 1.2).

1.3.2 miRNAs transcription

As previously mentioned in 1.2, miRNAs can be classified, based on their genomic context, in intragenic (exonic or intronic) or intergenic. According to their localization, miRNAs transcription can occur dependently on the host gene transcription or independently using a miRNA-specific promoter (Fig. 1.3).

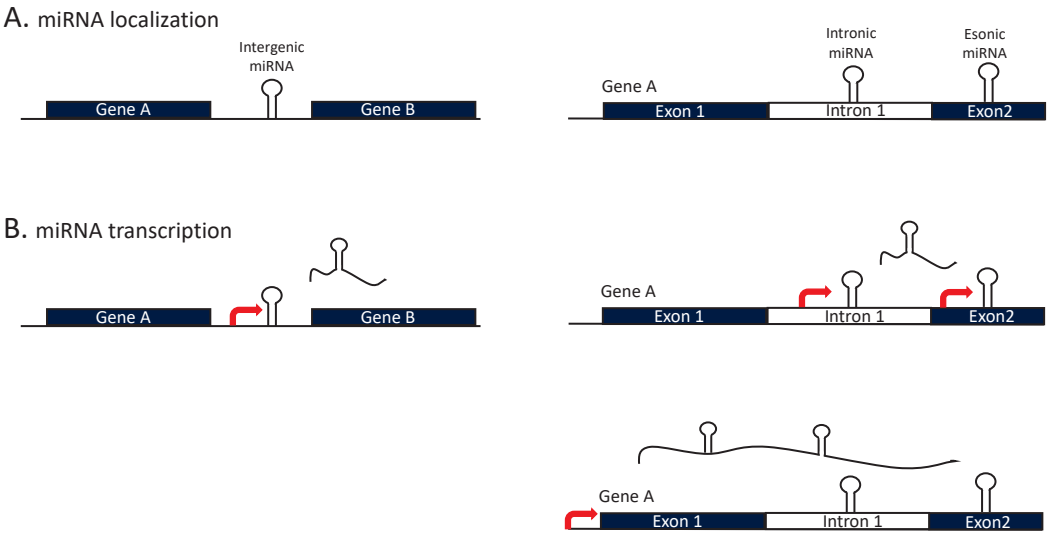


FIGURE 1.3: miRNA genome localization and their transcription.
A) Scheme of miRNA localization in the genome
B) Transcription of miRNAs and their promoters (red arrows).

However, because of the transient nature of pri-miRNAs, the study of their transcription is complex and the exact locations of promoters have not been mapped yet for most miRNA genes. Traditional transcriptome profiling techniques, such as RNA-seq, fail to sufficiently capture pri-miRNAs before their rapid cleavage by DROSHA. To overcome this technical barrier to large-scale detection of miRNA transcription starting sites (TSS), techniques to detect transient RNA transcription on a genome-wide scale are mainly used (Chien et al., 2011).

Intergenic miRNAs are transcribed from their own promoter which is recognized by PolII. Recently, many intergenic miRNA TSSs were identified using techniques such as CAGE-seq (Cap-Analysis of Gene Expression coupled with sequencing) or ChIP-seq (Chromatin Immunoprecipitation followed by sequencing) for marks

associated with actively transcribed genes, such as H3K4me3, H3ac and PolII occupancy (Chien et al., 2011; Hua et al., 2016).

Until recently, it was thought that intronic miRNAs and their host genes were always under the influence of common transcriptional regulatory mechanisms and that the mature miRNA was produced by processing from the host gene primary transcript. Interestingly, recent studies show a discordant expression of several intronic miRNAs and their host genes, suggesting that also intronic miRNAs might have their own promoter and therefore could be transcribed independently from their host gene (Q. Liu et al., 2017; Marsico et al., 2013; Ramalingam et al., 2014). Genome-wide analyses identified PolII intron-resident promoters for around 30% of intronic miRNAs, based on the presence of RNA PolII regulatory sequences (Monteys et al., 2010).

A more recent study suggests widespread independent transcription of intronic miRNAs showing that it's relatively common for intronic miRNAs and their host genes to display discordant expression. Intronic miRNAs showing discordant expression with their host genes were further analyzed to detect a putative miRNA TSS based on DNAse I hypersensitivity sites sequencing, H3K4me3 histone modification and other marks associated with promoter activity. miRNAs transcription was then classified into dependent or independent from host gene transcription according to the distance between the miRNA TSS and the host gene TSS (Hua et al., 2016; Steiman-Shimony et al., 2018; Y. Sun et al., 2017). However, the associations found in these studies are often not supported by conclusive experimental validations. One intronic-miRNA and host gene pair was recently validated using a 5' RACE (Rapid Amplification of cDNA Ends) assay. It was shown that the intronic pri-miR-26b was transcribed using a proximal promoter and its TSS aligned to the first intron of its host gene Carboxy-Terminal Domain Small Phosphatase 1 (CTDSP1) (Y. Sun et al., 2017).

On the other hand, discordant expression of intronic miRNA and its host gene could also be explained by other mechanisms. Several studies showed how

alternative splicing contributed to uncoupled expression of intronic miRNAs and host genes (Ramalingam et al., 2014). Moreover, the possibility should not be excluded that intronic miRNAs transcription might be controlled by both the host gene TSS and a downstream TSS at the same time in one cell type or that the TSS usage could vary in different cell types and/or conditions (Steiman-Shimony et al., 2018; X.-Q. Yang et al., 2019).

Exonic miRNAs and intronic miRNAs can be canonically transcribed by PolII using their host gene promoter as a single transcriptional unit (pri-miRNA) which can subsequently be processed by DROSHA. In this case, transcription is controlled by RNA Pol II-associated transcription factors, such as MYC and p53, and epigenetic regulators that contribute to miRNA gene regulation through DNA methylation and histone modification. Several studies support a p53-dependent regulation of miRNAs transcription (Cortez et al., 2016; Su et al., 2014; Y. Zhang et al., 2016). For instance, it was shown that miR-27b and miR-455, two intronic miRNAs involved in cancer cell quiescence, are both regulated by p53 which increases transcriptional activity in response to serum starvation. In addition, it was demonstrated that miR-27b and miR-455, found in human Chromosome 9 Open Reading Frame 3 (C9ORF3) and human collagen alpha-1 (XXVII) chain (COL27A1) genes respectively, are co-transcribed with their host genes and co-regulated by p53 during serum starvation. Indeed, knockdown (KD) of C9ORF3 and COL27A1 inhibited the expression of miR-27b and miR-455; while, following serum starvation, both host genes and miRNAs were upregulated (La et al., 2018). Another study (Heller et al., 2018) extensively analyzed the methylation pattern of some miRNA genes known to be involved in Non-small-cell Lung Carcinoma (NSCLC). For example, miR-1179 has tumor growth suppressor properties mediated by repression of its main target, Cyclin E1. In NSCLC, the expression of miR-1179 is affected by its methylation status which, in turn, can modulate cell viability and cell proliferation.

On the other hand, some examples show that intronic miRNAs transcribed independently of their host gene, can also be subjected to distinct transcriptional

control systems. For instance, transcription of the intronic miRNA, miR-26b, is suppressed by the transcription factor Myc, whereas the latter does not control transcription of the host gene, CTDSP1 (Y. Sun et al., 2017).

Despite the examples cited and the recent progress in the detection of transient pri-miRNA transcripts, miRNA transcription and its regulation is still largely unknown.

1.3.3 Microprocessor cleavage

The nuclear processing of the transcribed pri-miRNAs starts with the cleavage by a complex called Microprocessor. The product of this process is a pre-miRNA, a 65 nt long hairpin with the characteristic 2 nts overhang at the 3' end (Denli et al., 2004; Nguyen et al., 2015). As aforementioned, the Microprocessor is an heterotrimetric complex of around 364 kDa composed by DROSHA, an RNAse III enzyme, and a DGCR8 dimer (Quick-Cleveland et al., 2014).

Microprocessor structure

DROSHA is a predominantly nuclear protein of around 159 kDa. It consists of N-terminal proline-rich (P-rich) and arginine/serine-rich (R/S-rich) domains, a central domain (CED), and two RNase III domains (RIIIDa and RIIIDb) followed by a dsRNA-binding domain (dsRBD) (Fig. 1.4).

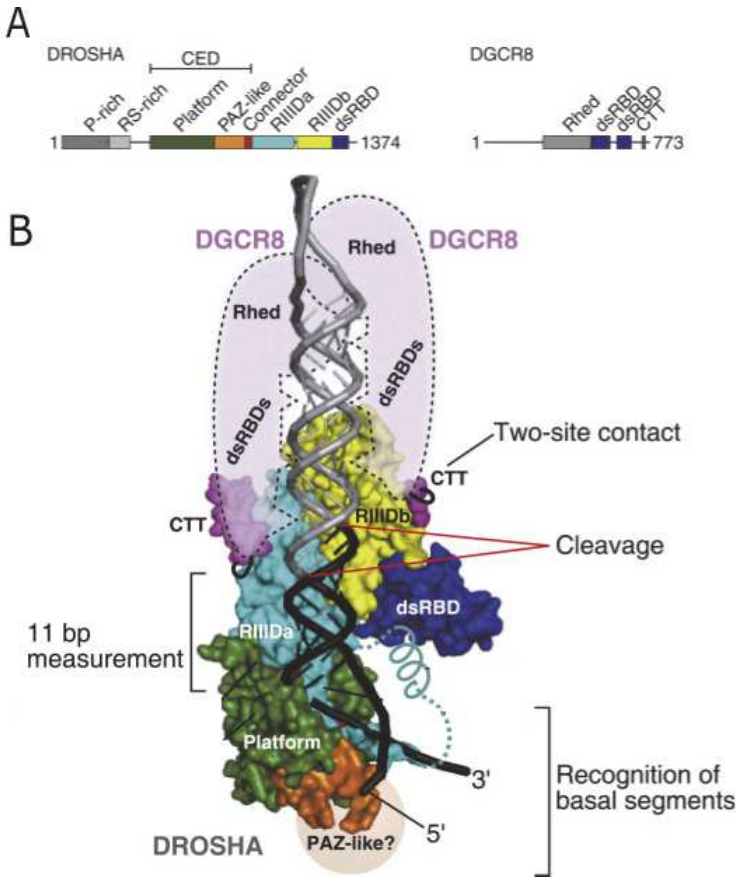


FIGURE 1.4: Microprocessor protein structure.
A) Domain composition of human DROSHA and DGCR8.
B) Model of human Microprocessor bound to a pri-miRNA.
Adapted from Kwon et al., 2016

The N-terminal domains are important for the processing of pri-miRNAs, while the first and second RIIIDs interact with each other forming an intramolecular dimer to cut the 3′ (RIIIDa) and 5′ (RIIIDb) strands of the stem (Kwon et al., 2016). The central domain is functionally uncharacterized but it appears to be essential for DROSHA function (Han et al., 2004). The two RIIIDs domains are able to recognize and cut the 3′ and the 5′ strands independently of each other suggesting that DROSHA is able to orient itself on the pri-miRNA. It was proposed that the CED might help DROSHA orienting the RIIIDs domains at the 3′ and 5′ strands of the

stem or, alternatively, this function could be assured by DGCR8 (Han et al., 2004).

DGCR8 is a nuclear protein of around 86 kDa that consists of an N-terminal region including the nuclear localization signal (NLS), the central RNA-binding heme domain (Rhed), two dsRBDs, and the C-terminal tail region (CTT). It was shown that depletion of the CTT region inhibits the interaction of DGCR8 with DROSHA suggesting that the CTT domain is essential for binding and stabilizing DROSHA binding to DGCR8 (Yeom et al., 2006).

Recognition and cleavage of the pri-miRNA stem loop by DROSHA and DGCR8

The mechanism by which Microprocessor chooses cleavage sites on pri-miRNAs is still unclear. It is now accepted that during the cleavage of the pri-miRNA to form the pre-miRNA, DROSHA functions as a "molecular ruler" measuring 11 and 13 bps distance from the 5′ and 3′ ends of the basal ssRNA-dsRNA junction, determining the cleavage site and cutting the pri-miRNA stem-loop (Han et al., 2006). However, for different pri-miRNAs, different mechanisms have been observed and models proposed. It was shown that extending the lower stem or altering the upper stem caused alternative DROSHA processing for a subset of pri-miRNAs, suggesting that both lower and upper ssRNA-dsRNA junctions contribute to determine DROSHA cleavage site (H. Ma et al., 2013). Therefore, the existing models do not fully explain cleavage site choice and further investigation of this mechanism is needed.

The binding of DROSHA to its cofactor, DGCR8, is also essential for accurate and efficient processing of the pri-miRNA, promoting Microprocessor activity. In particular, the dsRBDs of DGCR8 enhance RNA binding affinity of the Microprocessor while the dimerization of DGCR8 ensures the accuracy of the process by inhibiting non-canonical processing, also called "unproductive processing" (Nguyen et al., 2015).

Both DROSHA and DGCR8 are thought to contact the pri-miRNA hairpin by recognizing specific primary sequence features (Fig. 1.5).

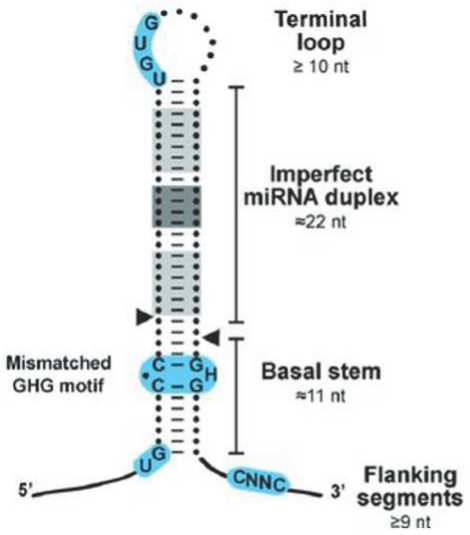


FIGURE 1.5: Sequence motifs on the pri-miRNAs that are recognized by effecter RBPs. Adapted from Creugny et al., 2018

In particular, DROSHA binds the basal ssRNA-dsRNA junction of the pri-miRNA hairpin recognizing an UG motif (basal UG motif) at position -14 and -13 of the Microprocessor cleavage site (Auyeung et al., 2013). Another primary sequence important for the recognition of pri-miRNA stem-loops is the CNNC motif (flanking CNNC motif), found in most pri-miRNAs and located on the 3′ ssRNA, around 17 bps downstream the Microprocessor cleavage site. The CNNC motif was recently found to be recognized by the splicing factor SRSF3 (SRp20) and to recruit DROSHA to the basal junction of the pri-miRNA. Moreover, it was shown that the distance of the CNNC domain from the basal junction influences DROSHA cutting site through its binding to SRSF3 (Auyeung et al., 2013; K. Kim et al., 2018).

On the apical region, an UGU conserved motif (apical UGU motif) can be recognized by DGCR8 through its Rhed domain (Auyeung et al., 2013; Michlewski & Cáceres, 2019). However, evidence suggests that this primary sequence motif doesn't influence cleavage site determination but instead enhances fidelity and efficiency of the processing (Kwon et al., 2019).

Among all human pri-miRNAs, 79% contain at least one of the aforementioned motifs (Creugny et al., 2018; K. Kim et al., 2018). Interestingly, it was found that

these motifs are not enriched in *C. elegans* pri-miRNAs but, when artificially added, they significantly increased pri-miRNA processing in mammalian cells. Thus, although very conserved among mammals, these primary-sequence motifs differ in *nematodes* (Auyeung et al., 2013).

More recently, a mismatched GHG motif (mGHG) was identified in the 3p arm and frequently mismatched at position 8 from the basal junction (Fang & Bartel, 2015). This motif is recognized by the dsRBD of DROSHA and it seems to be crucial for precise processing, at least for a subset of pri-miRNAs. Furthermore, sequence alignment studies show that mGHG is conserved from *sea anemones* to humans, implying its relevance for DROSHA processing. Therefore, the mGHG motif evolved much earlier that the basal and the apical motifs and plays a critical and highly conserved role in the determination of cleavage sites (Kwon et al., 2019).

The aforementioned primary sequences features have been found to be important for the correct and precise cleavage of pri-miRNAs and also enhance processing. Some exceptions to this observation are, for example, pri-miR-16-1 and pri-miR-30a. pri-miR-16-1 does not contain the UGU motif in the apical junction, suggesting that its basal junction is strong enough to ensure precise processing. On the other hand, the processing of pri-miR-30a, lacking the UG motif in the basal junction, has been shown to be strongly dependent on the apical DGCR8 binding motif. This indicates that sequence motifs and structural modules contribute to varying degrees to the recognition of each pri-miRNA by Microprocessor (Nguyen et al., 2015).

The ability of these numerous sequence motifs to guide cleavage by Microprocessor and their conservation confirms that the efficiency and the single-nucleotide precision of this process are crucial for miRNA abundance and targeting specificity. Because miRNAs bind their target mRNA specifically through their "seed region", located close to the 5′ end of the miRNA, a change in the cleavage site, even of one nucleotide only, can completely modify the target specificity of a miRNA (Chiang et al., 2010; H. Wu et al., 2009). Interestingly, it was shown that some pri-miRNAs can be processed by DROSHA at multiple sites, producing more than one

pre-miRNA from a single hairpin and, therefore, different mature miRNA isoforms (isomiR) (H. Wu et al., 2009). Different isomiRs, produced from the same precursor, can have different seed sequence and therefore regulate different mRNAs (Tan et al., 2014). As detected by sequencing, IsomiRs can vary by one or multiple nucleotides at the 5′ or 3′ ends (R. C. Lee et al., 1993) and their concentration was shown to fluctuate considerably after drug exposure (Giuliani et al., 2020), infection (Rotival et al., 2020), development (Fernandez-Valverde et al., 2010; Yuan et al., 2015), diasease and cancer (Liang et al., 2020; Martí et al., 2010), suggesting that their biogenesis might be actively regulated.

Microprocessor non-canonical substrates

Recently, DROSHA cleavage sites have been annotated on a genomic scale using formaldehyde crosslinking, immunoprecipitation, and sequencing (fCLIP-seq) (B. Kim et al., 2017). This approach allowed the discovery of DROSHA cleavage sites in an unbiased manner revealing numerous non-canonical substrates. Among them, several hairpin *loci* were found to be target of DROSHA cleavage but not yet annotated as miRNAs nor conserved among species (B. Kim et al., 2017). Thus, it is possible that these hairpin *loci* may be evolving miRNA genes with limited or no biological activity (section 1.2.2). On the other hand, these hairpins may represent *cis* acting elements for the regulation of mRNAs. It was already reported that DROSHA can destabilize mRNAs hosting hairpin structures in a miRNA-independent fashion by cleaving hairpins embedded in exons or UTRs (Han et al., 2009; Triboulet et al., 2009).

DROSHA miRNA-independent regulation was also shown for mRNAs of factors important for neurogenesis and myelopoiesis (Johanson et al., 2015; Knuckles et al., 2012; Rolando et al., 2016). For instance, the transcription factor Neurogenin 2 (Ngn2) mRNA, an important regulator of mammalian neurogenesis, contains evolutionary conserved hairpin structures that can be cleaved by DROSHA leading to transcript destabilization. This mechanism is particularly important for mouse development since it provides a molecular inhibition of Ngn2 accumulation, preventing abnormal neurogenesis (Knuckles et al., 2012).

Interestingly, Drosha was found to have a role in splicing regulation and alternative splicing modulation for different mRNAs, including its own transcript (Havens et al., 2014; D. Lee et al., 2017). It was shown that DROSHA is involved in the regulation of alternatively spliced exon 5 of eIF4H gene (Havens et al., 2014). Exon 5 has a predicted hairpin structure that can be recognized and cleaved by the Microprocessor. However, binding of DROSHA can enhance splicing of exon 5 in a cleavage-independent manner.

Lastly, DROSHA miRNA-independent regulation was linked to the viral response. Upon RNA virus infection, DROSHA is exported to the cytoplasm and cleaves viral genomic RNA, thereby inhibiting viral replication (Shapiro et al., 2014).
Considering the plenitude of hairpin structures located throughout the whole transcriptome, it's not surprising that DROSHA is able to recognize and cleave them, contributing in a miRNA-independent fashion to the regulation of transcript stability.

pri-miRNA cleavage is co-transcriptional

Cleavage of pri-miRNAs by Microprocessor was shown to occur co-transcriptionally on chromatin-associated pri-miRNAs (Y.-K. Kim & Kim, 2007; Morlando et al., 2008). It was demonstrated that coupling of PolII-transcription and pri-miRNA processing generally enhances the kinetics and efficiency of pri-miRNA processing. Moreover, in presence of flanking exons, miRNA biogenesis is enhanced, probably due to prolonged retention at the site of transcription (Pawlicki & Steitz, 2008; Yin et al., 2015). In fact, endogenous pri-miRNAs that are efficiently processed are enriched in chromatin-associated nuclear fractions, while pri-miRNAs inefficiently processed at transcription sites are recognized as defective transcripts and accumulate in splicing factor SC35 (SC35) domains, as a result of quality control mechanisms (Pawlicki & Steitz, 2008). Interestingly, many proteins have been shown to inhibit or promote the co-transcriptional association of the Microprocessor to the nascent pri-miRNA, thereby promoting or inhibiting efficient pri-miRNA processing.

For instance, fused in sarcoma/translocated in liposarcoma (FUS/TSL), a ubiquitously expressed RNA-binding protein, is recruited to chromatin at sites of pri-miRNA transcription and facilitates co-transcriptional DROSHA recruitment (Morlando et al., 2012).
Another important factor associated with co-transcriptional pri-miRNA processing is the RNA-binding protein LIN-28. During early development in embryonic stem cells, LIN-28 was found to inhibit the precocious maturation of let-7 miRNA by binding to its nascent transcript and inhibiting its processing by DROSHA (Newman et al., 2008).

It is plausible that proteins involved in Microprocessor activity regulation at this stage might also possess DNA-binding activity in order to interact with chromatin factors and promote Microprocessor association with nascent pri-miRNAs. For example, the heterochromatin protein 1 binding protein 3 (HP1BP3) is a histone H1-related protein that was recently associated with co-transcriptional miRNA processing (H. Liu et al., 2016). HP1BP3 was found to bind chromatin DNA, nascent pri-miRNA transcript and to specifically associate with the Microprocessor in order to retain pri-miRNA on chromatin and promote miRNA biogenesis.

pri-miRNAs transcription and host pre-mRNAs splicing The coupling of transcription and pri-miRNA processing is particularly interesting for intronic miRNAs where the cleavage of the pri-miRNAs could potentially interfere with the splicing of the pre-mRNA, also occurring co-transcriptionally. Notably, many splicing factors have been linked to the Microprocessor activity or have been found in the same complex (Treiber et al., 2017). For instance, the splicing regulator KHSRP, a key mediator of mRNA decay, binds the terminal loop of a subset of pri-miRNAs and favours processing by promoting their interaction with DROSHA (Trabucchi et al., 2009). The coordination of pri-miRNA processing and splicing may also be facilitated by the presence of DROSHA and DGCR8 in the suprasplaceosome, (Agranat-Tamir et al., 2014) a complex of around 21 MDa composed of 4 native spliceosome and the nascent pre-mRNA (Azubel et al., 2004).

In a subset of pre-miRNAs investigated, pri-miRNA processing was shown to occur before splicing, exposing the cleaved intron to rapid exonucleolytic degradation. This was also supported by the observation that the chromatin at these specific introns interacts with exosome and the 5′-3′ exonuclease XRN2 (Morlando et al., 2008). On the contrary, another study revealed that an artificial intron containing a pre-miRNA was spliced before pre-miRNA processing (Agranat-Tamir et al., 2014).

Concerning splicing efficiency, some studies found that the cropping of intronic pre-miRNAs by the Microprocessor did not significantly affect the production and maturation of the host mRNA (Janas et al., 2011; Y.-K. Kim & Kim, 2007; Pianigiani et al., 2018). Indeed, KD of DROSHA did not diminish splicing although introns without pre-miRNAs were spliced more rapidly than introns containing pre-miRNAs (Kataoka et al., 2009). This observation could be explained by a model where the pri-miRNA processing occurs between the splicing commitment step and the excision step. Therefore, despite the discontinuity of the intron that follows the cropping by the Microprocessor, splicing catalysis could still occur efficiently because the exons have already been paired and tethered to each other. However, in more recent studies, it was found that splicing inhibition increases miRNAs processing while KD of DROSHA increases the level of novel spliced isoforms, suggesting a competition between pri-miRNA processing and splicing (Agranat-Tamir et al., 2014; Barbier et al., 2018).

Data supporting both hypotheses could suggest that the interplay between pri-miRNA processing and splicing might be dependent on some other features that have not been taken into consideration yet. For instance, one could speculate that the competition or cooperation between Microprocessor and spliceosome could somehow depend on the length of the intron, on its position or on the specific protein composition of both complexes. Needless to say, more evidence needs to be gathered in order elucidate this mechanism.

Transcriptional and post-transcriptional regulation of DROSHA and DGCR8

In order to efficiently process miRNAs and control their abundance, Microprocessor activity must be tightly regulated. Multiple mechanisms that control the expression level, the activity and the specificity of DROSHA and DGCR8 have been recently discovered. DROSHA N-terminal region contains proline-rich (P-rich) and arginine/serine-rich (RS-rich) domains that provide a regulatory platform for DROSHA post-translational modifications. For instance, glycogen synthase kinase 3β (GSK3β) phosphorylates Serine300 and Serine302, located in the RS-rich domain which facilitates nuclear localization of DROSHA (Tang et al., 2011). Serine300 residue can also be a target of p38 mitogen-activated protein kinase (MAPK) under stress conditions, inducing DROSHA nuclear export and degradation (Q. Yang et al., 2015).

On the other hand, critical lysine residues of DGCR8 dsRBDs can be deacetylated by Histone deacetylase 1 (HDAC1). The deacetylation of these residues increases the affinity of DGCR8 for primary miRNA transcripts and enhances miRNA processing (Wada et al., 2012). Another example of post-translational modification that regulates Microprocessor activity is the phosphorylation of the residue Tyr^{267} by the kinase ABL. This residue is found immediately N-terminal to the DGCR8 Rhed domain, involved in the interaction with DROSHA. This modification has been shown to enhance the processing of some pri-miRNAs, including pri-miR-34c, probably by promoting the recruitment of DROSHA to the miRNA transcript (Tu et al., 2015).

Interestingly, examples of autoregulation between DROSHA and DGCR8 have been also shown. In the absence of DGCR8, the level of DROSHA protein is downregulated, suggesting that DGCR8 might stabilize DROSHA through protein-protein interactions. On the other hand, DROSHA was shown to suppress the expression of DGCR8 by targeting two hairpin structures found in the 5'UTR and in the coding region of DGCR8 mRNA (Han et al., 2009). Furthermore, this cross-regulatory loop that enables the homeostatic maintenance of Microprocessor

activity, is evolutionary conserved.

In addition, DROSHA pre-mRNA hosts a pri-miRNA-like hairpin, between exon 7 and intron 7, that can be a target of the Microprocessor itself (D. Lee et al., 2017). The binding of the Microprocessor to this structural feature in DROSHA pre-mRNA was shown to compete with the splicing machinery promoting exon 7 skipping in human cells. This mechanism impairs DROSHA protein nuclear localization or facilitates nuclear export. In fact, DROSHA proteins generated from exon 7-skipped isoforms were exclusively nuclear, while those from exon 7-containing isoforms were also present in the cytoplasm (Link et al., 2016). This alternative spliced isoform of DROSHA seems to have emerged very recently since it only occurs in placental mammals and was shown to be predominantly abundant in the brain (D. Lee et al., 2017; Link et al., 2016).

RNA-binding proteins involved in Microprocessor activity

Although the minimal components of Microprocessor are DROSHA and its co-factor DGCR8, many other proteins have been implicated in the pri-miRNA processing step of the miRNAs biogenesis at different stages. As previously discussed, a number of proteins are able to enhance or inhibit Microprocessor activity by modulating its co-transcriptional recruitment on chromatin DNA. Many other trans-acting factors, especially proteins with RNA-binding activity, have been shown to positively or negatively influence Microprocessor activity. For instance, RBP hnRNP A1 is able to enhance pri-miR-18a processing by binding to its terminal loop and driving an allosteric destabilization of the base-pairing in the stem that promotes more efficient DROSHA binding (Kooshapur et al., 2018; Michlewski et al., 2010). This is particularly interesting because it suggests that the tertiary structure and the sequence of the stem and the loop regions are both important for DROSHA binding and Microprocessor activity.

In addition to hnRNP A1, many other RBPs can recognize the terminal loop of miRNA precursors and positively or negatively influence their biogenesis (Castilla-Llorente et al., 2013; Michlewski et al., 2008). In fact, proteins that were

previously cited in this text, such as LIN28 and FUS/TSL, exploit their function by binding to the terminal loop of their target miRNAs precursors. Another interesting factor is the KH-type splicing regulatory protein (KSRP). KSRP was shown to directly interact with G-rich stretches in the terminal loops of a subset of precursor miRNAs, optimizing the position/recruitment of both Microprocessor and the downstream DICER miRNA processing complex (Trabucchi et al., 2009).
Importantly, these trans-acting factors have different effects on Microprocessor activity depending on different miRNA precursors or cell types. For instance, in human cells that lack LIN28 expression, hhnRNP A1 can act as a negative regulator for some miRNAs, such as let-7a, by competing with the activator protein KSRP for the binding to its terminal loop (Michlewski & Cáceres, 2010).

Helicases, such as the DEAD box RNA helicases p72/DDX17, can also act as cofactors for the Microprocessor activity (Dardenne et al., 2014). DDX17 is able to promote the maturation of a subset of miRNA precursors by remodelling the 3' flanking region of the pri-miRNA, therefore enhancing DROSHA processing (Ngo et al., 2019; Remenyi et al., 2016). Similarly, p68/DDX5 was found to be a component of the Microprocessor complex for the efficient processing of some pri-miRNAs (Gregory et al., 2004; Hong et al., 2013).

More recently, other classes of proteins have been linked to the recognition of the pri-miRNA stem by DROSHA and DGCR8. For instance, Hemin, a ferric ion-containing porphyrin, can be incorporated in DGCR8, which modifies its conformation and enhances the specific interaction between the DGCR8 dimer and the apical UGU motif on the pri-miRNA (Nguyen et al., 2018; Partin et al., 2017). In addition, this was shown to be particularly important for the differential expression of certain pri-miRNAs as Hemin-binding positively influences only UGU-containing pri-miRNAs (Nguyen et al., 2018).
More factors associated with Microprocessor activity are continuing to be identified, suggesting the existence of a diversified, but yet largely unexplored, regulatory interactome surrounding DROSHA and DGCR8 (Davis et al., 2008; J. Song et al., 2020). More importantly, different cofactors already identified have

the ability to specifically modulate the processing of a set of miRNAs, sometimes leading to divergent outcomes, depending on the cell type or the cellular environment. These observations reveal a much more convoluted mechanism than previously thought, leading to a plethora of different outputs to finely tune miRNA expression and maturation. Nevertheless, the exact mechanism of pri-miRNA processing regulation is still largely unknown.

1.3.4 Stability of pri-miRNA hairpins

Hairpins are very stable and recurrent structures that can be widely found in the human transcriptome (Bentwich et al., 2005). Several features contribute to the stability of hairpin structures, such as the length of the duplex and the loop, the number and the type of mismatches included in the stem (Groebe & Uhlenbeck, 1988; Meroueh & Chow, 1999). Mismatches found in the duplex are destabilizing and increase the free energy of the hairpin structure. Moreover, mismatches with at least one purine base (e.g. C/A, A/C, G/A) show higher stability (and lower free energy) than two pyrimidine bases (e.g. U/U). On the other hand, RNA containing G/U and U/G mismatches are the most stable compared to all the other mismatches (Meroueh & Chow, 1999).

Concerning structural features, hairpin loops seem to be a target for many RBPs. It was found that approximately 14% of all human pri-miRNAs have highly conserved loops that are often bound my RBPs, which influences their stability and association with Microprocessor (see previous section for details). For example, pri-miR-7-1, whose mature form, miR-7, that is enriched in brain and pancreatic tissue, has a conserved terminal loop that was shown to be bound by Musashi homolog 2 (MSI2) and Hu antigen R (HuR) proteins. In particular, in non-neural cells, although its host mRNA is expressed, the cleavage of pri-miR-7-1 is inhibited by MSI2/HuR binding to its terminal loop, leading to tissue-specific control of its expression (N. R. Choudhury et al., 2013). In the same study, it was shown that, upon binding of MSI2/HuR to the loop, the stem of pri-miR-7-1 becomes more stable and rigid, inhibiting Microprocessor activity.

On the other hand, hnRNP A1, already cited in this text (section 1.3.3), enhances

pri-miR-18 processing by binding to the hairpin and inducing a conformational change in its terminal loop that leads to a more relaxed stem and efficient DROSHA cleavage (Michlewski et al., 2008).

Aside from structural characteristics, also sequence motifs recognized by RBPs can increase or diminish the hairpin stability, depending on the RBP involved. For instance, LIN28, already found to bind terminal loop of pri-miRNAs influencing their processing (section 1.3.3), was shown to recognize a GGAG motif in several human pri-miRNAs (Rybak et al., 2008; Urbanek-Trzeciak et al., 2018).
Finally, enzymes part of the adenosine/cytidine deaminase family, such as ADAR1 and ADAR2 (Adenosine Deaminase acting on RNA), can trigger single nucleotide changes in the pri-miRNA sequence, impacting their stability. Indeed, nucleotide modifications (A-to-I or C-to-U), typical of the RNA editing mechanism, especially occurring in the stem region of pri-miRNAs, have the potential to change their secondary structure and therefore influence their stability, maturation and, importantly, their target repertoire (Correia de Sousa et al., 2019). For example, the edited form of miR-376, a miRNA expressed in gliomas, lacks its ability to bind its target mRNAs, leading to higher invasive capacity in this type of cancer (Y. Choudhury et al., 2012).

1.3.5 pre-miRNAs export

After the first step of miRNAs maturation, carried out by the Microprocessor, the pre-miRNAs formed are exported to the cytoplasm through the Exportin 5 (XPO5). XPO5 is one of the karyopherin β family of nucleocytoplasmic transport factors, which mediate protein export to the cytoplasm by recognizing specific transport signals in the cargo proteins. In particular, XPO5 was shown to shuttle small RNAs and RBPs, such as NF90 and ADAR1, to the cytoplasm, interacting with transport signals in their dsRBDs (Fritz et al., 2009; Gwizdek et al., 2003). Interestingly, the export of RBP was shown to be facilitated by their binding to RNA which increases their affinity to XPO5 (Fritz et al., 2009).

Several members of this family have been shown to shuttle ncRNAs; XPO5 was

initially shown to transport the adenovirus VA1 RNA (Gwizdek et al., 2003). Because of the "mini-helix" structure and the presence of two nucleotides 3′ overhang, characteristic of the adenovirus VA1 RNA and other ncRNAs, it was hypothesized and subsequently shown that XPO5 could mediate miRNAs nuclear export. XPO5 nuclear cargo binding requires the GTP-bound form of the Ran GTPase in the nucleus: pre-miRNAs are transported by Ran/GTP/XPO5 to the cytoplasm through the nuclear pore complex, released upon GTP hydrolysis to GDP, thereby allowing free XPO5 to subsequently return to the nuclear compartment (Lund et al., 2004; Yi et al., 2003).

The structure of XPO5:RanGTP:pre-miRNA reported by Okada et al., using X-ray technology, shows that the 2 nt 3′ overhang of the pre-miRNA functions as a transport signal, allowing XPO5 to recognize and bind its cargo pre-miRNA (Fig. 1.6).

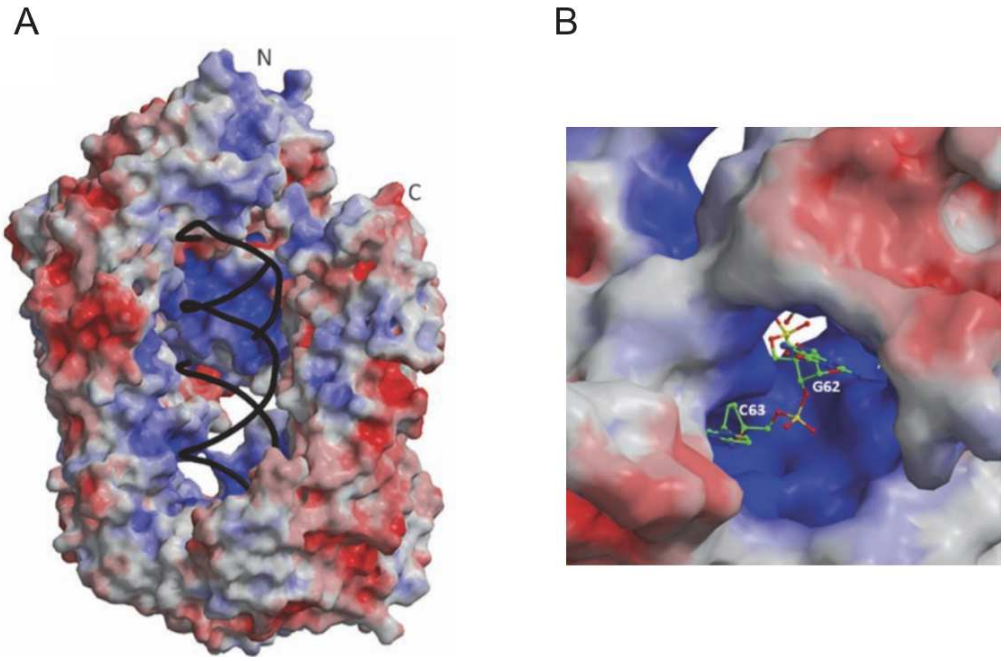


FIGURE 1.6: Structure of XPO5 complexed with a pre-miRNA. A) The electrostatic surface potentials are represented in a color gradient from red to blue while the black backbone represents the stem of the pre-miRNA. The inner surface of the baseball mitt-like structure is positively charged. B) The inside of the tunnel is electrostatically basic and shows many hydrophilic interactions with the 2-nt 3 overhang structure. Adapted from Okada et al., 2009

The 2 nt 3′ overhang is bound to a tunnel-like structure of XPO5 and stabilized

through a number of hydrogen bonds and salt bridges. On the other hand, the pre-miRNA stem is bound by a baseball mitt-like structure of XPO5. The inner surface of the tunnel and the baseball mitt-like structure is rich in positive charges, which stabilizes the interaction with the negatively charged RNA. Moreover, since both 3′ and 5′ ends of the pre-miRNA are shielded by XPO5, it also prevents the degradation of pre-miRNAs by exonucleases. Interestingly, all pre-miRNA/XPO5 interactions were found to be mediated by atoms of the sugar-phosphate backbone, suggesting that pre-miRNAs recognition is independent on their sequence (Okada et al., 2009). On the contrary, XPO5 binding to pre-miRNAs was shown to be sensitive to the RNA structure, particularly around the basal elements. In fact, the pre-miRNA 3′ overhang was important but not essential for XPO5 binding. Blunt-ends, 5 or 1 nt 3′ overhangs with or without a 2 nts 5′ overhang were able to bind XPO5, while only a 5′ overhang compromised XPO5 binding (Zeng & Cullen, 2004).

Pre-miRNAs export is a rate-limiting step of miRNA biogenesis and therefore an efficient cytoplasm shuttling is critical for miRNA maturation and miRNA-dependent gene regulation. As expected, downregulation of XPO5 inhibits miRNAs maturation and induces precursor degradation (Lund et al., 2004). In some types of colon cancers, the C-terminal region of XPO5 was found to be truncated. Because this region is essential for the formation of XPO5/Ran-GTP/pre-miRNA complex, the truncated XPO5 can not bind pre-miRNA, leading to nuclear accumulation of pre-miRNAs (Melo et al., 2010). On the other hand, overexpression of XPO5 enhances expression of mature miRNAs by enhancing cytoplasmic shuttling and increasing pre-miRNA availability for subsequent maturation steps. As a result, overexpression of XPO5 enhances miRNA-mediated inhibition of target mRNAs (Yi et al., 2005). Therefore, XPO5 regulation is essential for miRNAs biogenesis.

It was shown that ERK activation globally reduces miRNA expression by phosphorylating XPO5 and reducing its ability to bind pre-miRNAs (H.-L. Sun et al., 2016). In the same study, they showed that in hepatocellular carcinoma (HCC), phosphorilation of XPO5 by ERK inhibits miR-122 maturation. Loss of

miR-122 results in reduced chemosensitivity and confers resistance to the drug Taxol (Xu et al., 2011). Recently, XPO5 was also found to bind closely clustered polycistronic pri-miRNAs in the nucleus in a GTP-independent manner. In particular, XPO5 is able to bind pre-miR-19a hairpin in a GTP-dependent manner and pri-miR-19a basal stem region independently of Ran-GTP. By binding outside of pre-miR-19a hairpin region, XPO5 was found to promote Microprocessor cleavage *in vitro* (J. Wang et al., 2020).

Some non-canonical pre-miRNAs, such as the Microprocessor-independent m^7G capped pre-miRNAs, have been shown to utilize a different export pathway. For instance, in quiescent cells, m^7G-capped pre-miRNAs can be exported by Exportin-1 (XPO1 or CRM1) (Martinez et al., 2017). CRM1 is able to export a variety of RNA subtypes, by recruiting a phosphorylated adapter RNA export protein (PHAX) as an anchor for pre-miRNA binding (Xie et al., 2013). However, the mechanism of PHAX binding to pre-miRNAs is still unclear. It's possible that PHAX recognizes the cap structure at the 5′ end of the pre-miRNAs, since the interaction between PHAX and cap-binding-complex (CBC) was already shown for snRNAs nuclear export by CRM1 (Boulon et al., 2004).
Thus pre-miRNA export is a crucial part of the miRNA biogenesis pathway, representing a rate limiting step that is subjected to several regulatory mechanisms.

1.3.6 DICER cleavage

After export to the cytoplasm, pre-miRNAs maturation proceeds with DICER cleavage to produce small RNA duplexes, the miRNAs. Like DROSHA, DICER is a member of the RNAse III family and it specifically hydrolyzes phosphodiester bonds found in double-stranded RNAs, generating 21-25 nucleotides long products having 5′-phosphate, 3′-hydroxyl and 2-nt 3′-overhangs (Macrae et al., 2006).
Evolutionary studies indicate that animals developed two Dicer genes: Dicer-1, responsible for pre-miRNA recognition and cleavage to produce miRNAs, and Dicer-2 that cleaved dsRNA to generate siRNAs (Y. S. Lee et al., 2004). Dicer-2 has an important role in the response to viral infections by cleaving viral dsRNAs, but

it was subsequently lost from vertebrates that developed complex immune systems (Mukherjee et al., 2013).

DICER structure

Mammalian DICER is a 220 kDa multidomain protein that is produced as a single isoform from Dicer gene (Ciechanowska et al., 2021). A crystal of full-length mammalian DICER is not yet available and, so far, its structure has been inferred from numerous biochemical studies (E. Ma et al., 2008), comparison with DICER homologous proteins (Macrae et al., 2006) and crystallographic structure analyses of separate domains (Du et al., 2008; J.-B. Ma et al., 2004; Takeshita et al., 2007).

Thanks to a recent cryogenic Electron Microscopy (cryo-EM) analysis, we know that the overall shape of DICER resembles the letter "L" and it is conventionally divided in three parts: head, body and base (Fig. 1.7) (Taylor et al., 2013).

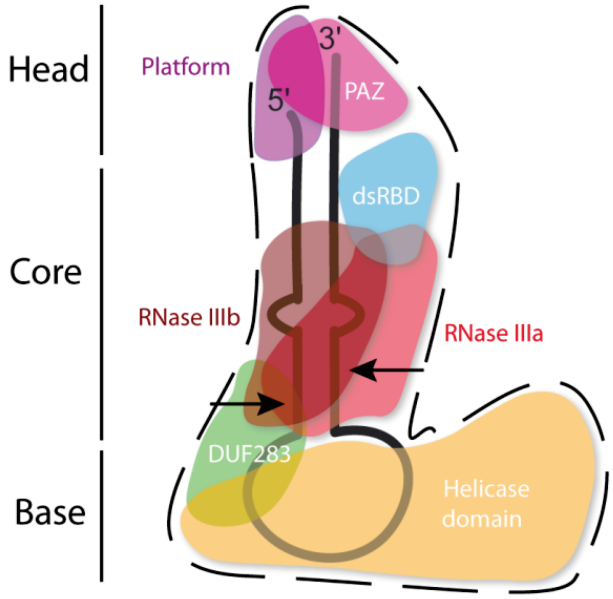


FIGURE 1.7: Schematic representation of DICER structure in complex with a pre-miRNA. The arrows indicate the cleavage sites. From Ciechanowska et al., 2021

The main functional domains are: PAZ (Piwi-Argonaute-Zwille) domain, located in the head of the protein, the RNAse IIIa and IIIb, located in the body together with the dsRNA binding domain, the helicase domain and a domain of unknown function (DUF283) which constitutes the base (Svobodova et al., 2016).

The PAZ domain is a dsRNA-terminus binding motif, common to Argonaute and DICER proteins (K. S. Yan et al., 2003). It contains a highly conserved 3′ overhang binding pocket to specifically anchor the two-nucleotide 3′ overhang of the RNA duplexes and it seems to be involved in the transfer of cleaved RNA to the PAZ domain of Argonaute in the RISC (J.-B. Ma et al., 2004; Macrae et al., 2006) . In particular, it was shown that modification of the 3′ overhang structure of the target RNA duplexes leads to significant reduction of DICER cleavage and RNAi efficiency (Elbashir, Lendeckel, et al., 2001; Harborth et al., 2003). By contrast, many other studies have reported that 3′ modifications of RNA duplexes do not interfere with RNAi, suggesting that PAZ might not be the only RNA duplex recognition domain (Amarzguioui et al., 2003; Elbashir, Martinez, et al., 2001). Specifically, DICER PAZ domain contains a large extended loop enriched in basic amino acids that dramatically changes the electrostatic potential of the 3′ overhang binding pocket and could be involved in the transfer of cleaved RNA to other proteins (Macrae et al., 2006).

An extension of the PAZ domain, the Platform-PAZ-connector helix cassette, contains a phosphate-binding pocket (5′ pocket) positioned in close proximity to the 3′ pocket to ensure simultaneous accomodation of the substrate 5′- and 3′-ends (J.-E. Park et al., 2011; Tian et al., 2014). The helicase domain of human DICER contains a DExD/H-box motif capable of unwinding RNA or DNA duplexes. While this activity is usually ATP-dependent, this has not been proven for vertebrate DICERs containing DExD/H-box motifs. The helicase domain can also function as a platform for dsRBPs binding (Hansen et al., 2019).

The DUF283 domain has still no characterized function but its structure adopts an alpha-beta-beta-beta-alpha (α-β-β-β-α) topology, typical of canonical dsRBDs (Dlakić, 2006; Z. Liu et al., 2018). Nevertheless, a dsRNA-binding activity has not been demonstrated for this domain, although it was reported that DUF283 is capable of binding single-stranded nucleic acids *in vitro* and enhancing hybridization between short RNAs and their complementary regions in longer

RNAs (Kurzynska-Kokorniak et al., 2016).

The two RNAse IIIa and IIIb domains, found in the body of DICER, form an intramolecular dimer and constitute its catalytic core. The mechanism of cleavage and substrate recognition has, so far, been studied only in *E. coli*. In particular, it was shown that the RNAse IIIa domain cleaves the pre-miRNA 3′ arm that contains the 3′-overhangs, while the RNAse IIIb domain cleaves the arm containing the 5′ phosphate (H. Zhang et al., 2004). The role of the dsRBD in human DICER is still under study. In fact, it doesn't seem essential for DICER action in standard assays but, when isolated, it binds strongly to dsRNA (Wostenberg et al., 2012) and enables substrate processing by DICER in absence of the PAZ domain (E. Ma et al., 2012). This suggests that the dsRBD plays an auxiliary role in dsRNA binding and cleavage or it might recognize alternative substrates that are not recognized by the PAZ domain.

Recognition and cleavage of the pre-miRNA by DICER

As aforementioned, mammalian DICER cleaves the Microprocessor products, the pre-miRNAs, generating RNA duplexes of 21 to 28 nucleotides.
Crystallographic structure studies suggest that DICER anchors both ends of its pre-miRNA substrates (Fig. 1.8).
It was reported that the region surrounding the PAZ domain contains a phosphate 5′-binding pocket in close proximity to the already known 2-nt 3′-overhang binding site (Tian et al., 2014). Interestingly, the distance between the 5′ and 3′-binding pockets is approximately 20 Å, which corresponds to the distance between the 5′ and the 3′-ends of a dsRNA with a 2-nt 3′-overhang. Therefore, the distance between the two binding pockets could explain the strong preference of DICER for dsRNAs containing a 2-nt 3′-overhang (J.-E. Park et al., 2011).

It is possible that the 5′ and 3′-ends of the substrate dsRNAs are simultaneously accommodated in the phosphate and the 3′ pockets, respectively (J.-E. Park et al., 2011; Tian et al., 2014). The binding of both ends seems to be extremely important for the precise and uniform processing of the substrates. Indeed, mutations of the

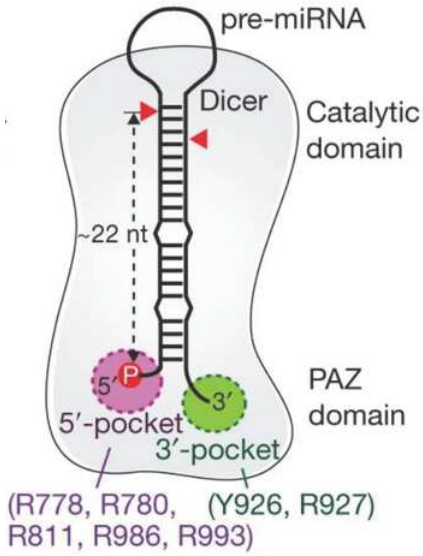


FIGURE 1.8: Model for DICER processing of pre-miRNA. The arrows indicate the cleavage sites. Adapted from J.-E. Park et al., 2011

5′-binding pocket lead to reduced accuracy and efficiency of human DICER, while 3′-pocket mutations reduce its processing activity (Tian et al., 2014). Moreover, mutations in the 3′ or 5′-ends of DICER substrates shows that, according to the substrate used, the product length might vary. Tests on several pre-miRNAs showed that DICER acts as a molecular ruler, cleaving the dsRNA substrate of a length that is calculated from the 3′-binding pocket (3′ counting rule) or the 5′-binding pocket (5′ counting rule) or both to the RNAse III domains (J.-E. Park et al., 2011). The relative contribution of the 5′ and 3′ ends to the length of the cleaved product seems to vary among pre-miRNAs. It was shown that pre-miRNAs following the 3′ counting rule are relatively stable at the stem termini, whereas the 5′ counting rule seems to occur when the pre-miRNA end is thermodynamically unstable (mismatch, G–U, or A–U pair) (Soifer et al., 2008).

The processing efficiency of thermodynamically unstable duplexes, such as pre-miRNAs harboring mismatches in the stem region, seems to depend on DICER helicase domain, since mutation in this domain results in poor processing of these substrates (Soifer et al., 2008). In addition, it was speculated that the helicase domain of DICER could be involved in the rearrangement of the enzyme to release the processed miRNA and handing it off to the RISC. Importantly, another level of

recognition of the substrate is carried out by the helicase domain itself which binds the hairpin RNA loop (E. Ma et al., 2012).

The cleavage center of DICER is formed by the intramolecular dimerization of the two RNAse III domains that independently cleave the two strands of the substrate (Takeshita et al., 2007). In fact, mutation of single RNAse III domain prevents cleavage of one RNA strand but has no effect on the second strand (H. Zhang et al., 2004). Moreover, dicing assays indicated that DICER processing rate varies with different substrates (Chakravarthy et al., 2010). Cleavage catalysis speed depends not only on the substrate binding affinity but it could depend on an inefficient release of the product or a sub-optimal loading due to particular structural properties of the substrates. Considering only pre-miRNAs, many variations in structure such as number of mismatches or loop size naturally exist and they might lead to considerable differences in their recognition and processing by DICER (Chakravarthy et al., 2010).

Other proteins involved in DICER processing

It was shown that DICER processes different pre-miRNAs at distinct rates, according to their structural properties (Chakravarthy et al., 2010). In fact, the interaction of DICER with dsRNA-binding proteins, such as the human *trans*-activation (TAR) RNA-binding protein (TRBP), can modulate its processing efficiency, its accuracy or fine tuning the length of mature miRNAs (Chendrimada et al., 2005). However, the effect of TRBP on DICER activity is still very controversial.

It is possible that the binding of TRBP to DICER induces a conformational change that influences DICER substrate specificity and kinetics. For instance, a dicing assay performed on pre-miR-31 or a chimeric substrate showed that association of TRBP with DICER inhibited the processing of these substrates by several fold. Results from the same study suggest that the effect of TRBP on the substrates might be mediated by their stem-loop conformation, since insertion of mismatches in the stem-loops of the targets reversed the effect observed previously (H. Y. Lee &

Doudna, 2012; Zhu et al., 2018). Along the same lines, TRBP binding to Dicer is able to alter isomiR production, contributing to guide strand selection and mRNA target silencing activity (H. Y. Lee & Doudna, 2012; Zhu et al., 2018). In fact, according to the substrate sequence and stem flanking nucleotides, mismatches in the stem-loop of a DICER/TRBP-bound pre-miRNA can change the length of the duplex included between the PAZ domain and the RNAse III domains, leading to isomiR production (Zhu et al., 2018).

Therefore, TRBP is not a general enhancer of DICER activity, but it can have divergent functions according to the substrate structure or even cellular context and signalling. For example, phosphorylation of TRBP, mediated by the mitogen activated protein kinase (MAPK) Erk, has been suggested to extend DICER-TRBP complex half-life and enhance growth promoting miRNA processing activity (Paroo et al., 2009). On a different note, it was recently shown that TRBP is involved in immunodeficiency virus type 1 (HIV-1) translation and replication by binding to the TAR RNA and inhibiting the 5'UTR TAR RNA-mediated translation suppression of TAR RNA while DICER processes HIV-1 TAR RNA to generate TAR miRNAs (Komori et al., 2020).

Importantly, TRBP was found to regulate DICER products loading onto Ago during RISC assembly in *Drosophila*, while in mammals its implication in this process is rather controversial. It has been demonstrated that DICER and TRBP regulate RISC formation and strand selection (Daniels et al., 2009; Noland & Doudna, 2013). In contrast, a more recent publication, showed that TRBP KO altered DICER-processing sites in a subset of pre-miRNAs but had no effect on miRNA abundance, DICER stability or AGO2 loading (Betancur & Tomari, 2012).

Another protein involved in DICER processing is the dsRNA-binding Protein Kinase R (PKR) activating protein (PACT), a cellular protein activator of the kinase PKR and human paralog of TRBP (R. C. Patel & Sen, 1998). Coimmunoprecipitation analysis showed that AGO2 interacts with PACT similarly to TRBP but the precise role of PACT in DICER processing has not yet been fully defined (Y. Lee et al.,

2006). PACT was shown to associate with DICER and it seems to have a function similar to TRBP. Their conserved sequence and domain organization suggest that TRBP and PACT may be functionally redundant although they appear to act on different miRNAs (Wilson et al., 2015). They are both conserved among vertebrates supporting the observation that PACT and TRBP function at the same step in the small RNA biogenesis pathway, but on different substrates (Heyam et al., 2015). Moreover, crystallographic structure of DICER-TRBP complex showed that binding of PACT and TRBP to DICER is mutually exclusive although it cannot be excluded that PACT and TRBP interact with each other and only one of them binds to DICER (Heyam et al., 2015; Kok et al., 2007).

DICER together with TRBP or PACT are often essential to achieve optimal strand selection, but they show different and relative contribution to this process (Wilson et al., 2015). In addition, double KO of TRBP and PACT indicated that the two proteins do not functionally compensate for each other (Y. Kim et al., 2014).
Interestingly, variations of TRBP or PACT expression and their effect in miRNA biogenesis is linked to several types of cancers (Caramuta et al., 2013; Sand et al., 2012). Overall, the role of TRBP and PACT in DICER-mediate miRNA processing is still controversial and more evidence is needed to draw conclusions.

Non-canonical DICER substrates and functions

In addition to pre-miRNAs, DICER can process endogenous and exogenous long dsRNAs although these have only a partial experimental support in mammals. Exogenous sources of substrate dsRNA of DICER include viral dsRNA, documented in plants, *Drosophila* and *C. elegans* (Vance & Vaucheret, 2001; X.-H. Wang et al., 2006; Wilkins et al., 2005). For instance, viral infection in adult flies triggers *Drosophila* Dicer protein to act as a sensor of the viral double-stranded RNA and produce siRNAs to activate the siRNA-dependent RISC (siRISC), leading to viral RNA silencing (X.-H. Wang et al., 2006). However, this pathway seems to have a very limited functionality in mammals (Stein et al., 2005; Svoboda, 2014).

Although DICER is mainly localized in the cytoplasm where it exerts its function in

the miRNAs biogenesis pathway, it was found in the nucleus of mammalian cells, suggesting a nuclear role as well. Nuclear DICER was found to process non canonical damage-induced dsRNAs that promote DNA repair, taking part in the DNA-damage response (DDR) (Burger et al., 2017; Francia et al., 2012).
Importantly, DICER was shown to be implicated in the biogenesis of tRNA-derived fragments (tRFs), a class of small RNAs derived from tRNAs in plants and *Drosophila* (Cole et al., 2009).

DICER regulation

In addition to the RBPs implicated in DICER processing that modulate miRNA maturation, DICER itself can be modified to regulate miRNA processing. For instance, it was shown that the DExD/H-box domain, included in the helicase domain of DICER, inhibits the rate of perfect-duplex dsRNA cleavage. Interestingly, this domain functions as a protein interaction domain with TRBP. It is plausible that the binding of TRBP to DICER leads to a conformational change that enhances its processing activity, suggesting that the DExD/H-domain acts as a structural switch that controls DICER activity (E. Ma et al., 2008).

DICER localization can also strongly impact miRNA processing. For example, studies indicated that phosphorylation of DICER by ERK affects oocyte development in *C. Elegans* and humans. In particular, ERK can phosphorylate two conserved residues of DICER RNAseIII and dsRNA-binding domains resulting in its nuclear translocation (Drake et al., 2014). DNA damage was also shown to induce phosphorylation and translocation of DICER into the nucleus where, as previously mentioned, it is involved in the DDR by generating non canonical small RNAs (Burger et al., 2017).
A more recent study suggests that regulation of DICER is important for specific miRNA production in a sub-cellular localization. It was shown that local synaptic activity can stimulate DICER-dependent maturation of pre-miRNAs leading to their maturation and a decrease in their target mRNA translation (Sambandan et al., 2017).

1.4 RISC-mediated gene silencing

The catalytic core of the human RISC is represented by AGO proteins. The Argonaute family can be divided into two subfamilies: Ago subfamily (composed by AGO1, AGO2, AGO3 and AGO4) and PIWI subfamily (which includes HIWI1, HIWI2, HIWI3 and HIWI4). PIWI subfamily is exclusively expressed in germ-line cells, while AGO family is widely expressed in different tissues. AGO proteins are extremely conserved among species from bacteria to archaea and eukaryotes.

In humans, while the four AGO proteins are expressed, AGO2 is essential for RISC-guided mRNA silencing by binding to miRNA or siRNA and mediating translational repression or cleavage and decay of the target mRNAs. AGO2 was thought to be the only AGO protein having mRNA slicing activity. However, more recently, AGO3 has been shown to slice target mRNAs when loaded with a subset of miRNAs (M. S. Park et al., 2017). In the majority of the cases, human AGO2 will silence target transcripts by inhibiting their translation or inducing mRNA decay (Flores et al., 2014).

1.4.1 AGO2 structure

The full length crystal structure of human AGO2 was determined in 2012 (Schirle & MacRae, 2012). It was observed that AGO2 has a bi-lobed structure formed by four core domains: N Domain, PAZ, MID and PIWI domains (Fig. 1.9).

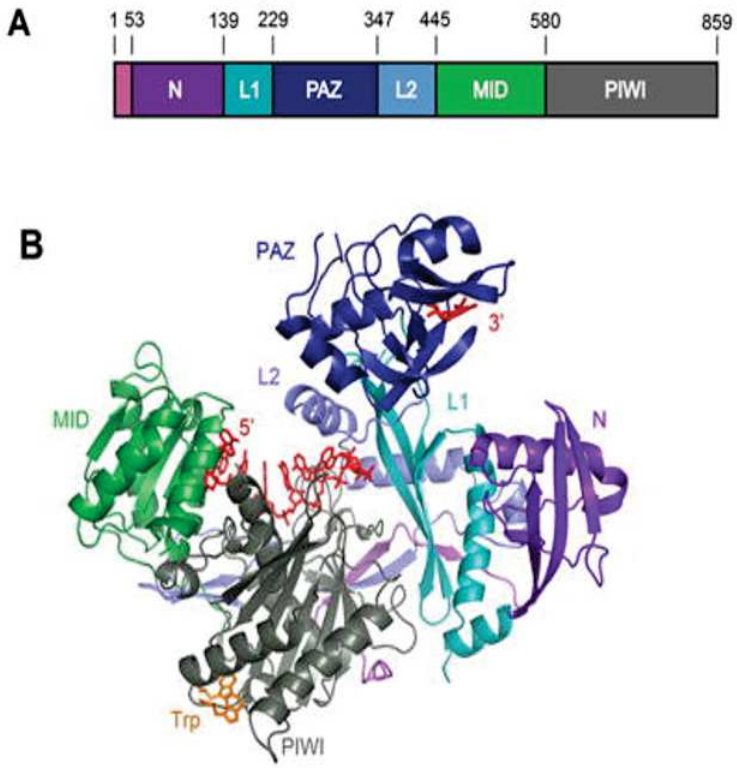


FIGURE 1.9: Schematic representation of AGO2 structure. A) Domain composition of human AGO2 primary sequence. B) Model of human AGO2. Adapted from Schirle and MacRae, 2012

The N domain is the least characterized of all AGO2 domains and it seems to be required for the unwinding of the RNA duplex but not for RNA loading (Kwak & Tomari, 2012). It was suggested that the N domain dynamically disrupts base pairing in the miRNA duplex loaded into AGO2. Furthermore, this domain might have a function during the target slicing step by suitably positioning the target RNA for endonucleolysis. Indeed, upon mutation of the N-domain, AGO2 fails to achieve activation of RISC and mRNA cleavage.

The PAZ domain is a very well conserved domain, which is also found in DICER,

that is able to bind the two-nucleotide 3′ overhang of miRNAs (J.-B. Ma et al., 2004). There are no sequence-specific contacts between the PAZ domain and the 3′ end of the miRNA, although the terminal nucleotide inserts in the aromatic ring of a conserved phenylalanine residue (Jinek & Doudna, 2009). In particular, it was shown that the PAZ domain cyclically binds and releases the 3′ end of siRNAs and that its action is essential for RNAi activity. Moreover, mutations at the 3′ end of siRNAs or blunt-ended substrates, have a detrimental effect in PAZ binding activity and they inhibit the formation of hydrogen bonds in PAZ active site (Alagia et al., 2018). However, the exact role of AGO2 PAZ domain in RISC is still in debate. It was also shown that AGO2 with a mutant PAZ domain could still interact with miRNA but it was unable to unwind the loaded miRNA (Gu et al., 2012).

The MID and PIWI domains form a very conserved interface, common with the prokaryotic structures, that behave like a channel to accommodate the miRNA duplex between the two lobes. In line with this, it was shown that MID-PIWI lobe is very sensitive to alterations, especially in their polar and hydrophilic interface. Mutations in this region lead to MID-PIWI interaction instability and the inhibition of substrate binding (Boland et al., 2011).
Specifically, the MID domain contains a binding pocket with a nucleotide specificity loop that contacts the 5′ end of the guide strand and it preferentially binds A or U nucleotides. On the other hand, the PIWI domain in AGO2 adopts a fold typical of endoribonucleases, such as RNAseH, and it exerts an endoribonuclease-like activity by slicing the target mRNAs. The PIWI domain also contains a tryptophan binding area, essential for the recruitment of GW182/TNRC6A, a scaffold protein important for AGO2 activity (Parker et al., 2005).
It is interesting to note that, although the core domain structure of AGO2 is well conserved between prokaryotes and eukaryotes, their relative positions in the structure differ, which may contribute to their major architectural differences (Schirle & MacRae, 2012).

1.4.2 RISC loading and interplay between DICER and AGO2

The final step in the assembly of a functional RISC is the loading of the miRNA onto AGO2 to form the pre-miRISC. A pre-miRISC composed of AGO2 and the miRNA is also called minimal RISC.
In this process, the partner strand (miR* strand) of the miRNA duplex, also called the passenger strand, is expelled from the pre-miRISC and degraded, while the other strand (miR strand) is selectively stabilized and assembled into AGO2 to form the miRISC.

A recent time-resolved analysis showed that amounts of miRs* and miRs do not show any statistically significant difference at early time-points after 4sU labeling, implying that miRNAs accumulate as duplexes. Interestingly, miRNA duplexes loading onto AGO2 occurs with 1 hour or more of delay, suggesting that RISC loading is a rate limiting step, representing a critical kinetic bottleneck for the accumulation of mature miRNAs and resulting in low yields in miRISC formation (Reichholf et al., 2019).
This suggests that RISC loading could require assistance from other factors such as chaperone proteins and ATP hydrolysis (Yoda et al., 2010). In particular, it was shown that the chaperone machinery composed of Hsc70 and Hsp90 uses ATP to mediate a high-energy conformational opening of AGO2 to accommodate the rigid miRNA duplex (Iwasaki et al., 2010).

AGO2 loading of miRNAs is followed by unwinding or cleavage of the duplex and the retention of a single-stranded RNA. If the duplex is perfectly matched, AGO2 is able to cleave the passenger strand and its expulsion is facilitated by the endonuclease C3PO (Sheu-Gruttadauria & MacRae, 2017). However, most miRNA duplexes present mismatches in the seed or 3'-mid regions that disfavor the cleavage of the passenger strand and favor miRNA unwinding. The miRNA unwinding seems to be independent of ATP (Yoda et al., 2010). In particular, it was proposed that the tension created by the chaperone-mediated opening could be released allowing AGO2 to return to a more closed conformation and unwind the

miRNA duplex (Kawamata & Tomari, 2010). Moreover, the structural asymmetry, characteristic of the miRNA duplex, is determinant for strand selection and the formation of a mature mi-RISC. In fact, data show that the strand with a thermodynamically less stable 5′ end becomes the miR strand and is retained in AGO2, while the other strand is expelled and degraded (Khvorova et al., 2003; Schwarz et al., 2003). The nucleolytic activity of AGO2 is critical for the cleavage of the passenger strand prior to its expulsion from the pre-miRISC but it was found to be important also for the maturation of DICER-independent miRNAs, such as miR-451 (Cifuentes et al., 2010).

The level of a miRNA does not strictly correlate with its association with RISC, suggesting that its abundance might also be controlled by other factors. *In vitro* studies indicated that human DICER might function as a nucleic acid annealer, facilitating interactions between complementary sequences in two RNA molecules; while AGO2 was found to have a limited RNA-annealing potential (Ameres et al., 2007; Kurzynska-Kokorniak et al., 2016). The role of DICER in the binding of AGO2 to target mRNAs is complex. On one hand, the binding of DICER to the miRNA-targeted mRNA has been shown to block association of AGO2 thereby preventing transcript degradation. On the other hand, DICER might also support AGO2 activity by loading the miRNAs or by facilitating the accessibility to the structured double-stranded fragments of transcripts, leading to degradation or translational repression (Pokornowska et al., 2020). Furthermore, this assumption is supported by reports showing that endogenous AGO proteins can bind to miRNAs that are already pre-annealed to mRNAs (Janas et al., 2012; B. Wang et al., 2006).

Other groups have previously identified a tripartite miRNA loading complex (miRLC), composed of DICER, AGO2 and TRBP that is able to transfer miRNA and siRNA from DICER to AGO2 (Maniataki & Mourelatos, 2005; H.-W. Wang et al., 2009). Interestingly, some studies suggest that this complex assembles spontaneously *in vitro* and prior to pre-miRNA recognition by DICER. Therefore, the miRLC displays several catalytical activities: it is able to recognize and bind pre-miRNAs, it processes them into miRNA duplexes and it cleaves the miRNA to

obtain a single stranded guide RNA (X. Liu et al., 2012; MacRae et al., 2008).
In addition, a miRNA Precursor Deposit Complex (miPDC) was shown to enhance the expression of certain miRNAs. This latter is composed by pre-miRNA directly binding to AGO2 and it's promoted by miRNA precursors beginning with 5'-U or a 5'-A and containing a 3'-mid base pairing and/or 5'-mid mismatches. Interestingly, miPDC seems to be particularly important for promoting RISC assembly for DICER-independent miRNAs. The assembly of DICER-dependent pre-miRNAs in miPDC promotes their maturation by escorting them to the miRLC, containing DICER. In the other hand, the assembly of DICER-independent pre-miRNAs in miPDC is crucial for their AGO2-dependent maturation (X. Liu et al., 2012).

1.4.3 Molecular mechanism of AGO2

After miRNA loading and elimination of the passenger strand, AGO2 is able to function as effector of the small-RNA-mediated silencing. As previously mentioned, an siRNA-guided AGO2 is able to slice the target mRNA thanks to the perfect complementarity between the "seed" region and the recognition element on the target RNA. In the majority of the cases, human AGO2 is loaded with a miRNA whose complementarity to the target mRNA is only partial. In this case, RISC induces translational repression or mRNA deadenylation, decapping and decay of the target mRNA (Meister, 2013). For pathways that rely on slicing of the RNAs, AGO2 cyclically binds, cleaves and releases the mRNA target while the guide strand remains bound within RISC. On the other hand, for slicer-independent pathways, AGO will repress target mRNA translation for as long as the protein is associated to it (Jinek & Doudna, 2009).

Interplay between translation and RISC mediated silencing

The miRNA recognition elements (MREs) on the target mRNAs are usually found in their 3'UTR but some MREs can be also found in 5' UTRs and ORFs of mammalian genes. However, no or only marginal efficacy was detected at these latter sites. This might be due to the implication of occlusive RNA structures and ribosome interference. Accordingly, MREs within 5'UTRs and ORFs would face ribosome interference because of the interaction of the poly(A) tail with the 5' cap

in the circular mRNA during translation (Bartel, 2009).

It was in fact demonstrated that translational repressed mRNAs can be detected at polysomes, suggesting that their repression occurs after cap recognition and, therefore, translational initiation. This implied that the mRNAs under miRNA control undergo at least one round of translation before or concurrently with their repression (Maroney et al., 2006). According to this line of evidence, RISC-mediated translational repression occurs by inhibiting translational elongation. In contrast with this, other studies showed that translational repression occurs at the initiation step by interfering with the cap structure and the cap-binding protein eIF4F (eukaryotic translation initiation factor 4E) (Zdanowicz et al., 2009).

It was also demonstrated that AGO proteins and their partners co-localize to processing bodies (P-bodies, PB), cytoplasmic *loci* that are involved in mRNA decay, nonsense-mediated decay, miRNA-mediated silencing, splicing and delayed translation. PB contain several proteins and factors such as decapping enzymes, translation regulating factors, including eIF4E, helicases and AGO proteins (Hubstenberger et al., 2017; Kanakamani et al., 2021; J. Liu et al., 2005). AGO proteins and their partners were found to co-localize and concentrate in P-bodies together with the mRNAs that undergo miRISC-mediated translational repression, in a miRNA-dependent manner, highlighting a connection between P-bodies and mRNA silencing (J. Liu et al., 2005; Meister et al., 2005). It was shown that the sequestration of repressed mRNAs to P-bodies enhances their interaction with AGO proteins as well as decapping and deadenylating enzymes, leading to target degradation or delayed translation (J. Liu et al., 2005). In fact, not all mRNAs that enter P-bodies are degraded but most mRNAs that are programmed for delayed translation can be stalled in P-bodies and, after the appropriate stimuli, can exit these foci and associate with polysomes or be recruited to stress granules to re-initiate translation (Bhattacharyya et al., 2006; Hubstenberger et al., 2017). Further evidence that confirms the link between P-bodies and RISC-mediated silencing comes from findings of a recent study where depletion of the nucleoporin Nup358 disrupts P-bodies and concomitantly impairs miRNA-mediated silencing

(Sahoo et al., 2017).

The interplay between P-bodies and RISC-mediated silencing is widely accepted and studied but the mechanism that regulates the fine tuning between degradation and mRNA translational delay is still not completely understood.

Target mRNA recognition by AGO2

The majority of MREs are hosted in 3′ UTRs, which are often several kilobases long. For this reason, it was suggested that AGO2 might use a lateral diffusion mechanism to identify its target MREs. According to this model, AGO2 is able to diffuse along the target mRNA scanning for a sub-"seed" recognition element of around 2-4 nts prior to a full "seed" motif binding (Chandradoss et al., 2015). Moreover, it was shown that this mechanism is more efficient in scanning unstructured RNAs since RNA secondary structures and the RBPs that recognize them most likely interfere with AGO2 diffusion (Ameres et al., 2007).

The context of the 3′ UTR containing the MRE can also influence the recognition and dwell time of RISC on a specific target mRNA. In fact, identical sites can lead to RISC-mediated repression in some UTRs but not in others (Bartel, 2009). For instance, it was found that positioning of miRISC at least 15 nt from the stop codon and away from the center of long UTRs enhances the MRE efficacy. This could be due to complex RNA structures within the mRNA that is the target of miRISC and substrate for the ribosome, at the same time. It's possible that long UTRs display occlusive RNA structures that could interfere with the RNAi machinery, especially in the middle of the UTRs where the least ribosome interference is detected.

Similarly, an AU-rich region around the MRE site can increase RISC efficiency, while the presence in proximity of several sites for the same or different co-expressed miRNAs might have a synergistic effect (Grimson et al., 2007). This could be explained by the fact that optimally spaced sites might favour cooperative contacts with the repressive machinery, or they might displace occlusive mRNA structures more easily. Interestingly, these closely spaced sites are the most co-conserved sites. Along the same lines, other than being more effective, sites

away from UTR center, at least 15 nt from stop codon or in AU-rich regions are also more conserved (Grimson et al., 2007).

Role of MOV10 in mRNA target recognition As previously mentioned, minimal RISC lacks unwinding activity and therefore scanning of structured RNAs is not efficient. For this reason, AGO2 requires a helicase to disrupt occlusive secondary RNA structures on the target mRNA 3′ UTR and to expose the MRE (Meister et al., 2005).
MOV10 (Moloney leukemia virus 10) is an ATP-dependent helicase that belongs to the UPF-like helicase superfamily 1 (SF1) that binds ssRNA and translocates in a 5′ to 3′ direction. MOV10 was initially described as an inhibitor of viral replication for HIV-1 and Hepatitis C viruses but also as an inhibitor of LINE-1 retrotransposition. In addition, MOV10 was found to co-localize in P-bodies together with AGO2 and other factors involved in RISC, identifying a role of MOV10 in miRNA-mediated regulation (Meister et al., 2005).

According to MOV10 PAR-CLIP results, it seems that MOV10 specifically binds in close proximity to UPF1 binding sites, to resolve structures and displace RBPs from the 3′ UTR of the target mRNA, exposing the MRE for AGO2 binding (Gregersen et al., 2014). Data shows that MOV10 is able to contact the 3′ UTR specifically at regions with low conservation and upstream of local secondary structures, consistent with its 5′ to 3′ directional unwinding activity (Gregersen et al., 2014). Moreover, for efficient loading and unwinding, MOV10, like many RNA helicases, needs a single-stranded region adjacent to a duplex (Q. Yang & Jankowsky, 2006).

In line with its role in RISC-mediated silencing, MOV10 can regulate the abundance of its bound mRNAs. In particular, it was shown that KD of MOV10 inhibits translational suppression leading to a global stabilization of its bound mRNAs (Gregersen et al., 2014). However, in contrast with this observation, MOV10 was also found to increase the expression of a limited subset of mRNAs by inhibiting AGO2 binding, in the presence of FMRP (fragile X mental retardation protein 1). Therefore, the concomitant binding of MOV10 and FMRP on the same mRNA can

block the canonical role of MOV10 in RISC by G-quadruplex binding(Kenny et al., 2020; Kenny et al., 2014).

Interaction of miRNA and target mRNA

The 6-nt long "seed" region of miRNA is the primary interaction interface between the miRISC and the target mRNA. Indeed, around 80% of miRNA-target interactions occur via this region (Grosswendt et al., 2014). In line with this observation, the "seed" region is the most evolutionary conserved portion of the miRNA (Krek et al., 2005). However, it was recently demonstrated that around 20% of miRNA-target interactions involve a supplementary region at the 3′ end of miRNAs (Marzec, 2020). This latter portion of miRNAs tends to be less conserved and it can vary even among miRNAs that belong to the same family (Wahlquist et al., 2014). It has been suggested that the binding of this so called 3′ compensatory site could enhance recognition of the "seed" matched target and it might compensate for a weak "seed" interaction or, in some cases, even replace it (Grimson et al., 2007). For instance, 3′ compensatory sites containing GC-supplementary sequences were shown to significantly increase target affinity (Sheu-Gruttadauria et al., 2019).

A recent crystal structure of human AGO2 bound to miRNA and target mRNA shows that the 3′ compensatory site bound to the target is accomodated in a supplementary chamber, physically separated from the "seed" chamber by a central gate. The supplementary chamber can host up to 5 miRNA–target base pairs and a linker bridge of 2-15 nt (central region) can be found between the two interaction chambers. Moreover, AGO2 PAZ domain needs to undergo a conformational change that is facilitated by recognition of the miRNA 3′end in order for the 3′ compensatory sites to bind the target mRNA. In fact, mutations at this sites (isomiRs) alter the stability of the supplementary interactions (Sheu-Gruttadauria et al., 2019). The same study suggests that the supplementary interaction could serve as a first step of pairing, interrogating for 3′ end complementarity before the commitment step and the binding with the "seed" region. Similarly, binding to a 3′ compensatory site is followed by the opening of the central gate and the release of

the 3’ end of the miRNA to allow "seed" region pairing (Sheu-Gruttadauria et al., 2019).

Interestingly, for the MREs with enhanced target repression, an A at position 1 appears to be important, regardless of the miRNA sequence. Structural analyses show that the first nucleotide of the miRNA is buried within Ago, making impossible for it to bind to the A in the MRE of the mRNA. Therefore, it was suggested that this particular A might be important for an RBP binding, possibly to increase the dwell-time on the target mRNA (Lewis et al., 2005; Schirle et al., 2015).

It was observed the presence of non-canonical sites that lack the contiguous 6-nt match to the "seed" region. High-throughput analyses revealed that AGO2 binds to many non-canonical sites, accounting for approximately 50% of crosslinked sites. However, these interactions don’t show a significant repression of the associated mRNA, supporting the hypothesis that AGO2 samples many low affinity sites by lateral diffusion before binding a site with sufficient dwell time to engage target repression (Bartel, 2009).

On another note, it was recently shown that not all miRNAs need to be loaded onto Ago prior to mRNA binding. In fact, it was observed that mRNAs associate with a seven-fold excess of miRNAs compared to Ago proteins (Janas et al., 2012). This suggests that Ago is able to bind pre-annealed miRNA-mRNA duplexes, therefore revealing an alternative mechanism of RISC assembly. According to this mechanism AGO2 is able to interact with preannealed miRNA-mRNA duplexes and exert its function of repression. Consequently, RISC activity is not limited to an equimolar ratio between AGO2 and miRNA, consistent with the finding that miRNAs are expressed in a 13-fold excess compared to AGO2 proteins (Janas et al., 2012).

RISC mediated translation inhibition, mRNA decapping, deadenylation and decay

The association of RISC with a mRNA induces gene repression, which can occur through two different mechanisms. As already mentioned before, RISC-mediated gene repression can be the result of target mRNA slicing and degradation or translational repression/mRNA decay (Bartel, 2018).
If the guide strand is perfectly complementary to the target mRNA and if it is loaded onto AGO2 with retained cleavage ability, the transcript will undergo endonucleolytic slicing. In this case, the cleavage occurs around 10 nts away from the 5′ end of the miRNA, leaving it intact and ready for another round of slicing. As previously discussed, slicing occurs in the PIWI domain of AGO2 which contains a catalytic motif able to coordinate a manganese ion, essential for cleavage (J. Liu et al., 2004). Although this slicing mode of repression is common for miRNAs in plants, in human this is the case for only a very limited subset of miRNAs and targets, around 20 cellular transcripts and a few viral RNAs (Barth et al., 2008; Shin et al., 2010).

On the other hand, the dominant mechanism for target repression in mammals is slicing independent and it occurs for miRNAs only partially complementary to their target transcripts. This mechanism is able to achieve gene silencing by inhibiting target transcripts translation or mediating mRNA decay after decapping and deadenylation (Bartel, 2018). This mechanism, if needed, can be finely tuned to assure translational re-initiation. This mechanism requires the adaptor protein GW182/TNRC6A (Trinucleotide repeat-containing gene 6A protein) that acts as a scaffold to recruit downstream effector proteins inducing translational repression, mRNA deadenylation, decapping and 5′-to-3′ exonucleolytic degradation. TNRC6A has different functional domains and, between them, the tryptophan (W) containing domain is important for its binding to AGO2 and to the subunits PAN3 and CNOT1, part of the cytoplasmic deadenylases complexes PAN2-PAN3 and CCR4-NOT, respectively.

In particular, target mRNAs are first deadenylated by the consecutive and partially redundant PAN2-PAN3 and CCR4-NOT deadenylase complexes. Several observations strongly suggest that CCR4-NOT complex can take over the whole deadenylation process in absence of PAN2-PAN3. For instance, TNCR6 can mediate promiscuous binding to PAN2-PAN3 and CCR4-NOT complexes, while the depletion of PAN2-PAN3 complex leads to only a minor change in silencing efficiency. According to the most recent model, the deadenylation process is started by PAN2-PAN3 by shortening the poly(A) tail, without causing RNA decay. Once the poly(A) tails are shortened to 110 nt poly(A) or less, CCR4-NOT complex can continue the deadenylation step in a processing manner, committing to RNA degradation (Jonas & Izaurralde, 2015; Wahle & Winkler, 2013).

Importantly, CCR4-NOT complex directly interacts with factors that catalyze translational repression and decapping, providing a platform that efficiently couples deadenylation and decapping. It interacts with DDX6 and PATL1, which function as a physical link with additional decapping proteins such as decapping protein 2 (DCP2) and cofactors like DCP1 and enhancer of decapping 3 (EDC3) and EDC4. Finally, it was shown that EDC3 and EDC4 interact with the major cytoplasmic 5'-to-3'nuclease exoribonuclease 1 (XRN1) which degrades deadenylated and decapped mRNAs (Y. Chen et al., 2014).
Moreover, TNRC6A interacts with proteins associated with the poly(A)-tail of the target mRNA, Poly(A)-binding proteins (PABP). Their interaction is thought to enhance RNA silencing by two different mechanisms. It is possible that TNCR6 binding to PABP selects poly(A) RNAs as targets, which was shown to enhance miRISC activity. On the other hand, it is also possible that by interacting with PABP, TNCR6 could interfere with its normal function of enhancing translation (Fabian et al., 2009). Finally, after repression of the target mRNA, the cluster of serine/threonine residues at the C-terminus of AGO2 are hyperphosphorilated and their acquired negative charges favor the release of the target mRNA. Therefore, the balance between the phosphorylation and dephosphorylation might be important for redirecting AGO2 to a new target mRNA (Quévillon Huberdeau et al., 2017).

To conclude, the question of how mRNA decay is chosen over translational repression and *vice versa* is still quite controversial but recent findings support the idea that the outcome of mRNA decay accounts for most of miRNA-mediated silencing events. Loss of translational efficiency is mediated by the recruitment of DDX6 by CCR4-NOT, however the effect is usually weak and, in most of cases, is followed by mRNA decay. RISC-mediated gene repression can occur exclusively via translational repression but this accounts for only about 6-26% of endogenous targets, making mRNA decay the major route for miRNA-mediated silencing (Jonas & Izaurralde, 2015). It is important to note that the choice between translational repression and mRNA decay largely depends on the post-transcriptional cellular context after the poly(A)-tail shortening mediated by RISC. For instance, it was shown that in early embryos the shortening of the poly(A)-tail leads to translational repression without affecting mRNA stability. On the other hand, in post-embryonic cells, the shortening of the tail reduces mRNA stability without changing its translational efficiency (Subtelny et al., 2014).

Non-canonical activities of Ago proteins

Although the main role of Ago proteins is exerted into the cytoplasm, a growing number of reports reveal a nuclear localization for Agos, suggesting other non-canonical roles for these proteins (Sala et al., 2020). For instance, it was shown that nuclear Ago is able to ensure correct chromosome segregation during cell division and it seems to be important for early embryogenesis in different species (Claycomb et al., 2009; Pushpavalli et al., 2014). This represents one of the best characterized non-canonical activity of Ago, whose absence can lead to infertility in animals and several defects of meiosis and mitosis (Yigit et al., 2006).

Less characterized functions involve splicing regulation and nuclear activation or repression of gene expression. Ago1 was shown to bind active transcriptional enhancers, interestingly leading to regulation of constitutive and alternative splicing rather than regulation of gene transcription (Alló et al., 2014). Moreover, Ago1 was also found directly associated with RNA Polymerase II on transcritionally active genes and it seems to be involved in the transcriptional

regulation of oncogenes in cancer cells (V. Huang et al., 2013; Janowski et al., 2006).
In line with its involvement in RISC activity, AGO2 was also shown to be important for the maturation of miRNAs that cannot be processed by DICER (Cifuentes et al., 2010). For instance, AGO2 is required for the maturation of miR-451, an erythroid-specific miRNA, important for erythrocyte maturation. Interestingly, it was shown that catalytic inactive AGO2 leads to perinatal lethality in mice due to a strong anemia caused by the lack of miR-451 (Papapetrou et al., 2010).
Moreover, AGO2 seems to be involved in miRNA stability since it was shown that its KD reduces the half-life of a subset of miRNAs. On the other hand, overexpression of AGO2 significantly diminishes miRNA degradation, increasing their half-life and their abundance post transcriptionally (Winter & Diederichs, 2011).

1.4.4 AGO2 regulation

The observation that the expression level of AGO2 correlates with the levels of mature miRNAs suggests that AGO2 expression must be strictly regulated in order to maintain miRNAs homeostasis (Diederichs & Haber, 2007).
Hydroxylation of Ago proteins was shown to increase their stability (Qi et al., 2008). In particular, hydroxylation of AGO2 can be mediated by hypoxia, leading to increased endonuclease activity, accumulation of AGO2 and miRNAs. Hydroxylation seems to be important also for AGO2 association with Hsp90 which, as previously mentioned, is essential for miRNA loading onto Ago (C. Wu et al., 2011).

AGO2 deregulation also correlates with a transformed phenotype in several types of cancers. It was shown that epidermal growth factor (EGF) receptor/MAPK signaling pathway is able to modulate AGO2 expression both transcriptionally and post-translationally, in particular by enhancing its stability. This modulation was linked to tumorigenic progression on breast cancer (Adams et al., 2009).
More recently, hypoxia-induced upregulation of EGFR was shown to mediate phosphorylation of AGO2 tyrosine 393 (Y393) in breast cancer samples. This phosphorylation impairs DICER interaction with AGO2 and therefore loading of a

subset of tumor suppressor miRNAs (J. Shen et al., 2013). Several phosphorylation sites have been mapped on Ago proteins and it's tempting to speculate that many signalling pathways are involved in post-translational regulation of AGO2 (Rüdel et al., 2011).

1.4.5 Other proteins involved in RISC-mediated mRNA silencing

In recent years, many proteins have been found to be involved in the regulation of the minimal RISC.

As previously discussed, miRNA loading represents a rate limiting step in the miRNAs-mediated silencing pathway. Besides the already mentioned chaperone proteins Hsc70 and Hsp90, that mediate a conformational change in AGO2 to accomodate the miRNA, other RBPs have been described to impact this step positively or negatively. For instance, the RBP AU-rich binding factor 1 (AUF1) was shown to enhance the loading of a subset of miRNAs onto AGO2, promoting decay of targets mRNAs (Yoon et al., 2015). AUF1, which recognizes U-/GU-rich regions in the target transcripts, was also shown to act on another level of the miRNA-mediated silencing pathway by binding to DICER mRNA decreasing its expression, thereby diminishing miRNA biogenesis (Min et al., 2017).

On the other hand, HuR was shown to inhibit AGO2 loading by sequestering a subset of miRNAs, such as miR-122, and promoting their extracellular export (Mukherjee et al., 2016).

The fragile X mental retardation protein (FRMP) was shown to be involved in RISC-mediated silencing by binding to a subset of MOV10-bound mRNAs. In particular, it was observed that , in the presence of a G-quadruplex structure, FMRP can bind close to MOV10 binding sites on the 3'UTRs. The N-terminus of MOV10 is required for FRMP stabilization on the mRNA which ultimately blocks MOV10 activity and protects the target mRNA from AGO2 binding (Kenny et al., 2014). More recently, it was shown that, for a subset of mRNA, FMRP is able to enhance RISC-mediated silencing by recruiting MOV10 and facilitating the resolution of RNA secondary structures (Kenny et al., 2020). However, the mechanism underling the bifunctional role of FMRP in mRNA decay and translation still needs to be

elucidated.

In the last years, many more RBPs have been shown to participate in RISC-mediated silencing at different steps but their description goes beyond the scope of this thesis and therefore they will not be further examined.

1.5 miRNAs stability and decay

While the miRNA biogenesis pathway and its regulation have been widely studied and characterized, little is known about miRNA stability and turnover.

Although it's tempting to think that the lack of protective covalent modification on miRNAs would make them particularly exposed to the action of RNA exonucleases, miRNAs duplexes are generally very stable (Gantier et al., 2011). In particular, the half-life of a miRNA duplex ranges from hours to days, making them around 10 times more stable than mRNAs (Gantier et al., 2011; Guo et al., 2015). This high stability translates into the persistence of miRNAs for hours or even days after their production is interrupted. Therefore, to avoid unwanted miRNAs effects, a mechanism of active specific decay is essential for situations that need rapid changes in miRNA abundance and function (Krol, Busskamp, et al., 2010) and its existence was demonstrated more than 10 years ago (Ameres et al., 2010; Rüegger & Großhans, 2012).

However, the mechanism of miRNA degradation is not fully understood particularly, how selective degradation is able to target specific miRNAs, even discriminating between sequence-related family members. Recent evidence suggests that specific miRNA degradation could be achieved by reverting the direction of the canonical miRNA-mediated silencing, an emerging pathway known as Target RNA-Directed MicroRNA Degradation (TDMD), in which the binding of the target mRNA could lead to miRNA degradation (Ameres et al., 2010). This mechanism was first described in *Drosophila* and it seems to be conserved in mammals, where it was shown that the active degradation of miRNAs occurs after their maturation and loading onto Ago proteins (Ameres et al., 2010). In particular, TDMD seems to be initiated by a specific miRNA-target binding

architecture that leads to target mRNA silencing evasion and destabilization of the bound miRNA. As for the canonical miRNA pathway, the extent of mRNA-miRNA binding seems to be crucial, requiring extensive pairing of the miRNA 3' compensatory site (Fuchs Wightman et al., 2018). The extensive binding of the 3' region of the miRNA to the target triggers the miRNA 3'-end tailing, with the addition of non-templated A or U nucleotides, followed by trimming of the 3'-end and decay.

Examples of TDMD in mammals are quite scarse but it seems to be more efficient in neuronal cells. For instance, the lncRNA *Cyrano* has been shown to trigger miR-7 decay through TDMD in the peripheral nervous system in mice and humans (Kleaveland et al., 2018). Indeed, *Cyrano* KO leads to increase of miR-7 and subsequent repression of miR-7 target mRNAs. Interestingly, *Cyrano* was first described as decoy for miR-7 because of their extensive complementarity (K. N. Smith et al., 2017). It's therefore tempting to speculate that the many already known ncRNAs acting as sponges or decoys for several miRNAs might affect their activity non only by sequestering the miRNAs but also by triggering their decay.
In addition to regulating the level of miRNA and, consequently, their target mRNA, TDMD might facilitate Ago recycling. Since miRNAs are expressed in excess compared to Ago proteins, degradation of Ago loaded miRNAs by TDMD might free Ago for the loading of other miRNAs (Hausser et al., 2013).

In contrast, more recent findings suggest that AGO2 binding protects miRNAs from TDMD (Kingston & Bartel, 2021).
However, it is possible that while AGO2 binding is the first step for TDMD only for a limited subset of miRNAs, miRNAs not bound by AGO2 might be target of other pathways of degradation. It was recently shown that around 40% of miRNAs duplexes are degradated before loading onto Ago (Reichholf et al., 2019). Several exonucleases have been found to non-specifically degrade unbound miRNAs. For instance, GW182/TNRC6A KD was shown to promote the non-specific degradation of several miRNAs by the 3'-5' RRP41 exoribonuclease (Yao et al., 2012). According to these observations, AGO2 binding of the miRNA would

prevent degradation by 3′-5′ exonucleases because the 3′ end of the miRNA should reside within the Ago protein.

Interestingly, recent findings suggest that miR* exhibit a shorter half-life compared to the miR strands, supporting the observation that expulsion of the miR* strand after Ago loading is followed by its degradation. On the other hand, miR that are loaded onto Ago proteins show significantly longer half-lives because Ago might protect them from degradation (Reichholf et al., 2019).
Given the extensive role of miRNAs in the regulation of gene expression, their fine tuning is essential for homeostasis. While miRNA biogenesis is a widely studied and well described pathway, miRNA turnover is much less investigated but is nonetheless important and needs to be further explored.

1.6 Non-canonical miRNAs

In recent years, studies have identified alternative miRNAs biogenesis pathways that generate non-canonical miRNAs. As for canonical miRNAs, non-canonical miRNAs seem to be involved in many biological pathways and their dysregulation leads to major changes of cellular phenotype and pathological effects, such as diabetes and cancer (Bronisz et al., 2011; Ling et al., 2009).

miRtrons

A small class of non-coding miRNA is represented by miRtrons, intronic miRNAs that are found in small introns. The size of the host intron is comparable to the length of the pre-miRNA itself since its 3′ and 5′ ends match the 3′ and 5′ splicing sites.
These miRNAs were initially reported in *Drosophila melanogaster* and *C. Elegans* and are also found in mammals. The mechanism that led to their evolution in mammals is however rather controversial (Berezikov et al., 2007; Ruby et al., 2007). The many differences between mammalian and invertebrate miRtrons suggest a convergent evolution of this non-canonical mechanism from the ancestral RNAi pathway. Initially, it was proposed that miRtrons arose in genomes proportionally to the

fraction of small introns. However this hypothesis can be confuted by the fact that human brain tissue expresses more miRtrons than flies, where such introns are more abundant, for instance. Moreover, mammalian miRtrons do not simply randomly derive from short introns since they show a significantly higher GC content than introns (Berezikov et al., 2007).

Similarly to canonical intronic miRNAs, miRtrons are not necessarily co-transcribed with their host gene (section 1.3.2). Once transcribed, miRtrons are generated from the primary transcript via a non-canonical DROSHA-independent mechanism that involves splicing and DICER cleavage. In fact, miRtrons precursor miRNAs are shorter than canonican pri-miRNAs and they are composed of the miR/miR* duplex while lacking the lower part of the stem that typically recruits the Microprocessor. As per all spliced introns, the initial product of miRtron biogenesis is a lariat in which the 3' branchpoint is ligated to the 5' end of the intron (Fig. 1.10). Subsequently, this structure is resolved and the pre-miRNA can adopt the typical hairpin structure that is recognized by XPO5 for export to the cytoplasm where it is further recognized and cleaved by DICER (Westholm & Lai, 2011).

However, in some cases, miRtron loci can reside on one end of the intron, sharing only the 5' or the 3' splicing site with the intron. These are called 5' or 3' "tailed" miRtrons and, after the debranching step, they require trimming by a 5'-3' exonuclease or 3'-5' exosome degradation, respectively (Flynt et al., 2010).
MiRtrons were found to be processed at a significantly lower rate than canonical miRNAs, probably due to the necessity of the trimming step. However, after loading onto Ago, they stability is comparable to canonical miRNAs (Reichholf et al., 2019). Interestingly, for the same primary transcript transcription level, miRtrons were found to yield fewer mature reads in small RNA libraries, suggesting that, being a recently evolved class of miRNAs, their biogenesis could be strictly regulated (section 1.2.2) (Westholm & Lai, 2011).

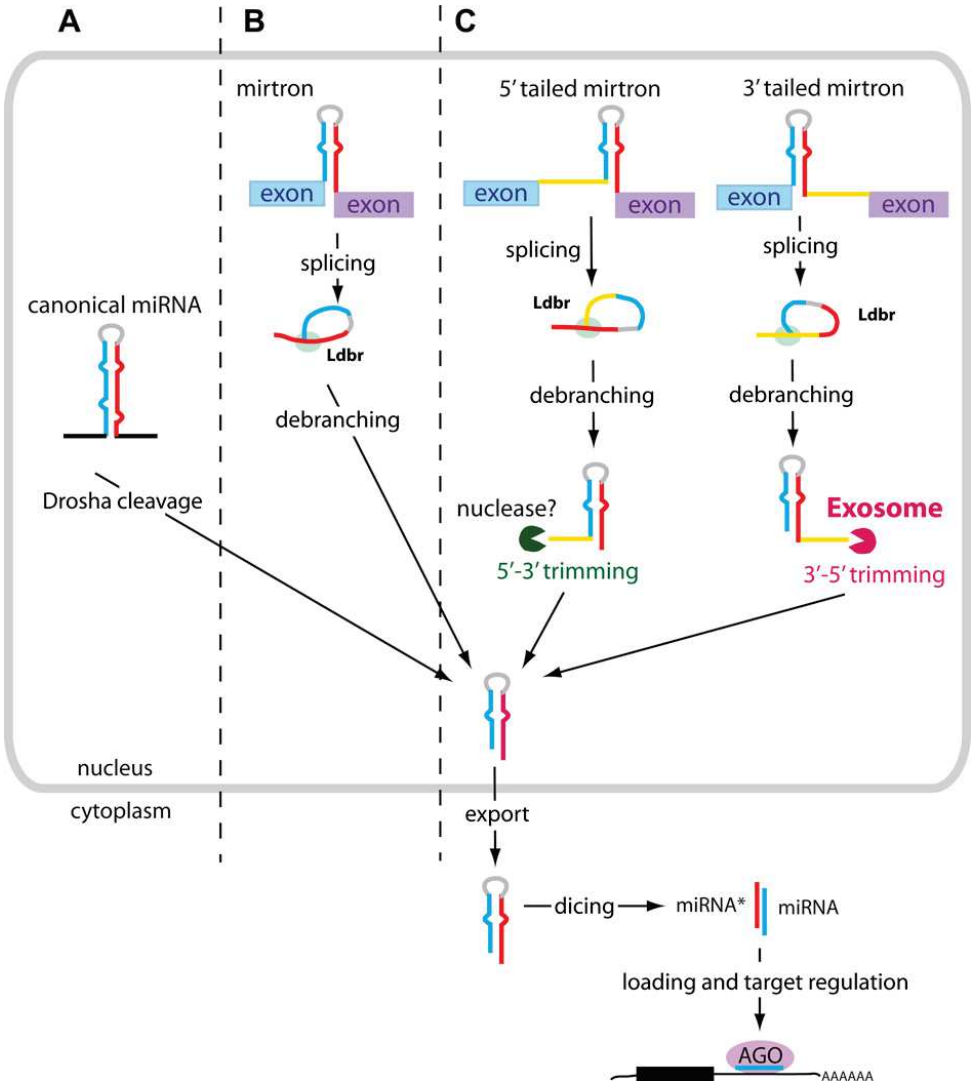


FIGURE 1.10: A) The canonical miRNA pathway. B) Mirtron pathway starts with splicing of small introns that are subsequently debranched by lariat debranching enzyme (Ldbr), before folding into pre-miRNA hairpins. (C) Tailed mirtrons undergo splicing and debranching, followed by trimming of the tails. 3′ tails are trimmed by the RNA exosome, while the enzymes responsible for 5′ trimming are not known. Figure from Westholm and Lai, 2011

Simtrons

Simtrons (splicing-independent miRtron-like miRNAs) can be classified as a subset of non-classical miRtrons. In fact, like miRtrons, simtrons can be found in small introns of a primary transcript and their 5′ and 3′ end often match with the 5′ and 3′ splicing sites. However, simtron biogenesis seems to be independent from splicing. In addition, manipulations of DGCR8, XPO5, DICER and Ago didn't influence simtron biogenesis while DROSHA seems to be essential for their maturation

(Havens et al., 2012). However, some aspects of this pathway still remain unclear. It was shown that, for a subset of simtrons, splicing could be used for their maturation in absence of DROSHA. Therefore, an interdependence of miRtrons and simtrons pathways is plausible (Curtis et al., 2012; Stavast & Erkeland, 2019).

snoRNA and tRNA-derived miRNAs

New findings suggest that ncRNAs themselves can be source of non-canonical miRNAs.

For instance, deep sequencing data suggests that small nucleolar RNAs (snoRNAs), an abundant class of nuclear ncRNAs with different functions, can associate with Ago proteins, giving rise to miRNAs that may follow the canonical biogenesis pathway and contribute to RNA silencing (Patterson et al., 2017) (Fig. 1.11).

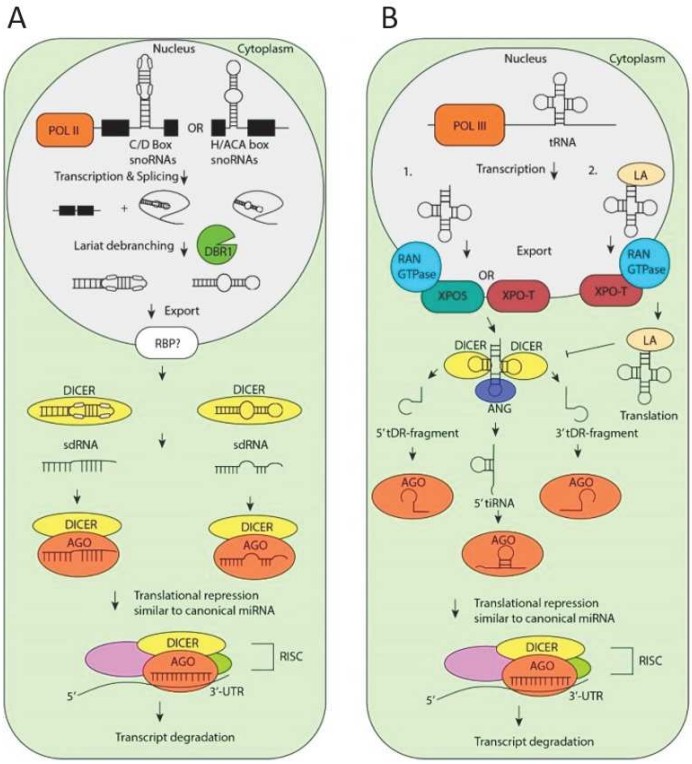


FIGURE 1.11: Schematic overview of snoRNA and tRNA-derived miRNAs biogenesis. A) Non-canonical processing of small nucleolar RNAs (snoRNAs) results in snoRNA-derived miRNAs. B) Non-canonical processing of transfer RNAs (tRNAs) results in tRNA-derived miRNAs. Adapted from Stavast and Erkeland, 2019

More studies identified tRNAs as a major source of small RNAs in Hela cells. In particular, they were found to be cleaved at their 5' ends by DICER to produce

miRNAs that were able to bind AGO2 (Cole et al., 2009). However, whether they are implicated in gene silencing has not yet been fully elucidated.

Nuclear miRNAs

In contrast to their main cytoplasmic function, some miRNAs have been shown to localize in the nucleus (Stavast & Erkeland, 2019). For example, miR-122, the most abundant tissue-specific miRNA in liver cells, has been observed to be actively transported from the cytoplasm to the nucleus (Földes-Papp et al., 2009).
Recent findings suggest the existence of an active and continuous shuttling of RISC and miRNAs between the cytoplasm and the nucleus. However, a subset of miRNAs seem to be particularly enriched in the nucleus, possibly because they interact with nuclear targets (Pitchiaya et al., 2017). On the other hand, it is also possible that a group of miRNAs, under specific circumstances, might be processed by DICER directly in the nucleus. In fact, there is strong evidence suggesting that functional DICER can be found in the nucleus together with the other fundamental factors involved in RISC-silencing, such as Ago, TRBP and GW182/TNRC6A (Burger et al., 2017).
As regards their function, nuclear miRNAs are involved in gene regulation not only through the canonical pathway, by targeting nuclear mRNAs, but also by interacting with gene promoters or antisense transcripts, leading to transcriptional silencing or activation (Miao et al., 2016; Place et al., 2008). For a few miRNAs, it was shown that their binding to chromosomal DNA might locally change the DNA structure, facilitating inactive or active chromatin histone marks (Matsui et al., 2013). However, the exact mechanism that results in transcriptional regulation is not fully known.

1.7 miRNAs in gene regulation

As previously discussed, miRNAs are predicted to regulate the expression of more than 60% of mammalian genes, being involved in almost all biological pathways. Gene regulation mediated by miRNAs is extremely important in physiological conditions, to maintain cellular homeostasis, but also in pathological conditions,

such as viral response and especially cancer.

Therefore, miRNAs are a widespread and efficient gene regulation mechanism that is able to dynamically regulate most genes by rapidly responding and changing according to cellular needs and stimuli (Bhattacharyya et al., 2006).

1.7.1 miRNAs in homeostatic gene silencing

Literature in recent years suggests the involvment of miRNAs in virtually all molecular pathways to ensure homeostatic levels of target mRNAs. For instance, it was recently shown that the hepatic miR-378 is able to control cholesterol metabolism. High miR-378 in mice leads to decreased serum cholesterol levels and resistance to hypercholesterolemia, while low miR-378 levels show defects of cholesterol homeostasis (C. Sun et al., 2021).

Many miRNAs have been linked to enbryonic development and hematopoiesis, such as miR-142. Lack of miR-142 in mice leads to abnormal hematopoietic lineage and extremely low immune response. On the other hand, miR-142-*null* mice exhibit decreased proliferation of mesenchymal cells during lung development (Shrestha et al., 2017).

Another miRNA involved in hematopoietic differentiation is miR-126 which was also linked to vascular homeostasis, being implicated in angiogenesis, vasculogenesis and endothelial inflammation (van Solingen et al., 2015).

It's interesting to note that, in order to maintain homeostasis and avoid deleterious off-targets effects, genes involved in basic cellular processes bear extremely short 3' UTRs that are specifically depleted of MREs to evade regulation by miRNAs (Stark et al., 2005).

1.7.2 miRNAs in cancer

The first evidence that miRNAs are implicated in cancer comes from Calin et al., 2002. In this study it was shown that the loss of a region in chromosome 13, where miR-15 and miR-16 are located, is involved in the pathogenesis of Chronic lymphocytic leukemia (CLL), identifying these two miRNAs as cancer suppressors. Since this first indication, many more studies have revealed a diversified and yet

strong link between miRNA expression and tumor development. Deregulated miRNA expression can be the output of several mechanisms such as aberrant transcription, genetic abnormalities, epigenetic factors and alteration of the miRNA biogenesis pathway (Di Leva et al., 2014). Despite the process that leads to their deregulation, miRNAs are generally globally depleted in cancers relative to healthy tissue (Lu et al., 2005). In agreement with this observation, loss of proteins involved in miRNA biogenesis contributes to cancer progression and metastasis (Barbier et al., 2018). Therefore, most human miRNAs can be classified as tumor-suppressor miR, since their depletion, as observed for miR-15 and miR-16, is sufficient to initiate cancer (Calin et al., 2002). On the other hand, transgenic expression of some miRNAs, called oncomiRs, such as miR-21, is directly linked to cancer initiation and progression (Medina et al., 2010).

Being widely involved in gene expression regulation, miRNA deregulation can impact different stages of cancer progression such as proliferation, apoptosis, migration and invasion, often controlling the epithelial-to-mesenchymal transition (EMT) phenotype (Gollavilli et al., 2021; Higuchi et al., 2016). For example, miR-95 has been shown to repress the expression of AHNAK, a tumor suppressor protein, promoting EMT in gastric cancer (E. Shen et al., 2020). In fact, deregulated miRNAs are frequently linked to development, progression and metastasis, acting as well-known hallmarks for several types of cancers (Y. S. Lee & Dutta, 2009). For instance, the miR-34 family has been extensively linked to cancer. These miRNAs are direct p53 targets and their downregulation was shown to correlate with higher proliferation and metastasis formation in several types of cancers (Bommer et al., 2007).

In the last decade, growing evidence for the involvement of miRNAs in cancer and the improvement of the techniques used to detect them, has encouraged their use in diagnosis and therapy. The sequencing of miRNAs, called miRNA profiling, can easily discriminate between tumor and non-tumor tissues and even between different subgroups of the same cancer (Drusco et al., 2014). Furthermore, miRNAs are important indicators of drug resistance and their presence circulating in patient

serum opened the possibility to use them for cancer prevention, as biomarkers for early diagnosis or for personalized therapy (Acunzo et al., 2015; Mollaei et al., 2019). Various strategies, currently tested in clinical trials, have been investigated to restore tumor suppressor miR, using synthetically derived oligonucleotides duplexes that mimic the endogenous miRNAs. On the contrary, different strategies have been tested to suppress oncomiRs using antimiRs based on antisense oligonucleotides (ASOs) (Rupaimoole & Slack, 2017).

To conclude, the growing evidence associating miRNAs and cancer encourages researchers in the world to investigate new solutions for the use of miRNAs as biomarkers and therapeutic targets. Therefore a deeper knowledge on miRNAs and miRNAs pathways is essential to this end.

1.8 NF90, double stranded RNA binding protein

Nuclear Factor 90 (NF90) is an RNA-binding protein which is part of a family of proteins produced from the *Ilf3* gene. *Ilf3* is localized on the human chromosome 19 and it contains 22 exons, giving rise to at least five distinct transcripts, generated by alternative splicing. The most abundant isoforms produced from *Ilf3* gene are NF90 and NF110, of apparent molecular mass of 90 kDa and 110 kDa, respectively (Fig. 1.12)(Castella et al., 2015; Masuda et al., 2013).

NF90 and NF110 are ubiquitous, and generally abundant, proteins expressed in animals but not in archea, eubacteria, unicellular organisms nor plants (Castella et al., 2015).

The main protein partner of NF90 and NF110 is NF45, which is transcribed from the *Ilf2* gene. Their heterodimerization is important for their stabilization and function. In particular, it was shown that the binding of NF45 to NF90 significantly leads to thermodynamic stabilization and, importantly, it improves the RNA-binding ability of NF90, enhancing its affinity for RNA substrates (Schmidt et al., 2017). Moreover, in most cell types NF90 is largely found tightly complexed with NF45, in an RNA-independent manner, and predominantly in the nucleus. However, NF90 was shown to be bound by XPO5, which promotes its nuclear

export in an RNA-dependent manner (Gwizdek et al., 2004). Furthermore, NF90 and NF45 have both been shown to shuttle between the nucleus and the cytoplasm, according to their phosphorylation status (Parrott et al., 2005) or as a result of several stimuli. For instance, viral infection of cells was shown to induce rapid NF90 translocation from the nucleus to the cytoplasm (X. Li et al., 2017).

1.8.1 NF90 structure and binding mode

NF90 and NF110 share the same N-terminal and central regions but they differ at the C-terminal region. In particular, 17 out of 22 exons are shared between these two isoforms. However,the C-terminus of NF90 is coded by exon 19 while that of NF110 is coded by exon 22 (Fig. 1.12) (Castella et al., 2015).

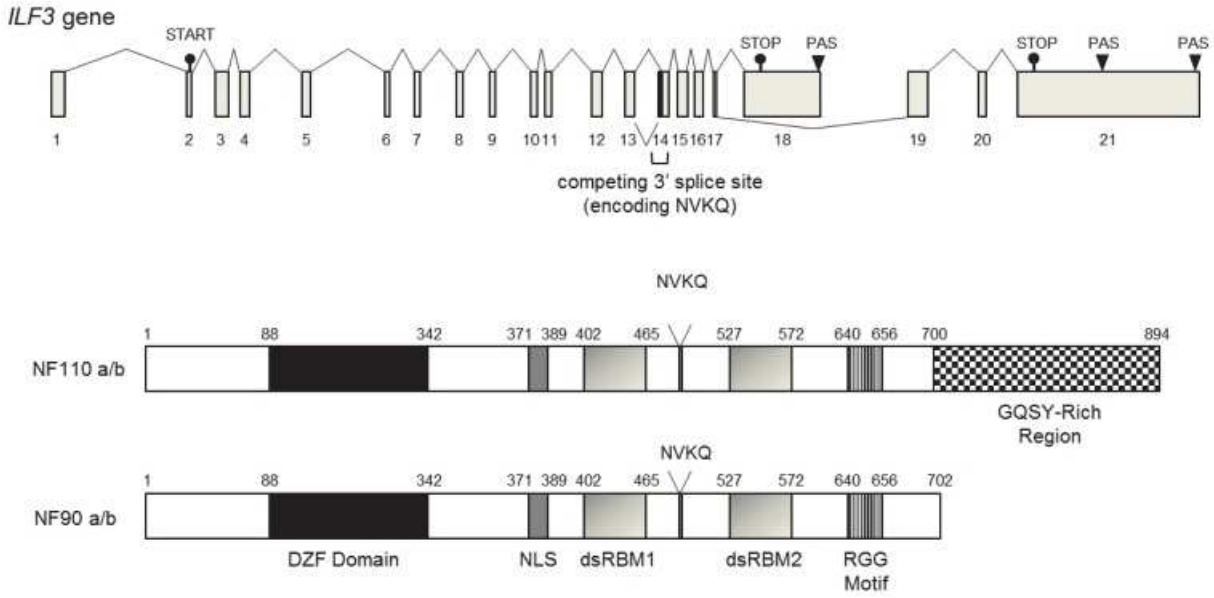


FIGURE 1.12: Schematic representation of *ilf3* gene and its two main products, NF90 and NF110, from Masuda et al., 2013

Ilf3 gene is a substrate for several observed alternative splicing events. For instance, exon 3 contains an alternative splicing site generating, for both NF90 and NF110, a long isoform containing exon 3 and a short one which lacks it (Viranaicken et al., 2006). Another alternative splicing event occurs at the 3' splicing site between intron 13 and exon 14. If the competing 3' splice site is recognized instead of the canonical one, four additional amino acids (NVKQ) are translated, encoding for NF90a/b and NF110a/b, where b forms present the insert (Duchange et al., 2000; Patiño et al., 2015).

The region common to NF90 and NF110 contains a DZF domain (dsRBM and zinc finger associated), a nuclear localization signal (NLS), two dsRBMs and a RGG motif.
The RGG motif typically consists of several copies or arginine and glycine repeats and they interact with ssRNAs. The RGG motif, together with the two dsRBMs cooperate to dynamically determine NF90 RNA-binding ability (Schmidt et al., 2016). The two tandem dsRBMs are separated by a natively unstructured segment and they participate in RNA binding in two different ways. Nevertheless, both dsRBMs contribute to the binding of the same molecule simultaneously and co-operatively. dsRBM2 was shown to be the major determinant in the interaction with dsRNAs, while the involvment of dsRBM1 is minimal. Interestingly, the interaction of dsRBMs with the target RNA molecule largely recalls the RNA binding mode of ADAR2. This analogy suggests that, like for ADAR2, a sequence motif in the target dsRNA might influence NF90 binding (Jayachandran et al., 2016).

In contrast with this observation, it was shown that NF90 is able to recognize dsRNAs exclusively based on their structure. In particular, NF90 binding to the adenovirus-expressed VA1 RNA did not show any specificity for nucleotide sequence, upon extensive mutational analysis. On the other hand, the structure of the RNA seemed to be largely determinant of NF90 binding, with a requirement for a minihelix-like structure (Gwizdek et al., 2004). These discordant findings suggests that NF90 might recognize target RNAs based on either structure, sequence motif or both, possibly depending on the target itself.

The DZF domain is also conserved in NF45 and it is responsible for dimerization and consequent stabilization of the heterodimer. However, NF45 itself does not bind RNA efficiently and does not participate in RNA binding achieved by the heterodimer (Schmidt et al., 2017). It was shown that NF45 binding to NF90 is able to drive tertiary structural changes that result in an enhanced interplay between NF90 dsRBMs. Another similarity to RNA editing enzymes comes from the DZF

domains of NF90 and NF45 which present a nucleotidyltransferase fold, typical of the RNA modifying enzymes. However, both NF90 and NF45 have lost the critical catalytic residue and therefore are not functional RNA editing enzymes (Wolkowicz & Cook, 2012).

Post translational modifications and regulation of NF90

NF90 and NF110 contain a large number of residues that can be phosphorylated in both the common region and in the isoform-specific regions. Phosphorylation is one of the main translational modifications that can be found on NF90/NF110 and it has been shown to efficiently regulate NF90 compartimentalization and function. In fact, numerous residues were reported to be phosphorylated and involved in mRNA stabilization and in translational regulation (Harashima et al., 2010; Pei et al., 2008). Moreover, specific phosphorylation patterns are observed during mitosis, but their role is still unknown (N. L. Smith & Miskimins, 2011).

Besides phosphorylation, asymetric dimethylation of an arginine contained in RGG motif was also described for NF90. Arginine methylation in the RGG motif of RBPs is usually associated to RNA metabolism, but its role in NF90 function is still unclear (Rajyaguru & Parker, 2012).
Recently, it was shown that, in HCC, where high NF90 was found to be associated with bad prognosis, ubiquitin-specific protease 11 (USP11) promotes NF90 deubiquitination, thereby stabilizing it (C. Zhang et al., 2020).

As regards NF90 regulation, it was shown that lncRNA-Low Expression in Tumor (lncRNA-LET) is able to increase intracellular degradation of NF90 by changing its conformation and exposing ubiquitination sites (F. Yang et al., 2013). More recently, in esophageal squamous cell carcinoma, lncRNA-LET, which acts as tumor suppressor, was shown to be inhibited by highly expressed miR-548k (Z. Chen et al., 2018). These observations suggest that a negative feedback loop might take place where NF90 might potentially inhibit miR-548k leading to higher lncRNA-LET, which could in turn target NF90 itself (described more in detail later).

1.8.2 NF90 functions

NF90 is a polyvalent factor that, since its discovery, has been linked to a variety of functions such as transcriptional and translational regulation, viral replication and miRNAs biogenesis. Moreover, deregulation of NF90 was observed for several diseases such as cancer and muscle atrophy and it was implicated in the immune response, particularly against viruses. More recently, NF90 was also shown to regulate embryonic stem cell pluripotency and differentiation (Ye et al., 2017).
However, in recent years, NF90 was found to be particularly relevant for several types of cancers such as ovarian, breast, cervical, hepatocellular and nasopharyngeal carcinomas, leukemia and bladder cancer. While the impact of NF90 on cancer proliferation, progression, metastasis and drug resistance is evident, its role in these processed is divergent, and may depend on the cancer type. For instance, NF90 was shown to be a strong tumor suppressor for ovarian carcinoma while promoting proliferation and metastasis in hepatocarcinoma (Barbier et al., 2018; C. Zhang et al., 2020).

NF90 in transcription regulation

Although NF90 does not contain a known DNA-binding motif, evidence strongly suggests that it could have similar function to canonical DNA-binding proteins.
NF90/NF45 complex was originally described as a DNA-binding complex, acting as transcription factor for the cytokine interleukin 2 (IL2) during T-cells activation. In particular, it was shown that NF90/NF45 is able to bind the antigen recognition response element 2 (ARRE-2) contained in IL2 promoter and enhance *Il2* transcription (Corthésy & Kao, 1994).
Later, it was demonstrated that the interaction between the complex and target DNA is indirect, being mediated by several protein partners such as eIF2, Ku proteins and DNA-protein kinase (PK) (Shi et al., 2007; Ting et al., 1998). In line with its role in viral replication and T-cell activation, NF90 was also shown to regulate the transcription of another cytokine, IL13 (Kiesler et al., 2010), by binding to a DNase I hypersensitive site (DHS). On the other hand, transcription inhibition by NF90/NF45 was observed for the major histocompatibility complex class II

HLA-DR, mediated by DHS binding in B-cells (Sakamoto et al., 1999).

More recently, NF90/NF45 was reported to be involved in the increased transcription of *c-fos* gene upon serum induction. NF90/NF45 was found associated to chromatin and, in particular, to *c-fos* enhancer/promoter region while cooperating with general coativator factors (Nakadai et al., 2015). Similarly, ChIP-seq data in K562 erythroleukemia cells strongly suggests that NF90/NF110, by associating to promoter regions, significantly activates transcription factors that are known drivers of growth and proliferation (T.-H. Wu et al., 2018).
Therefore, NF90 seems to indirectly mediate transcription regulation, with a marked role in the immune response and cancer progression.

NF90 in the miRNA biogenesis pathway

The role of NF90 in the regulation of the miRNAs biogenesis pathway was recently described, with only few examples available.
The first evidence of the involvement of NF90/NF45 in the maturation of miRNAs comes from Sakamoto et al. (Sakamoto et al., 2009). In this study they showed that, in the nucleus, NF90/NF45 behaves like a negative regulator of Microprocessor activity for the maturation of pri-let-7a, competing with DROSHA binding on the pri-miRNA. Therefore, the maturation of pri-miRNAs to form pre-miRNAs is inhibited by the binding of NF90/NF45 complex, which impairs access of the Microprocessor on the pri-miRNAs.

Since this finding, different miRNAs have been shown to be modulated by NF90/NF45 binding activity. This complex was found to downregulate myogenic miRNAs, such as miR-133a, leading to significant loss and maturation of skeletal muscle and atrophy in NF90/NF45 double-transgenic mice (Todaka et al., 2015). More recently, NF90/NF45 was shown to inhibit the maturation of miR-7 in HCC. MiR-7 is a tumor suppressor miRNA and increased expression of NF90 leads to inhibition of its maturation followed by elevated proliferation rate in HCC (Higuchi et al., 2016). Therefore, NF90 was described as an oncogenic factor for HCC. Interestingly, the existence of a negative feedback loop between miR-7 and

NF90 was later shown, in which mature mir-7 is able to target the 3'UTR of NF90, leading to its translational repression (Higuchi et al., 2018).

However, the effect of NF90 on proliferation and metastasis seems to be different depending on the type of cancer. It was recently demonstrated by our team that NF90 is able to inhibit the maturation of miR-3173, a miRNA embedded in DICER first intron, by preventing the Microprocessor binding(Barbier et al., 2018). Furthermore, it was also shown that, in the absence of NF90, the level of pre-miR-3173 increases while DICER pre-mRNA exhibits splicing defects that lead to its downregulation. Increased progression and metastasis were observed in ovarian cancer cells. Therefore, it was established that NF90 can act as a tumor suppressor in ovarian cancer models. Interestingly, the mature form of miR-3173 is able to target NF90 mRNA by binding to its 3' UTR and leading to translational repression, mediating a feedback amplification loop that controls DICER expression and ovarian carcinoma progression.

NF90 was recently linked to chemotherapy resistance in bladder cancer. In particular, Gemcitabine treatment was shown to directly repress lncRNA-LET, leading to higher NF90 stability, which in turn inhibited maturation of miR-145 and enhanced the accumulation of chemotherapy-induced cancer stem-like cells (Zhuang et al., 2017).
With the exception of these few examples, the role of NF90 in regulating the biogenesis of miRNA is still largely unknown. For instance, it would be useful to understand the extent of NF90 effect on miRNAs maturation and whether this role is extensive or restricted to specific miRNA families or cellular pathways.

NF90 in mRNA translation, stability and degradation

In addition to controling mRNA fate by binding and modulating the processing of miRNAs, NF90 can directly regulate mRNAs translation, stability and decay. As previously discussed, AU-rich regions are frequently found in 3' UTRs of mRNAs and their recognition by RBPs is often responsible for their fate. Thanks to a ribonucleoprotein immunoprecipitation analysis, it was shown that NF90 is able to

bind a AU-rich, 25-to-30 nucleotide long signature motif (called NF90m) on a large subset of mRNAs (Kuwano et al., 2010). However, the post-transcriptional consequence of NF90 binding can vary depending on the target mRNA or on the condition studied. For example, NF90 binding to 3' UTRs is able to both increase the stability of mRNAs and positively or negatively regulate their translation. As regards translation regulation, NF90 can inhibit or activate translation, depending on the mRNA target, by affecting the initiation step or by retaining target mRNAs in the nucleus (Castella et al., 2015).

For instance, insertion of the AU-rich NF90m on a reporter gene did not affect mRNA stability but rather inhibited its translation by preventing its association with actively translating ribosomes (Kuwano et al., 2010). Along the same lines, following viral infection, NF90 was shown to be phosphorylated by PKR which leads to its dissociation from NF45 and export from the nucleus. Phosphorylated NF90 accumulates on ribosomes where it associates with viral RNA, inhibiting their translation (Harashima et al., 2010). On the other hand, NF90 was shown to activate translation of a limited subset of mRNAs, such as VEGF and cyclin T1 mRNA. Under hypoxic conditions, NF90 is able to interact with the 3' UTR stem-loop hypoxia stability region in VEGF mRNA promoting its loading onto polysomes and increasing its stability (Vumbaca et al., 2008). After HIV infection, NF90 promotes its replication and latency by binding to cyclin T1 mRNA 3' UTR and facilitating the recruitment of translation initiation factors (Hoque et al., 2011).

On a different note, NF90 is also able to increase the stability of its target mRNAs. For instance, after T cell activation, NF90 is phosphorylated by AKT and translocates to the cytoplasm where it can bind the 3'UTR of *Ilf2* mRNA stabilizing it and inhibiting its degradation (Pei et al., 2008). Similarly, it was shown that NF90, after its translocation to the cytoplasm mediated by cyclin-dependent kinase (CDK) 2 phosphorylation, promotes HCC proliferation by stabilizing and upregulating cyclin E1 mRNA (Ding et al., 2020; Jiang et al., 2015). In HCC, NF90 was also shown to regulate the stabillity of PARP1 mRNA and, therefore, enhance tumor development (D. Song et al., 2017).

More recently, several studies suggested the existence of signaling pathways or regulatory axis involving NF90 and, often, miRNAs that highly impact on the development of pathologies and especially cancer progression and metastasis (J. Lin et al., 2019; Wen et al., 2018; Y. Zhang et al., 2020; Zhuang et al., 2017). For instance, NF90 was shown to suppress the cytokine B-cell activating factor (BAFF) mRNA translation by recruiting miR-15a to BAFF 3' UTR. Moreover, a variant of BAFF mRNA (BAFF-*var*), that lacks the NF90 binding site, was found to have enhanced translation and it is associated with an elevated risk of developing autoimmune diseases. Interestingly, it was shown that the binding of NF90 and miR-15 to the mRNA correlates with higher AGO2 binding, which suggests that translational inhibition of BAFF mRNA might be mediated by RISC (Idda et al., 2018).

Despite the numerous examples of mRNAs regulated by NF90 at the post-transcriptional level, the exact mechanism that results in translation inhibition/activation or mRNA stabilization is still largely unknown.

NF90 in viral replication

Besides acting as a cellular mRNA binding factor that controls RNA metabolism and translation, NF90 also binds viral RNA or DNA (Pfeifer et al., 2008). Moreover, the consequence of its binding can vary, supporting or inhibiting viral replication and viral genome expression, depending on the type of virus (Patiño et al., 2015). Numerous viruses were shown to exploit NF90 to support their amplification, such as hepatitis C virus (HCV), HIV, human papilloma virus (HPV) and Dengue virus (DV) (Gomila et al., 2011; Y. Li & Belshan, 2016; Y. Li et al., 2014; Shamanna et al., 2013). On the other hand, NF90 is able to act as a host antiviral factor for other types of viruses such as influenza A virus (IAV) and Ebola virus (EBOV) (T. Li et al., 2016; Shabman et al., 2011). For instance, NF90 was shown to bind the 5'-terminal sequence of HCV RNA genome upon infection and promote HCV replication by possibily associating with the replication complex (Y. Li et al., 2014). On the same lines, upon HIV infection, NF90 shows a pleiotropic effect stimulating the viral gene expression but also stabilizing HIV RNA (Y. Li & Belshan, 2016). Contrarily,

NF90 was found to suppress Ebola virus replication by associating with the viral protein VP35 and impairing the function of EBOV replication complex (Shabman et al., 2011).

It was recently shown that viral infection promotes NF90 translocation from the nucleus to the cytoplasm. In particular it was found that NF90 is a key factor for the biogenesis of circRNAs and these molecules might act as a reservoir of NF90 in the cytoplasm, to avoid off-targets effect. In response to a viral infection, NF90 translocates into the cytoplasm reducing circRNAs expression while cytoplasmic NF90 is able to bind viral RNAs (X. Li et al., 2017).
Despite the numerous and diversified examples of NF90 acting during the viral response, the exact mechanism underlying its role is yet to be fully elucidated. However, it is possible that the complicated contribution of NF90 to antiviral immunity might be exploited through different mechanisms depending on the type of virus.

To conclude, while the miRNA pathway is a well-studied mechanism, its regulation by several RBPs is still largely unknown. The role of NF90 in the miRNA biogenesis pathway was recently brought to light but the extent of its function is yet unexplored.

Chapter 2

NF90 modulates processing of a subset of human pri-miRNAs

Giuseppa Grasso[1], Takuma Higuchi[2], Victor Mac[1], Jérôme Barbier[1], Marion Helsmoortel[1], Claudio Lorenzi[3], Gabriel Sanchez[1], Maxime Bello[1], William Ritchie[3], Shuji Sakamoto[2] and Rosemary Kiernan[1]*

[1] UMR9002 CNRS-UM, Institut de Génétique Humaine-Université de Montpellier, Gene Regulation lab, Montpellier, 34396, France

[2] Laboratory of Molecular Biology, Science Research Centre, Kochi Medical School, Kochi University, Kochi, 783-8505, Japan

[3] UMR9002 CNRS-UM, Institut de Génétique Humaine-Université de Montpellier, Artificial Intelligence and Gene Regulation lab, Montpellier, 34396, France

* To whom correspondence should be addressed. Tel: 33 4 34359939; Fax: 33 4 34359901; Email: Rosemary.Kiernan@igh.cnrs.fr

Abstract

MicroRNAs (miRNAs) are predicted to regulate the expression of more than 60% of mammalian genes and play fundamental roles in most biological processes. Deregulation of miRNA expression is a hallmark of most cancers and further investigation of mechanisms controlling miRNA biogenesis is needed. The double stranded RNA-binding protein, NF90 has been shown to act as a competitor of

Microprocessor for a limited number of primary miRNAs (pri-miRNAs). Here, we show that NF90 has a more widespread effect on pri-miRNA biogenesis than previously thought. Genome-wide approaches revealed that NF90 is associated with the stem region of 38 pri-miRNAs, in a manner that is largely exclusive of Microprocessor. Following loss of NF90, 22 NF90-bound pri-miRNAs showed increased abundance of mature miRNA products. NF90-targeted pri-miRNAs are highly stable, having a lower free energy and fewer mismatches compared to all pri-miRNAs. Mutations leading to less stable structures reduced NF90 binding while increasing pri-miRNA stability led to acquisition of NF90 association, as determined by RNA Electrophoretic Mobility Shift Assay (EMSA). NF90-bound and downregulated pri-miRNAs are embedded in introns of host genes and expression of several host genes is concomitantly reduced. These data suggest that NF90 controls the processing of a subset of highly stable, intronic miRNAs.

Introduction

MicroRNAs (miRNAs) are short non-coding RNAs that negatively regulate the expression of a large proportion of cellular mRNAs, thus affecting a multitude of cellular and developmental pathways (Ebert & Sharp, 2012; Shenoy & Blelloch, 2014). The canonical miRNA biogenesis pathway involves two sequential processing events catalysed by RNase III enzymes. In the nucleus, the microprocessor complex, comprising the RNase III enzyme Drosha, the double-stranded RNA-binding protein, DGCR8 and additional proteins carries out the first processing event, which results in the production of precursor miRNAs (pre-miRNAs) (Denli et al., 2004; Gregory et al., 2004). These are exported to the cytoplasm, where a second processing event is carried out by another RNase III enzyme, DICER, leading to the production of miRNA duplexes. The duplexes are loaded into the RISC complex and the release of the 'passenger' strands leads to the formation of mature miRNAs and mature RISC complexes (Ha & Kim, 2014).

Due to the central role of miRNAs in the control of gene expression, their levels must be tightly controlled. Indeed, deregulation of miRNA expression is associated

with aberrant gene expression and leads to human disease (Finnegan & Pasquinelli, 2013; Krol, Loedige, et al., 2010; Mendell & Olson, 2012; Winter et al., 2009). Consequently, miRNA biogenesis is tightly regulated at multiple steps, both transcriptional and post-transcriptional. Increasing evidence suggests that RNA binding proteins (RBPs) act as post-transcriptional regulators of miRNA processing. Many RBPs modulate the processing efficiency of Microprocessor, either positively or negatively, by binding to regions of the pri-miRNA. A number of RBPs have been shown to bind the terminal loop, which can either facilitate or inhibit cropping by Microprocessor. For example, LIN28B binds the terminal loop of pri-let-7, which prevents its processing by Microprocessor (Viswanathan et al., 2008). Binding of hnRNP A1 to the terminal loop has been shown to exert either positive or negative effects on Microprocessor activity, depending on the pri-miRNA target. It promotes cropping of pri-miR-18A while it inhibits processing of pri-let-7. KSRP is another terminal loop-binding RBP that facilitates Microprocessor cleavage of several pri-miRNA targets, including pri-let-7 where it acts as a competitor of hnRNP A1 (Guil & Cáceres, 2007; Michlewski et al., 2008).

Several other RBPs, including SMAD, TPD-43, SRSF1 and RBFOX, have been shown to bind pri-miRNA terminal loops to influence Microprocessor activity (see Michlewski and Cáceres, 2019 for review). In most cases, they have been shown to bind specific pri-miRNAs, such as pri-let-7, or a limited subset of pri-miRNAs. To date, only NF90/NF45 heterodimer and ADAR1,2 have been shown to bind the double stranded stem region of pri-miRNAs (Higuchi et al., 2016; Sakamoto et al., 2009; W. Yang et al., 2006). Both factors negatively affect Microprocessor activity. Indeed, NF90 has been shown to bind double stranded RNA in a mode similar to that of ADAR2 (Jayachandran et al., 2016). Like terminal loop binding RBPs, binding of NF90/NF45 or ADAR1,2 has thus far been demonstrated for a very limited number of pri-miRNAs. NF90 has been shown to associate with pri-miR-7-1, pri-let-7A and pri-miR-3173 in human cells (Barbier et al., 2018; Higuchi et al., 2016; Sakamoto et al., 2009).

We have previously shown that NF90 associates with pri-miR-3173, which is

located in the first intron of DICER pre-mRNA (Barbier et al., 2018). Binding of NF90 prevented cropping of pri-miR-3173 by Microprocessor and promoted splicing of the intron, thereby facilitating expression of DICER. By modulating DICER expression, NF90 was found to be an independent prognostic marker of ovarian carcinoma progression (Barbier et al., 2018). Levels of NF90 are known to be elevated in hepatocellular carcinoma (HCC) and the effect of NF90 on processing of pri-miR-7-1 contributes to cellular proliferation in HCC models (Higuchi et al., 2016; Jiang et al., 2015).

Here, we have used genome-wide approaches to identify pri-miRNAs that are associated with and modulated by NF90 in HepG2 model of HCC. We identified 38 pri-miRNAs that are associated with NF90, in a manner that is for the most part exclusive of Microprocessor. Of these, 22 showed increased abundance of mature miRNAs products upon loss of NF90. NF90-targeted pri-miRNAs appear to be highly stable, having a lower free energy and fewer mismatches compared to all pri-miRNAs. Destabilization of the structures by mutation reduced NF90 association as determined by RNA EMSA. Of the 22 NF90-modulated pri-miRNAs, 20 are embedded exclusively in introns of host genes. Transcriptomic analysis revealed that the expression of the host gene is concomitantly downregulated for several, including an oncogene implicated in metastasis of hepatocellular carcinoma, TIAM2. These data suggest that NF90 controls the processing of a subset of intronic miRNAs, which in some cases affects the expression of the host gene.

Material and methods

Cell culture

Human HepG2 cell line was grown in Dulbecco's Modified Eagle's medium – high glucose (Sigma-Aldrich®, D6429) supplemented with 10% fetal bovine serum (PAN Biotech, 8500-P131704), 1% penicillin-streptomicin (v/v) (Sigma Aldrich®, P4333) and 1% L-glutamine (v/v) (Sigma Aldrich®, G7513). Human HEK-293T cells were grown in Dulbecco's Modified Eagle's high glucose medium with HEPES

(Sigma-Aldrich®, D6171) supplemented with 10% fetal bovine serum, 1% penicillin-streptomicin and 1% L-glutamine. Cells were cultured at 37°C in a humidified atmosphere containing 5% CO_2. To perform small RNA-seq and RNA-seq, HepG2 were seeded at 1.5×10^6 cells in 6-well plates the day of siRNA transfection while HEK-293T were seeded at 6×10^5 cells in 6-well plates. To perform RNA Immunoprecipitation, HepG2 were seeded at 8×10^6 cells in 100 mm culture dishes the day of siRNA transfection.

Transfection of small interfering RNAs

Double-stranded RNA oligonucleotides used for RNAi were purchased from Eurofins MWG Operon or Integrated DNA Technologies. Sequences of small interfering RNAs (siRNAs) used in this study have been described previously (Barbier et al., 2018) and are shown in Supplementary Table S1.

HepG2 or HEK-293T cells were transfected with siRNA (30 nM final concentration) using INTERFERin® siRNA transfection reagent (Polyplus Transfection) according to the manufacturer's instructions. To perform small RNA-seq and RNA-seq, two rounds of transfection were performed. The first transfection was carried out the day of seeding; on the fourth day cells were passaged and a second round of transfection was performed. Cells were collected for RNA extraction or protein purification approximately 65 h after the second transfection. To perform RNA Immunoprecipitation, one round of siRNA transfection was carried out, as explained, the day of seeding. Cells were collected approximately 65 h after siRNA transfection.

Immunoblot

HepG2 were lysed using RIPA buffer (50 mM Tris-HCl pH=7.5, 150 mM NaCl, 1% NP40, 0.5% Sodium Deoxycholate, 0.1% SDS, Halt™ Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific)). Protein extracts (30 mg for NDUFS8, 50 mg for TIAM2 and 5 mg for all other proteins) were immunoblotted using the indicated

primary antibodies (Table S2) and anti-mouse, anti-rabbit or anti-rat IgG-linked HRP secondary antibodies (GE Healthcare) followed by ECL (Advansta).

Small RNA-seq and RNA-seq

Total RNA was extracted using TRIzol (Thermo Fisher Scientific) according to the manufacturer's instructions. Small RNA-seq (single end, 50 bp) was carried out by BGI Genomic Services (HepG2) or Fasteris (HEK-293T) in triplicate samples. Raw data were processed using the Subread package (version 1.6.0) as previously described (Seco-Cervera et al., 2018) and the reference annotation was obtained from miRBase release 22.1 database (Kozomara et al., 2019). Statistical analysis was performed using DESeq2 (version 2.11.40.2). RNA-seq (paired-end, 125 bp) was carried out by BGI Genomic Services in triplicates. Raw data were processed using HISAT2 (version 2.1.0) and featureCounts (version 1.6.3), statistical analysis was performed using DESeq2. Reference annotation was obtained from ENSEMBL (GRCh38.96).

RT-qPCR, Modified 5′ RLM RACE and RNA EMSA

Total RNA was extracted from HepG2 cells using TRIzol reagent (Thermo Fisher Scientific) and RNA was treated with DNAse I (Promega) according to the manufacturer's instructions. RNA was used for RT-PCR and modified 5′ RLM-RACE as described previously (Barbier et al., 2018).

For RT-qPCR, RT was performed using TaqMan™ Reverse Transcription Reagent or TaqMan™ Advanced miRNA cDNA Synthesis Kit (Thermo Fisher). qPCRs were performed using GoTaq® Probe qPCR Master Mix (Promega) or TaqMan® Fast Advanced Master Mix (Thermo Fisher).

Modified 5′ RLM RACE was performed according to the manufacturer's instructions (FirstChoice™ RLM-RACE kit, ThermoFisher Scientific). In order to detect premature miRNAs, the step using Calf Intestine Alkaline Phosphatase was omitted. Sequences of the primers used for PCR amplification are shown in Supplementary Table S3.

RNA EMSA was performed as described previously (Sakamoto et al., 2009) using

recombinant NF90 and recombinant DGCR8 dsRBDs (amino acids 484-773) in at least three replicates. The pri-miRNA probes were amplified by PCR using the primers shown in Supplementary Table S3. Sequences of mutant pri-miRNAs are shown in Supplementary Table S4.

RNA ImmunoPrepicipation (RIP)

RIP was performed as previously described (Bennasser et al., 2011). HepG2 were seeded in 100 mm culture dishes and transfected with siRNAs the day of seeding as aforementioned. Cells were harvested approximately 65 h after the treatment and lysed for 15 min in RIP buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 2.5 mM $MgCl_2 \bullet 6H_2O$, 250 mM sucrose, 0.05% (v/v) NP-40 and 0.5% (v/v) Triton X-100) containing 20 U ml^1 of RNasin (Promega), 1mM DTT, 0.1 mM PMSF and EDTA-free protease and phosphatase inhibitor. After centrifugation, lysates were incubated for 4 h at 4°C with 2 µg of antibodies recognizing NF90, Drosha and IgG control and then incubated for 1 h at 4°C with Dynabeads™ Protein A (ThermoFisher Scientific). After incubation, beads were washed 5 times with RIP buffer for 5 min at 4°C and RNA was extracted as previously explained. RNA was treated with DNAse I (Promega) and RT was performed using SuperScript™ III Reverse Transcriptase (ThermoFisher Scientific) according to the manufacturer's instructions. cDNA was treated with RNAse H (ThermoFisher Scientific) and the samples were used to perform qPCRs using QuantiTect SYBR® Green PCR Kit (Qiagen) according to the manufacturer's instructions.

Splicing analysis

Splicing analyses were carried out as previously described (Barbier et al., 2018). HepG2 were seeded in 6-well plates and transfected with siRNAs, as aforementioned. Approximately 65 h after the second transfection, RNA was extracted using TRIzol reagent (Thermo Fisher Scientific) and treated with DNAse I (Promega) according to the manufacturer's instructions. RT was performed using SuperScript™ III Reverse Transcriptase (ThermoFisher Scientific) and cDNA was treated with RNAse H (ThermoFisher Scientific). qPCRs were performed using

QuantiTect SYBR® Green PCR Kit (Qiagen) using primers overlapping exon-intron boundaries to detect unspliced pre-mRNAs or primers amplifying exon-exon boundaries to detect the spliced mRNA.

Bioinformatic analyses

Enhanced UV crosslinking followed by immunoprecipitation (eCLIP) data for NF90, DGCR8 and DROSHA obtained in HepG2 cells by Nussbacher and Yeo (Nussbacher & Yeo, 2018) were retrieved from the NCBI database (NF90 eCLIP: ENCSR786TSC; DGCR8 eCLIP: ENCSR061SZV; DROSHA eCLIP: ENCSR834YLD). Peaks were filtered based on Fold Change (FC 1.5) and p-value (Bonferroni-Adj P-val 0.05). Distribution of eCLIP reads along the miRNAs was evaluated using deeptools software (version 3.1.3). Bigwig files from different replicates were merged using bigWigMerge v2. The base pair probability at each position of miRNA hairpins was calculated using RNAfold software (version 2.4.7).
Free energy analysis was performed using RNAfold software, version 2.4.7. Statistical analysis was performed using R (version 3.5.1).
Validated targets of the double positive miRNAs were extracted from MirTarBase database, release 7.0 (Chou et al., 2018). Gene ontology was performed on the expressed validated target using DAVID Functional Annotation Tool database version 6.8 (https://david.ncifcrf.gov) (D. W. Huang et al., 2009). Motif search was performed using MEME (version 5.0.5).

Results

NF90 affects the abundance of a subset of human miRNAs

To determine the effect of NF90 on the abundance of miRNAs, we performed small RNA-seq of biological triplicate samples obtained from HepG2 cells that had been transfected with a non-targeting control siRNA (siScr) or an siRNA targeting NF90 (siNF90) (Figure 1A, top panel). Of 1917 miRNA precursors annotated in miRBase, 1105, which corresponds to 1661 mature 5p and 3p miRNA products, were found to be expressed in HepG2 cells. Following loss of NF90, differential expression analysis (fold change 1.5 or 0.667; AdjP-value 0.05) showed that 268 mature

miRNAs, corresponding to 212 precursor miRNAs, were upregulated while 149, corresponding to 126 precursor miRNAs, were downregulated (Figure 1B). The number of upregulated and downregulated miRNAs in HepG2 cells after loss of NF90 is summarized in Figure 1C. MiRNAs that have previously been shown to be repressed by NF90, miR-7-1 (Higuchi et al., 2016) and miR3173 (Barbier et al., 2018), were found to be upregulated in HepG2 cells following loss of NF90 (Figure 1B, red dots).

The effect of NF90 on the abundance of miRNAs observed by miRNA profiling were validated by RT-qPCR analysis of selected miRNAs, miR-3173-3p, miR-186-5p, miR-1273c and miR-3189-3p, from biological triplicate samples. The results obtained confirmed the effects observed by miRNA profiling (Figure 1B, 1D). In addition, RNA was extracted from cells transfected with an independent non-targeting siRNA (Scr2) and an NF90-targeting siRNA (NF902) that has been described previously (Barbier et al., 2018) (Figure 1A, lower panel). Quantification of miRNAs 3173-3p, -186-5p, -1273c and -3189-3p in biological triplicate samples (Figure 1D, lower panels) showed similar results to those obtained in Figure 1D upper panel, and also validated the results obtained by small RNA-seq. While we cannot exclude the possibility that a proportion of the small RNA-seq results could be due to off-target effects of the siRNAs, since only a single control and NF90-targeting siRNA were used, validation of a subset of the results using additional control and NF90-targeting siRNA suggests that the data are, to some extent, robust.

To evaluate whether the effect of NF90 on miRNA abundance might be cell type specific, we performed small RNA-seq in biological triplicate in HEK-293T cells transfected with control or NF90-targeting siRNA (Supplementary Figure S1A). Of 1917 annotated miRNA precursors, 1121, corresponding to 1647 mature miRNAs, were expressed in HEK-293T. Differential expression analysis (fold change 1.5 or 0.667; AdjP-value 0.05) revealed that 278 mature miRNAs, corresponding to 217 miRNA precursors, were upregulated following loss of NF90 while 84 mature miRNAs, corresponding to 77 precursors, were downregulated (Supplementary

Figure S1B, S1C). Comparing upregulated miRNAs in the two cell types, we found 139 miRNAs that were upregulated in both cell lines after NF90 knock-down (Supplementary Figure S1D). This represents more than 65% of miRNAs upregulated in HepG2 and 64% of those upregulated in HEK-293T. Thus, NF90 appears to regulate a common subset of miRNAs.

NF90 associates with a subset of pri-miRNAs

To determine which of the miRNAs upregulated upon loss of NF90 (Figure 1B) are direct targets of NF90, that is, pri-miRNAs that are bound by NF90, we took advantage of enhanced UV crosslinking followed by immunoprecipitation (eCLIP) dataset obtained in HepG2 cells (Nussbacher & Yeo, 2018). Analysis of HepG2 eCLIP data revealed 38 pri-miRNAs for which eCLIP peaks overlapped annotated pri-miRNA localizations +/- 25 nt of flanking region (FC 1.5 and Bonferroni AdjP 0.05), as depicted in Figure 2A and Supplementary Table S5. Pri-miR-3173 and pri-miR-7-1 were among the 38 NF90-associated pri-miRNAs (Figure 2A, red dots).

We next analysed eCLIP read coverage across the pri-miRNA hairpin +/- 200 bp for the 38 NF90-associated miRNAs compared to all pri-miRNAs (Figure 2B). As expected, analysis of all pri-miRNAs did not show significant read coverage for NF90 association. In contrast, NF90-associated miRNAs showed highest read coverage over the region having the strongest base pair probability and therefore likely corresponding to the double stranded pri-miRNA stem (Figure 2B). The region corresponding to the terminal loop, which has a low base pair probability, was not significantly bound by NF90. Interestingly, NF90 also appeared to bind to the pri-miRNA flanking region. Browser shots showing NF90 association with pri-miR-7-1, pri-miR-186 and pri-miR-1273c by eCLIP are shown in Supplementary Figure S2A.

To validate NF90 association with pri-miRNAs identified by eCLIP analysis (Figure 2A), we performed RNA EMSA using pri-miR-186, pri-miR-3173, pri-miR-1273c and pri-miR-3189 as radiolabelled probes together with recombinant NF90 (Figure S2B), as described previously for pri-miR-7-1 and pri-miR-3173 (Barbier et al., 2018;

Higuchi et al., 2016). RNA EMSA, performed in triplicate, confirmed NF90 association with pri-miR-186, pri-miR-3173, pri-miR-1273c and pri-miR-3189 (Figure 2C and S2C). Similarly, RNA EMSA confirmed that NF90 was not highly associated with pri-miR-200a, as indicated by eCLIP (Figure 2C). NF90 association with the pri-miRNAs identified by eCLIP analysis was also validated for several endogenous pri-miRNAs by performing RNA immunoprecipitation (RIP). RIP confirmed the association of NF90 with region proximal to the endogenous pri-miRNA (Figure 2D and S2D), while negative controls, pri-miR-200a and DALRD3, were not significantly associated with NF90. In contrast, pri-miR-200a was significantly bound by Drosha (Figure 2D and S2D). While not all NF90-bound pri-miRNAs identified by eCLIP have been tested, RIP analysis confirmed the association with NF90 in vivo for at least several.

Previous studies have indicated that NF90 may act as a competitor of Microprocessor for binding to pri-miRNAs (Barbier et al., 2018; Higuchi et al., 2016; Michlewski & Cáceres, 2019; Sakamoto et al., 2009). We therefore analysed eCLIP data for DGCR8 and Drosha performed in HepG2 cells (Nussbacher & Yeo, 2018). Association of DGCR8 was detected at 203 pri-miRNAs, while 147 pri-miRNAs were positive for Drosha binding (Figure 3A). Not surprisingly, there was a significant overlap between pri-miRNAs that were bound by both subunits of Microprocessor (Figure 3A). Indeed, 125 pri-miRNAs were associated with both factors, which represents approximately 60% and 85% of pri-miRNAs positive for DGCR8 and Drosha, respectively. Interestingly, only 10 pri-miRNAs bound by NF90 overlapped with those bound by either DGCR8 or Drosha, which represents approximately 24% overlap with DGCR8 and 13% overlap with Drosha (Figure 3A). This result indicates that NF90-associated pri-miRNAs are not highly associated with Microprocessor. Analysis of eCLIP reads showed association of DGCR8 with both apical and stem regions of pri-miRNAs (Figure 3B), as expected (Nguyen et al., 2015).

We further analysed eCLIP read coverage over the pri-miRNAs that were found to be associated with both NF90 and DGCR8. While the profile for DGCR8 was

similar to that for all DGCR8 positive pri-miRNAs (compare Figure 3C, top panel to Figure 3B), the profile for NF90 read coverage was somewhat different to that for all NF90-positive pri-miRNAs (compare Figure 3C, lower panel, to Figure 2B). Interestingly, for pri-miRNAs that are bound by both DGCR8 and NF90, the profiles appear to be complementary (Figure 3C, compare top and lower panels). Plot profiles of DROSHA and DGCR8 eCLIP data suggest that pri-miRNAs common with NF90 (shown with red dots) are not among the most enriched for Microprocessor binding (Figure 3D and Supplementary Figure S3).

To further explore the competition between NF90 and the Microprocessor for the binding of pri-miRNAs, we performed RNA EMSA on pri-miR-3189 and pri-miR-1273c using recombinant NF90 and the dsRNA-binding domains of DGCR8 (Figures S2B and S4A). Upon addition of rNF90, a shift corresponding to the formation of NF90-pri-miRNA complex and a reduction in the intensity of the band corresponding to DGCR8-pri-miRNA complex could be detected (Figure 4A). These results indicate that NF90 competes with Microprocessor for binding to certain pri-miRNAs, at least in vitro. Further analysis will be required to determine whether this competition also occurs in vivo. We next tested whether loss of NF90/NF45 or Drosha/DGCR8 complexes could affect the binding of the complexes to endogenous pri-miRNAs in vivo. We performed RIP of NF90, Drosha or IgG control after downregulation of either NF90/NF45 or Drosha/DGCR8. Drosha association with the region surrounding the target pri-miRNAs was significantly enhanced after downregulation of NF90/NF45, while NF90 association was significantly enhanced after downregulation of Drosha/DGCR8 only for pri-miR-1273c (Figure 4B and S4B). This could be explained considering that these miRNAs are already poorly bound by the Microprocessor. To test this hypothesis, we analysed the association of NF90 to two pri-miRNAs poorly bound by NF90, pri-miR-200a and pri-miR-425. Notably, NF90 association with these miRNAs was significantly increased after loss of Drosha/DGCR8 complex (Figure S4C). On the other hand, downregulation of NF90/NF45 complex did not significantly affect the association of pri-miR-200a and pri-miR-425 with Drosha (Figure S4C), possibly because these miRNAs are poorly bound by NF90/NF45

under control conditions. Taken together, these results suggest that target pri-miRNAs may have binding preferences for either NF90/NF45 or Microprocessor under wild-type conditions, but that the relative abundance of these complexes can also influence the observed binding to specific pri-miRNAs.

Pri-miRNAs that are bound and downregulated by NF90 are highly stable

We next asked whether NF90 association with pri-miRNAs might affect their cropping by Microprocessor. If so, loss of NF90 would be predicted to increase the abundance of the mature miRNA products, as observed previously (Barbier et al., 2018; Higuchi et al., 2016; Sakamoto et al., 2009). MiRNA profiling revealed that of the 38 NF90-associated pri-miRNAs, 22 showed an increase in mature miRNA products, representing more than 57% of NF90-associated pri-miRNAs, while only 2 were decreased (Supplementary Tables S6 and S7). Thus, we identified a subset of 22 pri-miRNAs that are bound by NF90 and whose abundance is increased following loss of NF90, which we named 'double-positive' pri-miRNAs. Both pri-miR-7-1 and pri-miR-3173 were identified within the double positive subset. Thus, NF90 downregulates the expression of most of its target pri-miRNAs.

Gene ontology of validated mRNA targets of double positive miRNAs revealed an implication particularly in cancer and infection by viruses, such as Epstein Barr Virus (EBV), hepatitis B virus (HBV), and human T lymphoma virus type 1 (HTLV1), as well as viral carcinogenesis (Supplementary Figure S5). This result is interesting given that NF90 translocates from the nucleus to the cytoplasm following viral infection of cells (X. Li et al., 2017). Thus, viral infection could result in the coordinated processing of the NF90-modulated subset of pri-miRNAs, whose target mRNAs are implicated in viral replication. Interestingly, several miRNAs upregulated following loss of NF90 in this study have been shown to target RNAs expressed by influenza A virus subtypes. For instance, miR-3682 is involved in viral replication by targeting the NS gene of pH1N1 and H3N2 subtypes (X. Zhang et al., 2018). Similarly, miR-4753 and miR-3145, which target PS and PB1 genes of H5N1 and H3N2 subtypes, are overexpressed in response to viral infection and inhibit viral transcription and replication (Khongnomnan et al., 2015).

We wondered whether pri-miRNAs that are associated with NF90 and downregulated upon its loss might share a common characteristic that would make them targets for NF90 binding. A MEME search did not reveal a simple binding motif common to the 22 pri-miRNA sequences. Compared to all human pri-miRNAs, the subset of 22 double-positive pri-miRNAs did not show any significant difference in their overall length (mean=82.5 nt compared to 81.88 nt) or in the size of the terminal loop (mean=7.87 nt compared to 7.92 nt) (Figure 5A). In contrast, however, the minimal stretch containing a mismatch 1 nt was significantly longer for double-positive pri-miRNAs compared to all pri-miRNAs, with a mean of 27.68 nt for double-positive pri-miRNAs compared to 21.11 nt for all pri-miRNAs (Figure 5A). This analysis suggests that double-positive pri-miRNAs might be more stable, having a longer duplex and less bulges compared to all human pri-miRNAs. To further investigate this possibility, we compared the free energy of the 22 double-positive pri-miRNAs compared to all pri-miRNAs. The 22 double-positive pri-miRNAs had a lower free energy (mean=-42.26) compared to all pri-miRNAs (mean=-38.19), as shown in Figure 5B. Taken together, these data suggest that double positive pri-miRNAs are more stable and have less mismatches than all pri-miRNAs. Predicted folding of double-positive pri-miRNA sequences also revealed highly stable structures with very few bulges, compared to pri-miR-200a, which is not highly associated with NF90 (Supplementary Figure S6A).

To test the idea that NF90 can bind to pri-miRNAs that have a stable structure with few bulges, we designed mutations within NF90-binding pri-miRNAs predicted to reduce stability and form bulge-like regions that might disrupt NF90 association. For each of the NF90-associated pri-miRNAs tested, we designed two mutant structures that would be less stable than wild-type structures. (Figure 6A). WT and mutated pri-miRNAs were tested for NF90 association by RNA EMSA. As shown in Figure 6B and Figure S6B, mutation of pri-miR-3173 or pri-miR-186 to less stable structures diminished NF90 binding. On the other hand, mutation of pri-miR-200a to a more stable structure enhanced NF90 binding. These data suggest that NF90 shows a preference for association with stable pri-miRNA hairpin structures having

few bulge regions.

We then wondered whether pri-miRNAs whose mature products increased following loss of NF90, but were not considered eCLIP-positive using the applied cut-offs, might share the characteristics identified for double-positive pri-miRNAs. We therefore calculated the longest duplex length, allowing a mismatch of 1 nt, for the group of 181 upregulated but eCLIP negative pri-miRNAs, and 124 downregulated pri-miRNAs, as well as for those falling outside these groups (other) (Figure 7A). Interestingly, pri-miRNAs upregulated after loss of NF90 and eCLIP negative have a significantly longer duplex than all pri-miRNAs or other pri-miRNAs. Indeed, the duplex length is similar to that observed for the double positive group. In contrast, pri-miRNAs downregulated upon loss of NF90 have a shorter duplex compared to all pri-miRNAs or other pri-miRNAs. We then calculated the mean free energy for the upregulated, eCLIP-negative group and the downregulated group of pri-miRNAs (Figure 7B).

Similarly, when compared to all pri-miRNAs, the upregulated, eCLIP-negative group of pri-miRNAs had a significantly lower free energy. Free energy of the downregulated group was similar to that of all pri-miRNAs. In contrast, terminal loop size was comparable between the 2 groups; 7.86nt (downregulated group) compared with 8.64 nt (upregulated eCLIP-negative group). Of note, total pri-miRNA length was higher for the upregulated eCLIP-negative group (87.01 nt) compared to the downregulated group (77.79 nt). These analyses suggest that upregulated, eCLIP-negative pri-miRNAs share some characteristics with double-positive pri-miRNAs. It is feasible that some NF90-associated pri-miRNAs were not detected by eCLIP analysis or did not pass the selection criteria used to identify eCLIP-positive pri-miRNAs. To test this idea, we selected 2 pri-miRNAs, pri-miR-4755 and pri-miR-4766, from the upregulated, eCLIP-negative group whose structure corresponds to the defined criteria for NF90 association, that is, having low free energy and few mismatches (Supplementary Figure S7A). NF90 binding to the pri-miRNAs was tested by RNA EMSA (Figure 7C and Figure and S7B). Indeed, both pri-miR-4755 and pri-miR-4766 were found to be significantly

associated with NF90.

NF90 modulates the expression of a subset of genes hosting NF90-associated pri-miRNAs

Approximately 70% of human miRNAs are located in an intron of a host gene. Out of 22 double-positive pri-miRNAs, 20 are exclusively intronic. Two double-positive pri-miRNAs are found in either the 3′ UTR or an intron depending on transcript usage (Supplementary Table S6).

To determine whether loss of NF90 also affected the expression or splicing efficiency of the host genes, we performed RNA-seq in HepG2 cells transfected with control siRNA or siRNA targeting NF90. Loss of NF90 significantly diminished expression of 3 genes containing NF90-associated pri-miRNA; growth differentiation factor 15 (GDF15) hosting pri-miR-3189, 1-acylglycerol-3-phosphate O-acyltransferase 5 (AGPAT5) hosting pri-miR-4659a and zinc finger RAN-binding domain containing 2 (ZRANB2) hosting pri-miR-186 (Figure 8A). Furthermore, the splicing efficiency of introns containing pri-miRNAs downregulated by loss of NF90 was determined by RT-PCR for several targets (Figure 8B). Splicing efficiency was diminished for 3 pre-mRNAs containing NF90-associated pri-miRNAs: T-cell lymphoma invasion and metastasis 2 (TIAM2), hosting pri-miR-1273c, Zinc Finger RNA binding protein (ZFR), hosting pri-miR-579, and DICER, hosting pri-miR-3173 (Figure 8B). Interestingly, the splicing defect was detected for the intron containing the pri-miRNA but not for another intron within the same transcript (Figure 8B). In contrast, no significant effect was observed for NDUFS8, which hosts pri-miR-7113 and pri-miR-4691 that are not bound by NF90 and whose abundance are not affected by NF90 (Figure 8B).

The expression of these genes was analysed by western blot of extracts obtained from HepG2 cells transfected with control (Scr and Scr2) and NF90-targeting (NF90 and NF902) siRNAs. All genes tested showed diminished expression upon loss of NF90, except NDUFS8 that showed no significant difference in expression (Figure 8C). Thus, NF90 modulates the expression of certain pri-miRNA host genes, including TIAM2, a known oncogene and metastasis factor in HCC (J.-S. Chen

et al., 2012; Yen et al., 2016).

Finally, to determine whether loss of gene expression correlated with increased pri-miRNA cropping following knock down of NF90, we performed modified RLM-5′ RACE as described previously (Barbier et al., 2018), using extracts of cells transfected with control (Scr and Scr2) and NF90-targeting (NF90 and NF902) siRNAs. Indeed, RLM RACE analysis showed enhanced cleavage of the intronic region of ZRANB2 hosting pri-miR-186 and GDF15 hosting pri-miR-3189 in extracts of NF90 knock down cells compared to controls (Figure 8D). This analysis indicates that loss of NF90 enhances transcript cleavage in the vicinity of the hosted pri-miRNA.

Discussion

We and others have previously shown that NF90 can inhibit the processing of certain miRNA precursors (Barbier et al., 2018; Higuchi et al., 2016; Sakamoto et al., 2009). However, it was unclear how widespread the impact of NF90 might be on human miRNA biogenesis. Here, we have used genome-wide approaches to address the effect of NF90 on the miRNA pool in HepG2 HCC cells. Our data indicate that NF90 modulates the processing of a specific subset of miRNA precursors. NF90 is associated with at least 38 human pri-miRNAs, as indicated by analysis of eCLIP data obtained by Nussbacher and Yeo (Nussbacher & Yeo, 2018). Of these, 22 showed increased abundance of mature miRNA products following knock-down of NF90. Thus, association of NF90 with a pri-miRNA is likely to influence its fate. Most NF90-associated pri-miRNAs did not overlap with those bound by either DGCR8 or Drosha. Moreover, results obtained by RNA-EMSA support the idea that NF90 and Microprocessor may compete for the binding of the subset of pri-miRNAs, at least in vitro. Further analysis will be required to determine whether the competition also occurs in vivo. Of note, RIP analysis showed that loss of NF90/NF45 complex led to increased binding of Drosha at pri-miRNAs that were highly bound by NF90 in control conditions. Conversely, loss of Microprocessor increased binding by NF90 to pri-miRNAs that were not

highly bound by NF90 in wild-type cells. Interestingly, for those pri-miRNAs that were bound by both NF90 and DGCR8, the binding profiles of the two factors were largely complementary. Furthermore, while the binding profile of DGCR8 was not noticeably different for this group compared to all pri-miRNAs bound by DGCR8, the binding profile of NF90 differed somewhat for this group compared to all pri-miRNAs bound by NF90. This could suggest that NF90 and DGCR8 might bind simultaneously to the pri-miRNA, and that the binding of DGCR8 may alter the binding mode of NF90 for such pri-miRNAs.

Since NF90 is a highly abundant and ubiquitously expressed protein, it might be expected that NF90-associated pri-miRNAs would be poorly processed in most cells. Indeed, the mature miRNA products of NF90 bound pri-miRNAs are very poorly expressed, or not expressed at all in control cells. They become readily detectable only upon loss of NF90. An exception is pri-miR-7-1, although interestingly, this miRNA shows tissue specific expression, being highly expressed only in brain and pancreas (Landgraf et al., 2007).

Our data suggests that pri-miRNAs upregulated after loss of NF90 share a common structure that might facilitate NF90 association with the stem region. This finding is consistent with a previous report showing structure-based recognition of adenovirus-expressed VA1 RNA by NF90 (Gwizdek et al., 2004). Extensive mutational analysis of VA1 association with NF90 showed no specificity for nucleotide sequence but rather the requirement for a minihelix structure within the stem region. The pri-miRNAs identified in this study also exhibit a minihelix-like structure that appears to be necessary for NF90 binding. Indeed, RNA EMSA showed that NF90 association with pri-miR-3173 and pri-miR-186 could be diminished by introducing destabilizing mutations, while NF90 association could be acquired by increasing the stability of the stem region, as for pri-miR-200a.

Interestingly, our data predict that the subset of NF90-associated pri-miRNAs may extend beyond those detected by eCLIP analysis. Using the characteristics determined from the eCLIP-positive, upregulated pri-miRNA group, that is duplex

length and free energy, we found that pri-miRNAs whose mature products were upregulated following loss of NF90 but were not positive by eCLIP analysis shared the same characteristics as the double positive group. The length of the duplex region and the free energy of the structure was comparable to that of double positive pri-miRNAs. RNA EMSA confirmed the predicted association with NF90 for two of these pri-miRNAs. Interestingly, both groups were significantly different to all pri-miRNAs or those that are unaffected by NF90 (other). Thus, it appears that the high specificity of eCLIP revealed a subset of pri-miRNAs that share a common structure. When this information was used to interrogate the group of pri-miRNAs who share the same biological response to loss of NF90, that is, upregulation of their mature products, we observed that both groups share the same characteristics. We predict that a certain number of the upregulated group likely do bind to NF90 but may escape detection by eCLIP. For example, as noted above, many of the pri-miRNAs are expressed at extremely low levels in control cells, which could make their association with NF90 difficult to detect. Interestingly, pri-miR-7-1 processing has been shown to be influenced by another RBP, HuR, which recruits MSI2 to the terminal loop. Binding of HuR/MSI2 was found to stabilize the stem region and led to diminished processing by Microprocessor (N. R. Choudhury et al., 2013). It would be interesting to determine whether binding of HuR/MSI2 to pri-miR-7-1 might facilitate NF90 binding to the stem region, and compete with Microprocessor. Similarly, it would be interesting to determine whether HuR/MSI2 can bind the terminal loop of other NF90-modulated pri-miRNAs in addition to pri-miR-7-1. NF90 may cooperate with other RBPs, such as HuR/MSI2 to control the processing of a subset of pri-miRNAs.

Another feature that the subset of NF90-modulated pri-miRNAs share is their restriction to human or primate lineages. Again, pri-miR-7-1 is an exception, being highly conserved throughout evolution. Thus, given that the subset of NF90-modulated pri-miRNAs are young and almost perfect hairpins, it is tempting to speculate that this group may have originated through recent insertion of repeat elements in the genome.

Interestingly, GO analysis of validated mRNA targets of the mature miRNAs showed significant enrichment for infection by viruses such as Epstein Barr Virus (EBV), hepatitis B virus (HBV) and human T lymphoma virus type 1 (HTLV1) and in viral carcinogenesis. Indeed, viral infection of cells induces translocation of NF90 from the nucleus to the cytoplasm (28). Thus, it is conceivable that pathological conditions such as viral infection could result in the coordinated processing of the NF90-modulated subset of pri-miRNAs, which target mRNAs important for viral replication.

Finally, transcriptomic analysis showed that association of NF90 with pri-miRNAs may diminish the expression of certain host genes, as described previously (Barbier et al., 2018). Among the pri-miRNA-hosting transcripts that are downregulated after loss of NF90, two are noteworthy. The expression of TIAM2, hosting pri-miR-1273C, is down-regulated upon loss of NF90. TIAM2 is a known oncogene and metastasis factor in HCC (J.-S. Chen et al., 2012; Yen et al., 2016). Levels of NF90 are elevated in HCC (Higuchi et al., 2016; Jiang et al., 2015) and it would be interesting to determine whether NF90-dependent modulation of TIAM2 might contribute to pathogenesis. Loss of NF90 also diminished expression of growth differentiation factor 15 (GDF15), hosting pri-miR-3189. GDF15 is expressed and secreted by a limited number of tissues, including liver. When complexed with its receptor, GFRAL, in brain and CNS, GDF15 supresses appetite (see (Tsai et al., 2018) for review). Cancer patients express high circulating levels of GDF15, which contributes to anorexia/cachexia. On the other hand, enhancement of GDF15 expression is a promising therapeutic strategy in the treatment of obesity. It would be interesting to determine whether high levels of NF90 in HCC may have a role in promoting expression of GDF15 from liver cells in cancer patients.

In summary, we have identified a subset of human pri-miRNAs that are bound by NF90. Analysis indicates that this subset shares a similar structure that appears to be favourable for NF90 binding. These data extend our knowledge of how processing of pri-miRNAs can be modulated by RBPs. This may be beneficial for understanding perturbations of miRNA levels in pathological conditions and could

also open up novel treatment strategies using nanotherapeutics.

Accession numbers

Small RNA-seq and RNA-seq data have been deposited at GEO (GSE132341).

Supplementary data

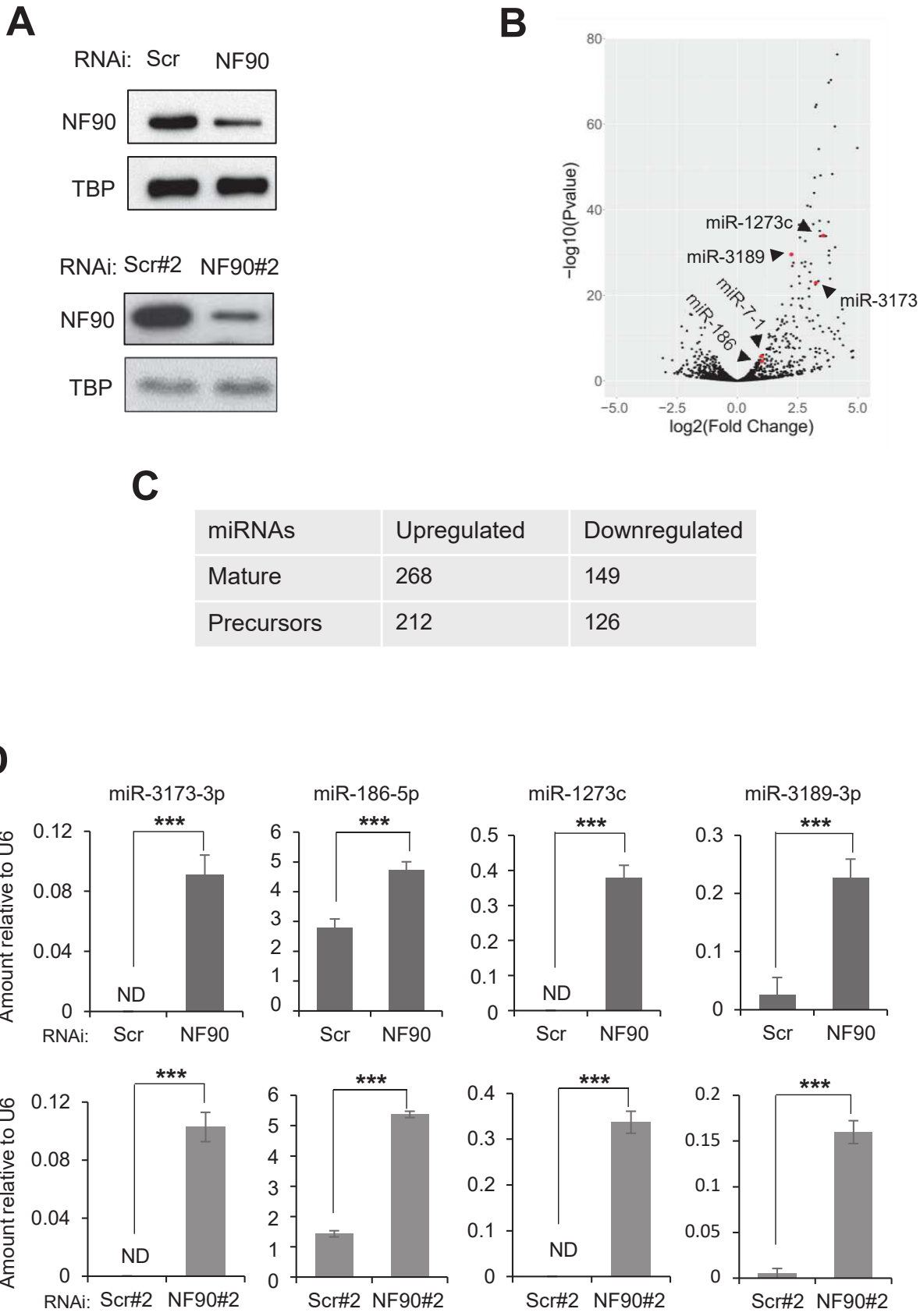
Supplementary Data are available at NAR online.

Acknowledgement

We wish to thank Catherine Dargemont and Xavier Contreras for critical reading of the manuscript, and the Gene Regulation lab and Hervé Seitz for helpful discussions.

Funding

This study was supported by funds from the European Research Council (RNAmedTGS to R.K.), Ministère de l'Enseignement Supérieur et de la Recherche et de l'Innovation scholarship (to G.G.), Japan Society for the Promotion of Science (Grant-in-aid for Young Scientists (B) 17K15601 and 19K16523 to T.H. and Grant-in-aid for Scientific Research (C) 16K08590 and 19K07370 to S.S.). Funding for open access charge: [ERC, RNA MedTGS]



miRNAs	Upregulated	Downregulated
Mature	268	149
Precursors	212	126

Figure 1. NF90 Modulates the Expression Level of a subset of miRNAs in HepG2 cells. **(A)** Extracts of HepG2 cells transfected with non-targeting control siRNAs (Scr, Scr#2) or siRNA targeting NF90 (NF90, NF90#2) as indicated were analyzed by immunoblot using the antibodies indicated. **(B)** Total RNA extracted from cells transfected with siScr or siNF90 were analyzed by small RNA-seq. Results are shown as $\log_2$ fold change versus $-\log_{10}$ p-value. **(C)** Table summarizing the number of mature miRNAs and pri-miRNAs modulated in HepG2 cell line upon loss of NF90, according to small-RNA seq. **(D)** Total RNA extracted from cells described in (A) were analyzed by Taqman RT-qPCR as indicated. Results were normalized by those obtained for U6 abundance in the same samples. ND indicates “not detected”. Data represent mean +/- SEM obtained from 3 independent experiments (***P < 0.001, independent Student’s t test).

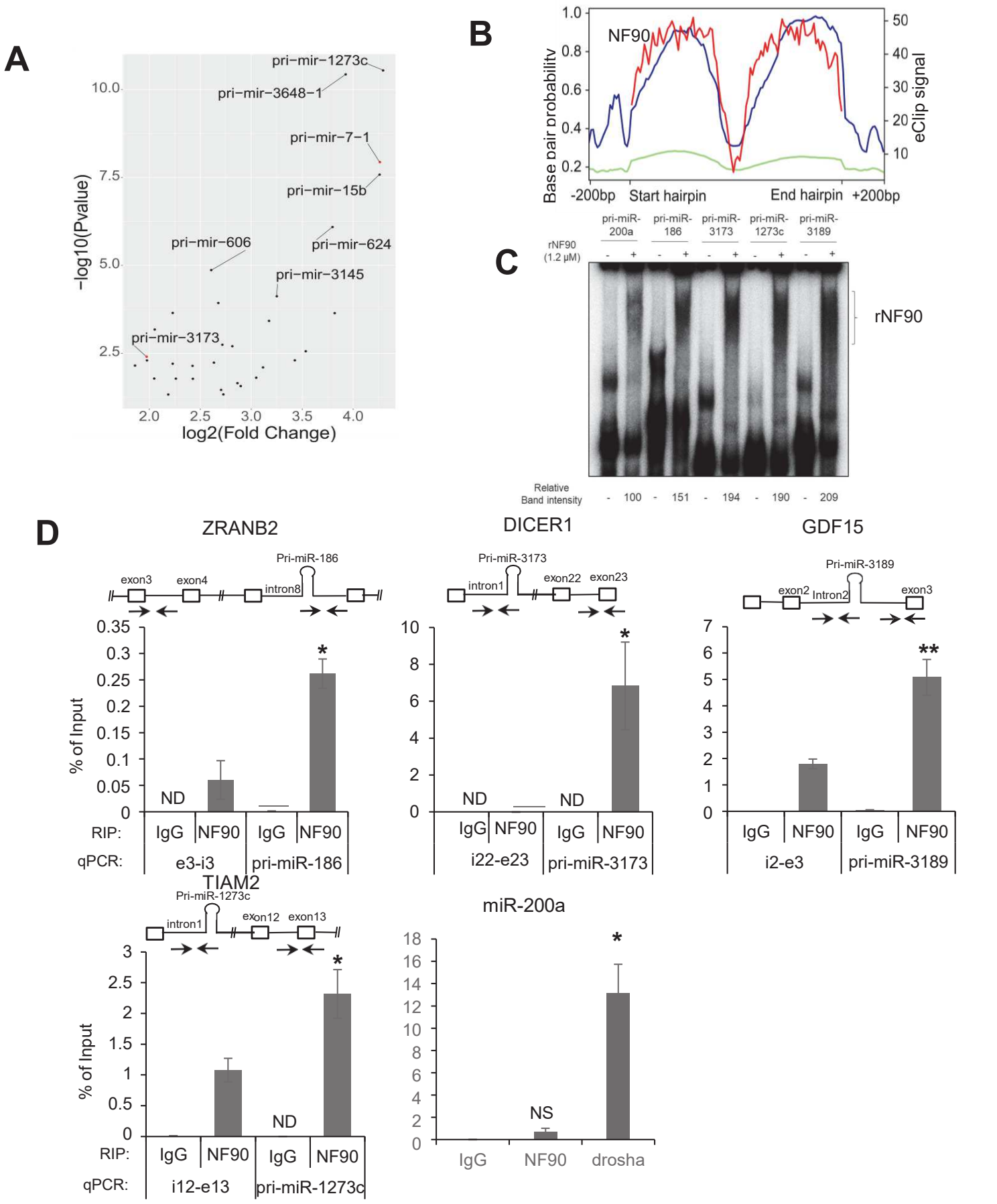


Figure 2. NF90 is associated with a subset of pri-miRNAs in HepG2 cells. **(A)** Dot plot representation of eCLIP data showing the 38 pri-miRNAs significantly associated with NF90 in HepG2 cells. Graph shows $\log_2$ fold change versus $-\log_{10}$ p-value. **(B)** Distribution of NF90 eCLIP reads along the region +/-200 bp of NF90-associated pri-miRNAs (blue) or all miRNAs (green) and base pair probability of NF90-associated hairpins (red). **(C)** RNA EMSA performed using recombinant NF90 was probed with radiolabelled pri-miRNAs as indicated. rNF90-pri-miRNA complexes are indicated on the figure. **(D)** RIP analysis of HepG2 cells transfected with NF90 targeting siRNA or a non-targeting control (Scr), as indicated using anti-NF90, anti-Drosha or a control IgG antibody. Immunoprecipitates were analyzed by RT-qPCR amplifying a region proximal or distal to the miRNAs. ND indicates 'Not Detected'. NS indicates 'Not Significant'. Data represent mean +/- SEM obtained from 3 independent experiments (*P < 0.05, **P < 0.01, independent Student's t test).

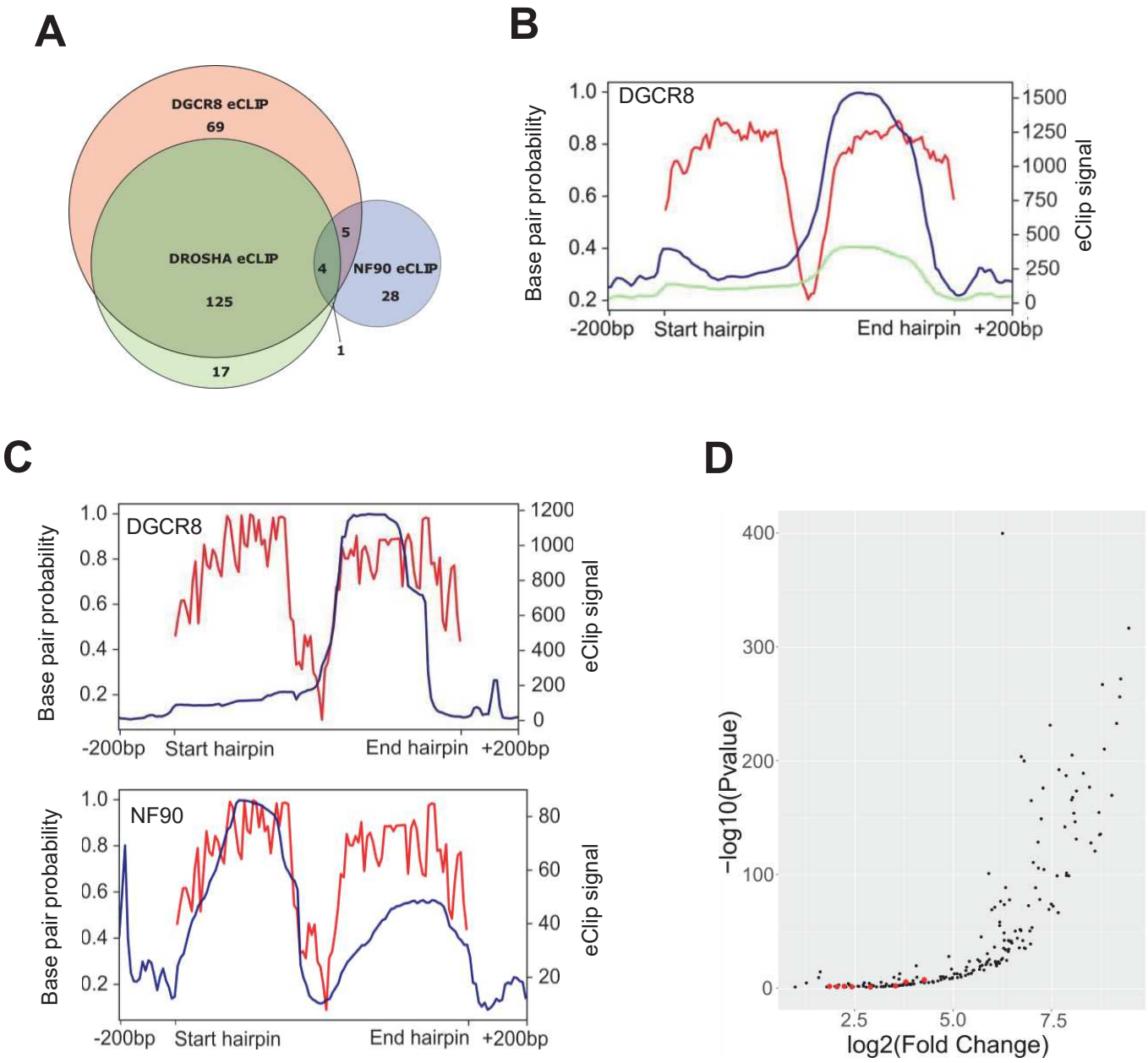


Figure 3. NF90-associated pri-miRNAs are poorly associated with Microprocessor. **(A)** Venn diagram showing the number of pri-miRNAs associated with DGCR8, Drosha or NF90 detected by eCLIP, as indicated. **(B)** Distribution of DGCR8 eCLIP reads along the region +/- 200 bp of DGCR8-associated pri-miRNAs (blue) or all miRNAs (green) and base pair probability of DGCR8-associated hairpins (red). **(C)** Distribution of eCLIP reads along the region +/-200 bp of pri-miRNAs associated with both DGCR8 and NF90 (blue) and base pair probability of the hairpins (red). Left panel shows DGCR8 eCLIP reads, right panel shows NF90 eCLIP reads in blue. **(D)** Dot plot representation of eCLIP data showing 203 pri-miRNAs significantly associated with DGCR8 in HepG2 cells. Graph shows $\log_2$ fold change versus $-\log_{10}$ p-value. Red dots indicate pri-miRNAs that are also significantly associated with NF90.

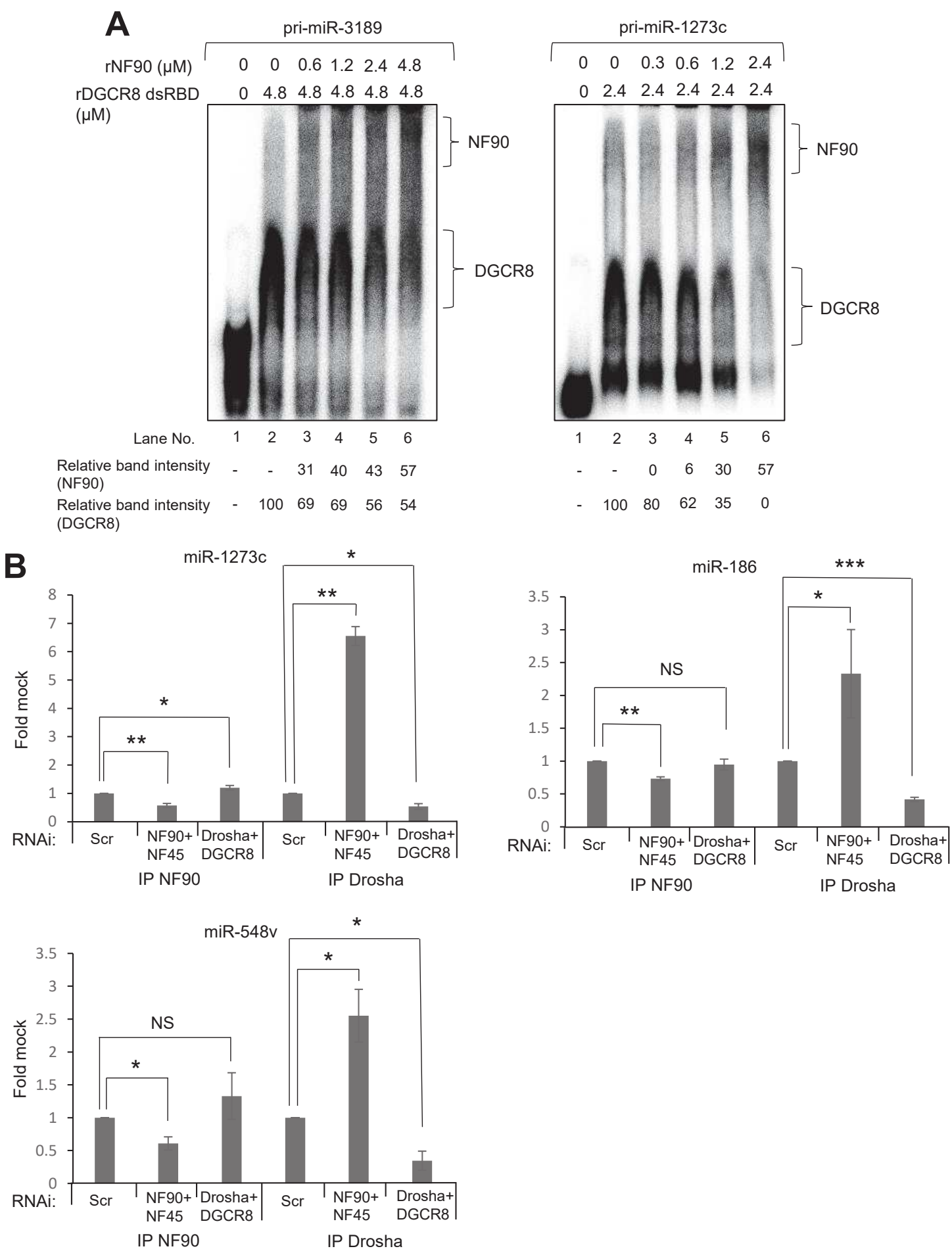


Figure 4. NF90 competes with the Microprocessor for the binding to pri-miRNAs. **(A)** RNA EMSA carried out using rDGCR8 dsRBD either alone or together with increasing amounts of rNF90 and probed with radio-labelled pri-miR-3189 or pri-miR-1273c. **(B)** Immunoprecipitates obtained using anti-NF90, anti-Drosha or a control antibody were analyzed by RT-qPCR amplifying a region proximal to the pri-miRNAs. The fold change relative to the control antibody sample was calculated and results are presented relative to the control sample (siScr), which was attributed a value of 1. Data represent mean +/- SEM obtained from 3 independent experiments (*P < 0.05, **P < 0.01, ***P < 0.001, NS 'Not Significant'. independent Student's t test).

A

Feature (nt)	All pri-miRNAs	Double positive pri-miRNAs
Length	81.88	82.5
Terminal loop size	7.92	7.87
Longest duplex	21.11	27.68

B

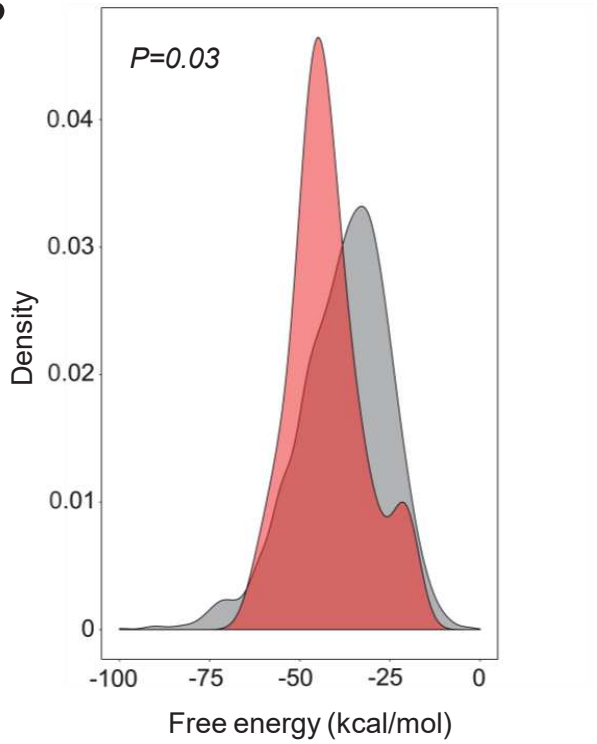


Figure 5. NF90 associates with a subset of highly stable pri-miRNAs. **(A)** Structural characteristics of all human pri-miRNAs and NF90 double positive pri-miRNAs. **(B)** Graph showing the free energy of all pri-miRNAs (grey) and NF90 double positive pri-miRNAs (red).

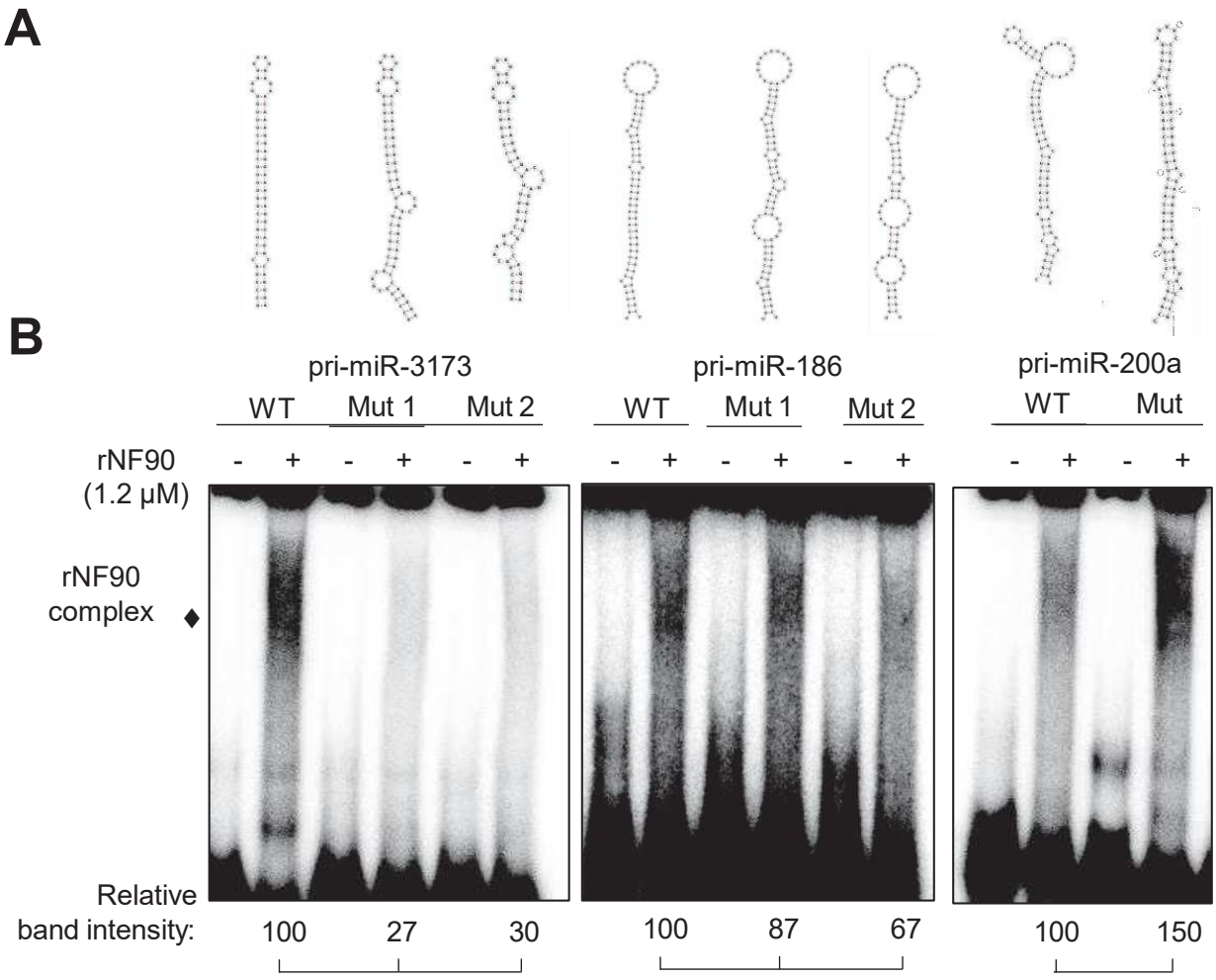


Figure 6. Modification of pri-miRNA structure alters NF90 binding. **(A)** Representations of wt or mutant pri-miRNAs sequences, as indicated. **(B)** RNA EMSA performed using recombinant NF90 and probed with radiolabelled pri-miRNAs as indicated. rNF90-pri-miRNA complexes are indicated on the figure. Relative band intensities (normalized to signal for wt) are shown below.

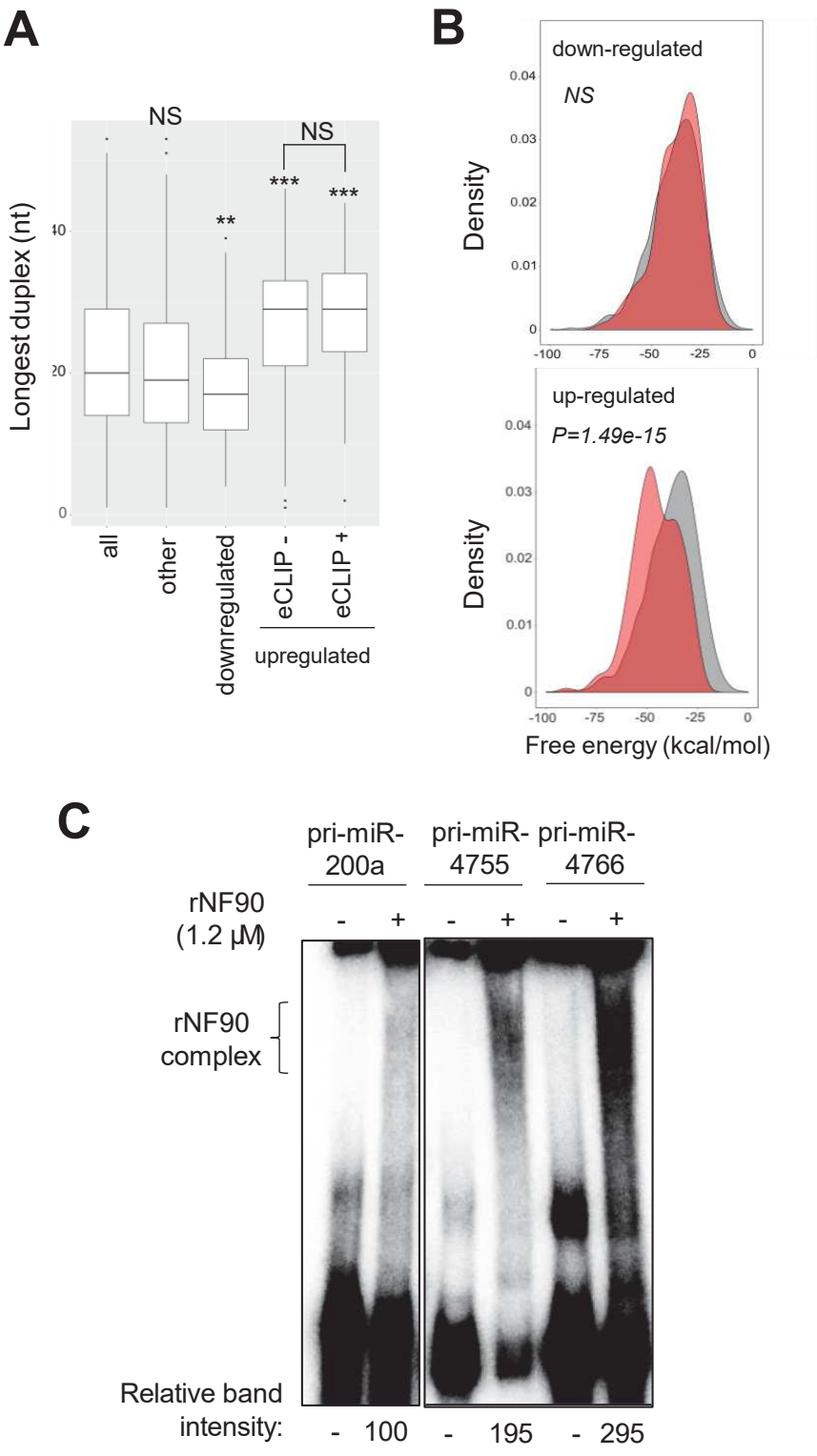


Figure 7. Pri-miRNAs whose mature products are upregulated following loss of NF90 share a similar structure. **(A)** Box plot representation of the longest duplex length of pri-miRNAs sorted into the indicated categories. (*P < 0.05, ***P < 0.001, NS, not significant, Wilcoxon test). **(B)** Graphical representation of the free energy of pri-miRNAs whose mature products are downregulated or upregulated as indicated following loss of NF90 (red) compared to all pri-miRNAs (grey). **(C)** RNA EMSA performed using recombinant NF90 and probed with radiolabelled pri-miRNAs as indicated. rNF90-pri-miRNA complexes are indicated on the figure. Relative band intensities (normalized to pri-miR200a) are shown below.

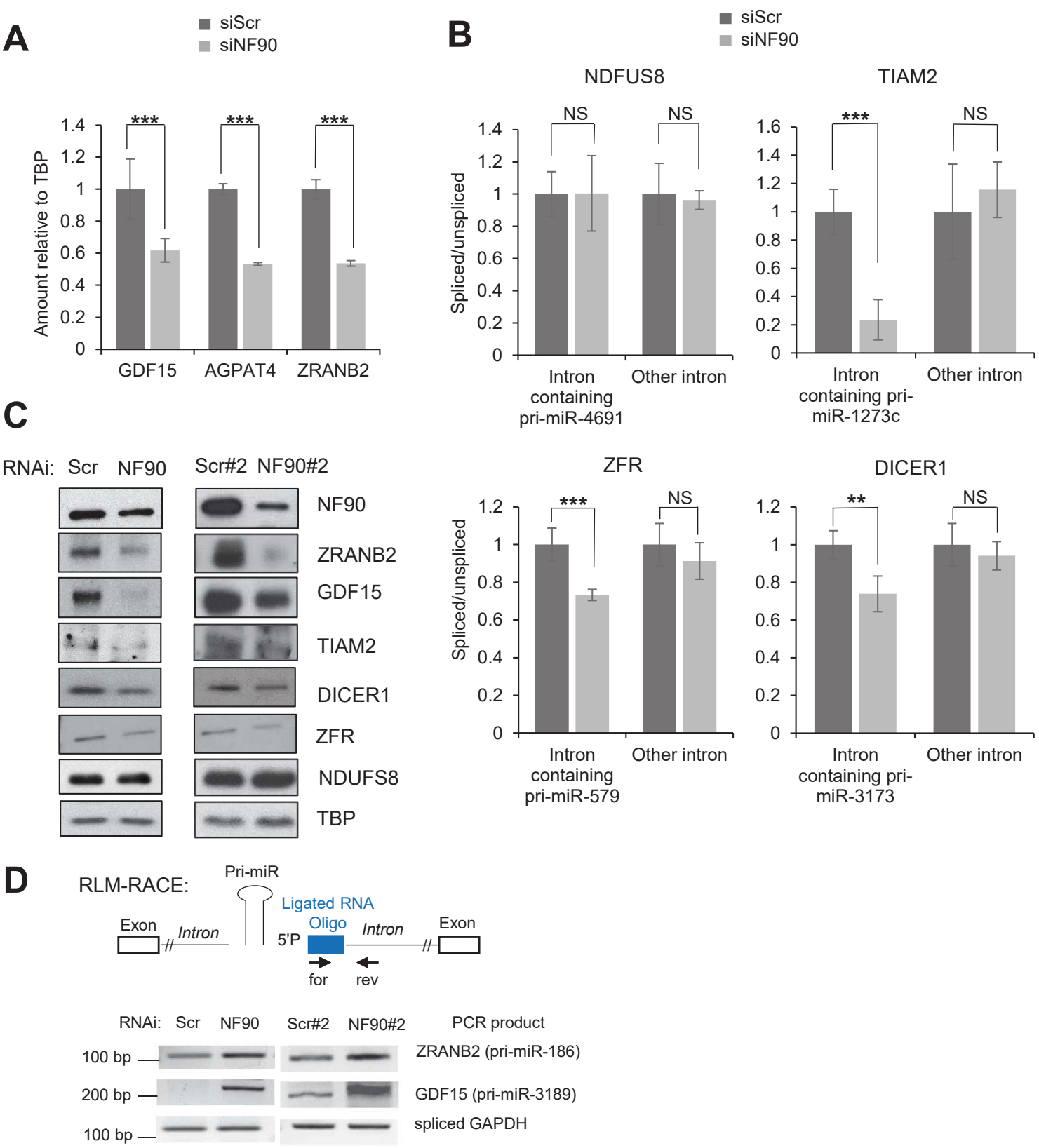
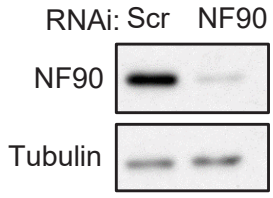


Figure 8. NF90 impacts expression of genes hosting pri-miRNAs. **(A)** Extracts of HepG2 cells transfected with siRNA targeting NF90 or a non-targeting control (Scr) as indicated were analyzed by RNA-seq and DESeq2. Data represent mean +/- SEM obtained from 3 independent samples (***P < 0.001, independent Student's t test). **(B)** The abundance of exon-intron junctions and exon-exon junctions in samples described in A was measured by RT-qPCR using PCR primers amplifying spliced or unspliced transcripts including introns containing pri-miRNAs or other introns. The splicing efficiency was calculated by the ratio of spliced to unspliced transcripts. Values obtained for the control sample (siScr) were attributed a value of 1. NS indicates 'Not Significant'. The graphs represent the mean +/- SEM obtained from three or more independent experiments (**P < 0.01, ***P < 0.001, independent Student's t test). **(C)** Extracts of HepG2 cells transfected with siRNA targeting NF90 (NF90, NF90#2) or a non-targeting control (Scr, Scr#2) as indicated were analyzed by immunoblot using the antibodies indicated. **(D)** NF90 modulates transcript cleavage at the region containing miRNA. Extracts of HepG2 cells transfected with siRNAs targeting NF90 (NF90, NF90#2) or non-targeting controls (Scr, Scr#2) as indicated were analyzed by modified 5' RLM-RACE. Forward and reverse primers used, and the predicted sizes of the PCR products are indicated.

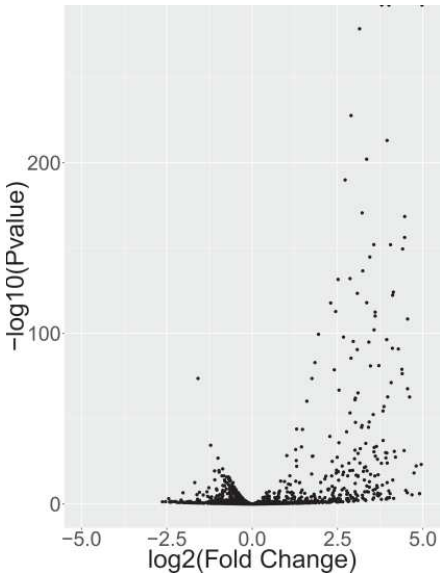
A



C

miRNAs	Upregulated	Downregulated
Mature	278	84
Precursors	217	77

B



D

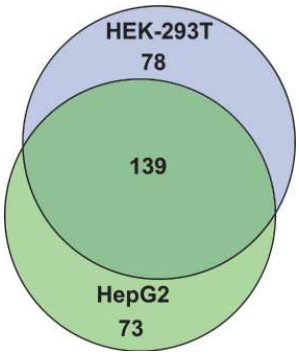


Figure S1. NF90 Modulates the Expression Level of miRNAs in HEK-293T cells. **(A)** Extracts of HEK-293T cells transfected with siRNA targeting NF90 or a non-targeting control (Scr) were analyzed by Western blot using the indicated antibodies. **(B)** Samples described in A were analyzed by small RNA-seq. Results are shown as log2 fold change versus –log10 p-value. **(C)** Table summarizing the number of mature miRNAs and pri-miRNAs modulated in HEK-293T cells upon loss of NF90, according to small-RNA seq. **(D)** Venn diagram representing the number of miRNAs upregulated following knock-down of NF90 in HepG2 versus HEK-293T cells.

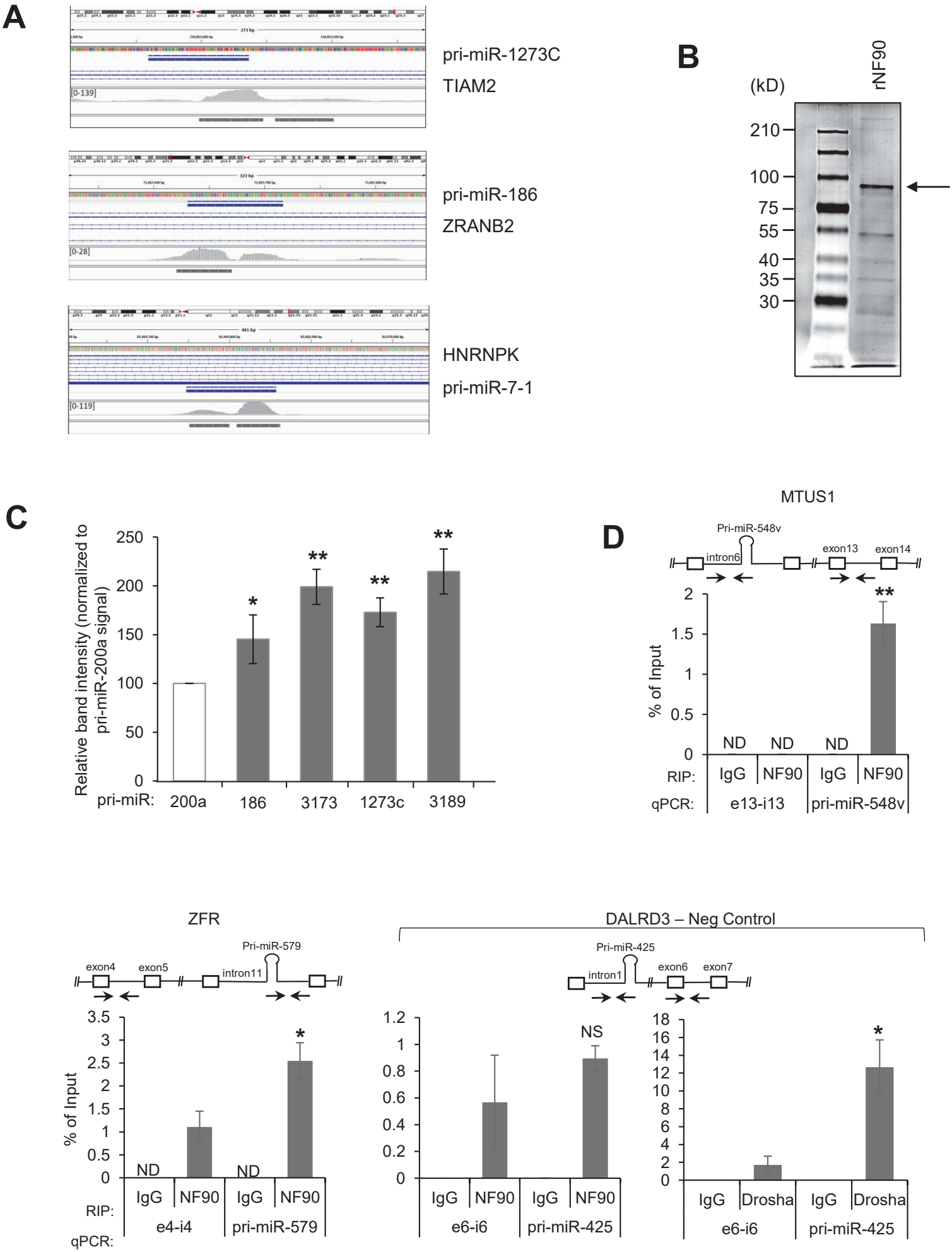

A
pri-miR-1273C
TIAM2
[0-139]
pri-miR-186
ZRANB2
[0-28]
HNRNPK
pri-miR-7-1
[0-119]
B
rNF90
(kD)
210
100
75
55
40
35
30
C
Relative band intensity (normalized to pri-miR-200a signal)
250
200
150
100
50
0
pri-miR:
200a
186
3173
1273c
3189
D
MTUS1
Pri-miR-548v
intron6
exon13
exon14
% of Input
2
1.5
1
0.5
0
ND
ND
ND
RIP:
IgG
NF90
IgG
NF90
qPCR:
e13-i13
pri-miR-548v
ZFR
Pri-miR-579
exon4
exon5
intron11
% of Input
3.5
3
2.5
2
1.5
1
0.5
0
ND
ND
RIP:
IgG
NF90
IgG
NF90
qPCR:
e4-i4
pri-miR-579
DALRD3 – Neg Control
Pri-miR-425
intron1
exon6
exon7
1.2
1
0.8
0.6
0.4
0.2
0
NS
IgG
NF90
IgG
NF90
e6-i6
pri-miR-425
18
16
14
12
10
8
6
4
2
0
IgG
Drosha
IgG
Drosha
e6-i6
pri-miR-425

Figure S2. (A) Browser shots of NF90 eCLIP read coverage over the pri-miRNAs indicated on the figure. Blue lines represent host gene showing localization of the pri-miRNA. eCLIP reads are shown in grey and locations of eCLIP peaks are shown as dark grey bars. The arrows indicate the strand from which the reads originated. **(B)** Recombinant NF90 used in RNA EMSA was analyzed by SDS-PAGE and Comassie brilliant blue staining. **(C)** RNA EMSA shown in Figure 2C performed using recombinant NF90 and probed with radiolabeled pri-miRNAs was carried out in three independent experiments. The graph shows the mean ± SD of the relative band intensity normalized to pri-miR-200a signal *(*P < 0.05 **P < 0.005, independent Student's t test).*

(D) HepG2 cells transfected with siRNA targeting NF90 or a non-targeting control (Scr) were subjected to RIP using anti-NF90, anti-Drosha or a control antibody. Immunoprecipitates were analyzed by RT-qPCR. ND indicates 'Not Detected'. Data represent mean ± SEM obtained from 3 independent experiments *(*P < 0.05, **P < 0.01, independent Student's t test)*.

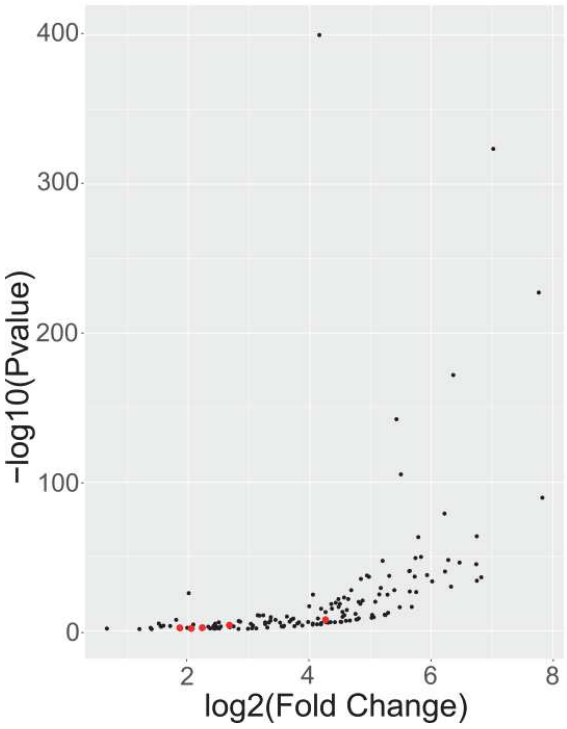


Figure S3. DotPlot of Drosha-associated pri-miRNAs, determined by eCLIP analysis. Red dots indicate the position of pri-miRNAs that are also positive for association with NF90.

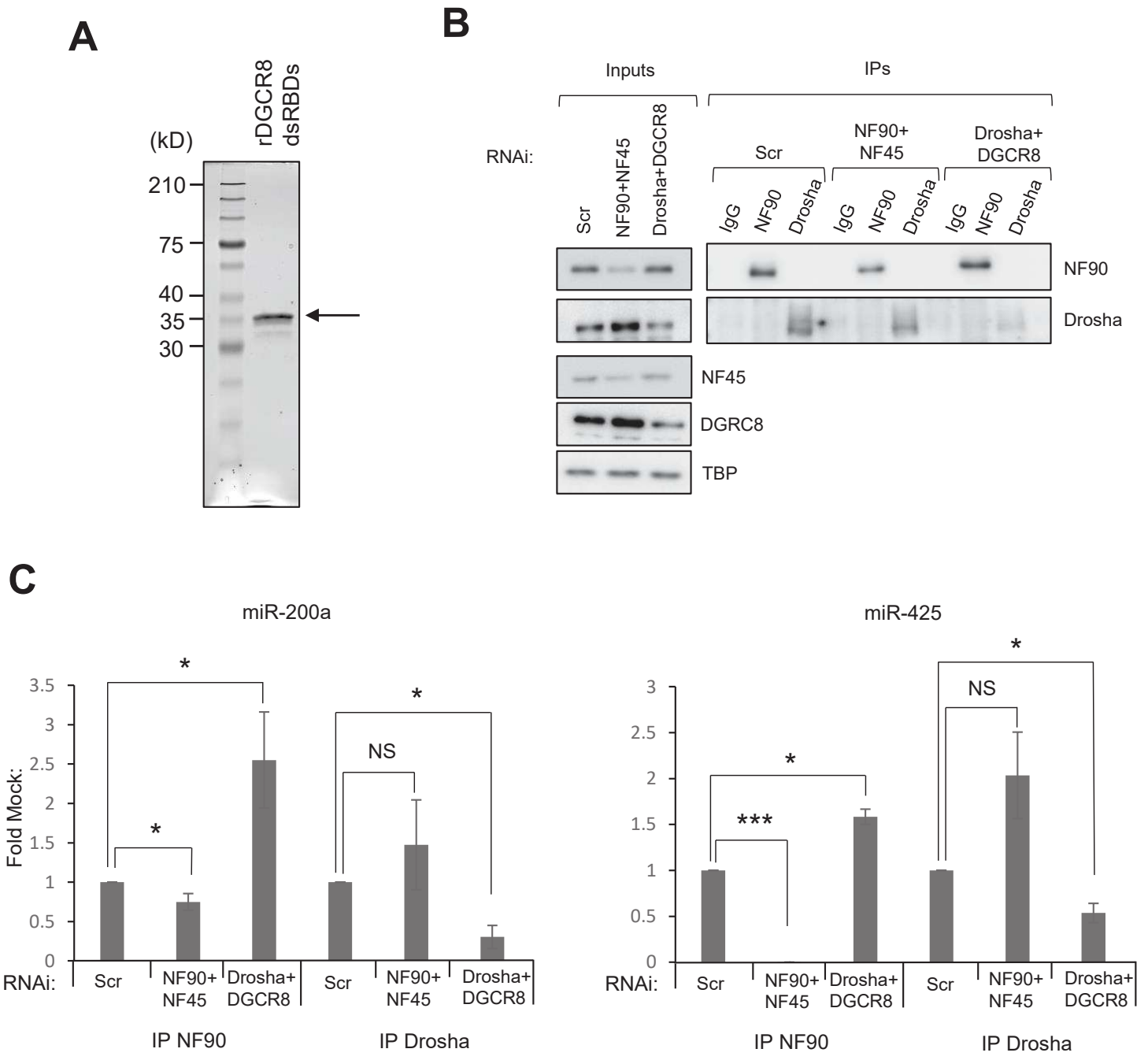


Figure S4. (A) Recombinant DGCR8 dsRBDs used in RNA EMSA was analyzed by SDS-PAGE and Comassie brilliant blue staining. **(B)** Extracts of HepG2 cells transfected with siRNAs targeting NF90 and NF45, Drosha and DGCR8 or a non-targeting control (Scr) were analyzed by Western blot using the indicated antibodies (left panel). The same extracts were used of RIP using antibodies anti-NF90, anti-Drosha or control IgG, as indicated (right panel). **(C)** HepG2 cells transfected with siRNA targeting NF90 and NF45 or Drosha and DGCR8 or a non-targeting control (Scr), as indicated, were subjected to RIP using anti-NF90, anti-Drosha or a control IgG antibody. Immunoprecipitates were analyzed by RT-qPCR. NS indicates 'Not Significant'. Data represent Fold mock (IgG) relative to the control sample (siScr), which was attributed a value of 1, obtained from 3 independent experiments *(*P < 0.05, ***P < 0.001, independent Student's t test)*.

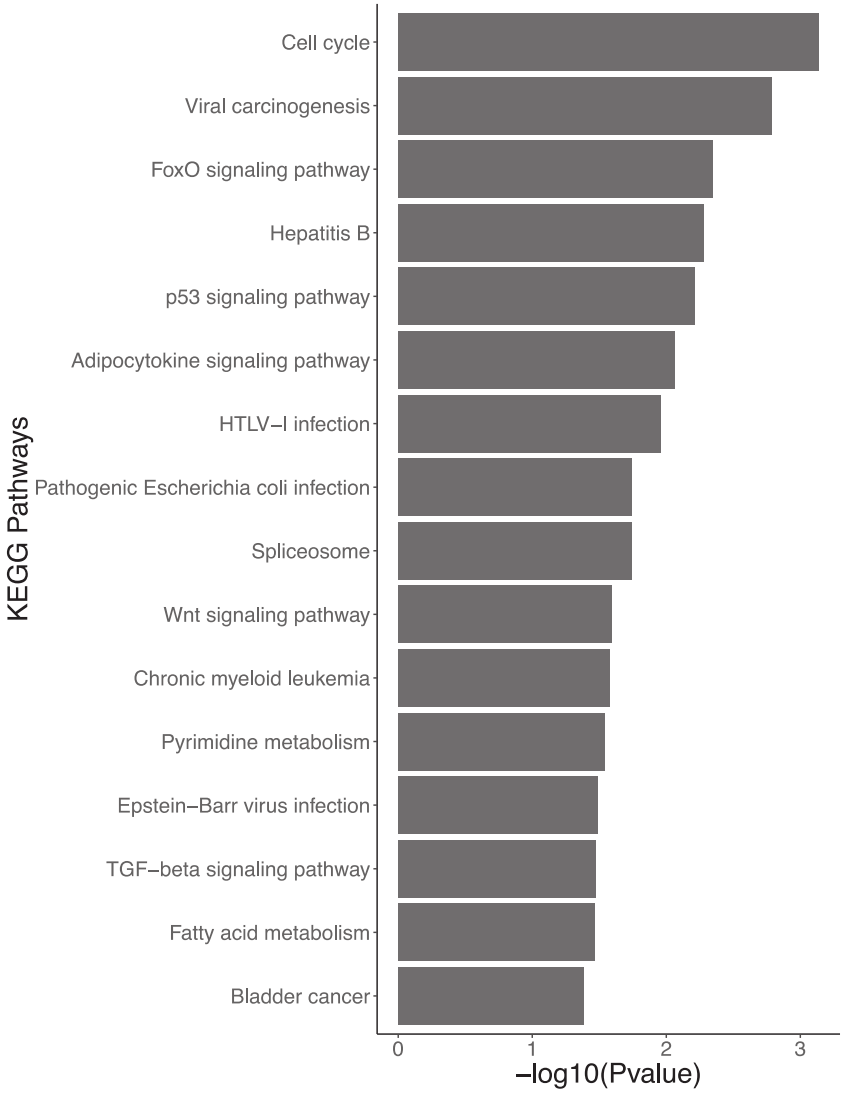


Figure S5. NF90 double positive pri-miRNAs target genes involved in viral infection and cancer. Gene ontology of validated targets of NF90-bound and upregulated miRNAs.

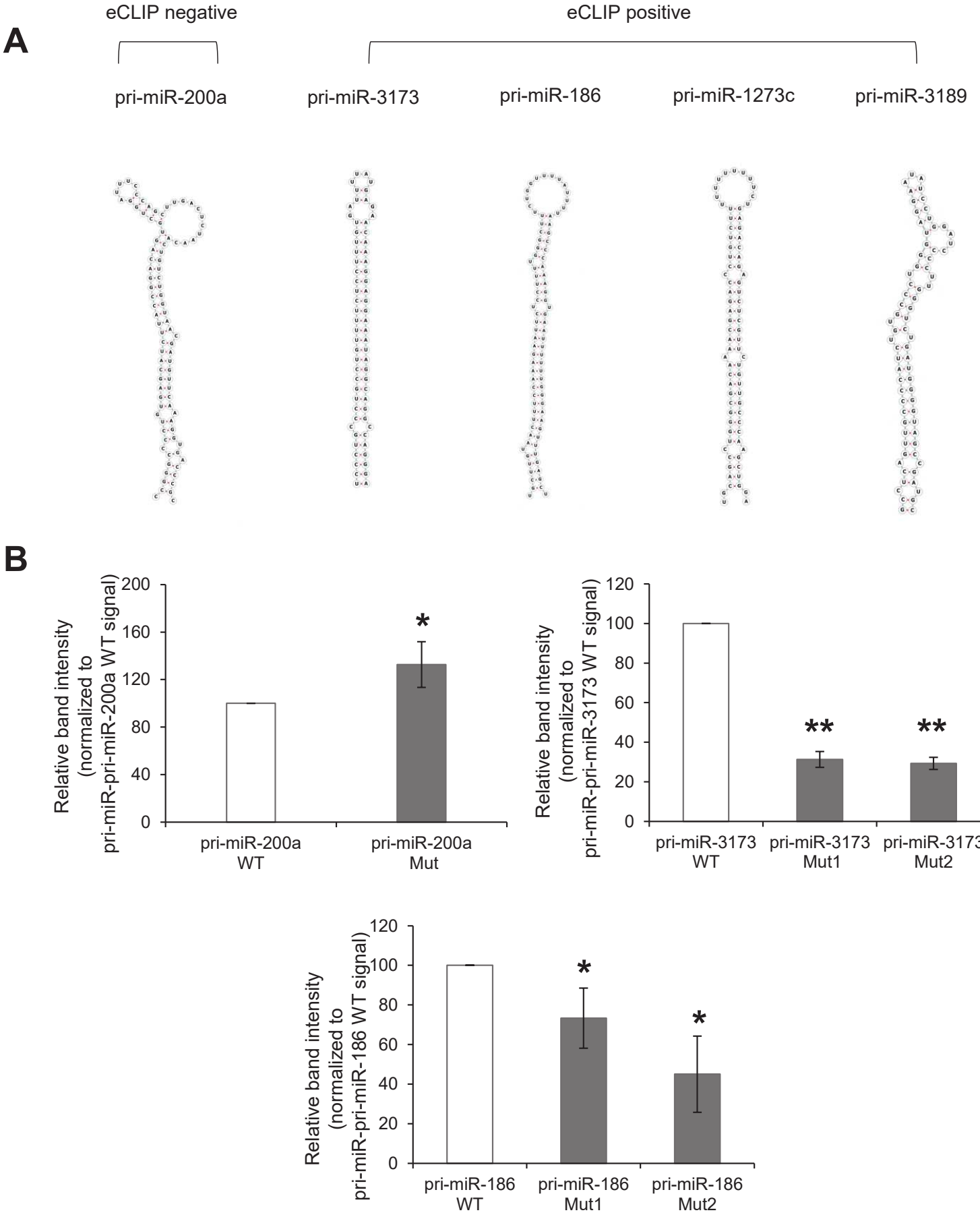


Figure S6. (A) NF90-associated pri-miRNAs are highly stable. Predicted folding of pri-miRNAs that are significantly associated or not with NF90 as indicated. RNA structures were predicted using FORNA. **(B)** RNA EMSA shown in Figure 5B performed using recombinant NF90 and probed with radiolabeled WT or mutant pri-miRNAs was carried out in three independent experiments. The graph shows the mean ± SD of the relative band intensity normalized to the WT signal *(*P < 0.05 **P < 0.005, independent Student's t test).*

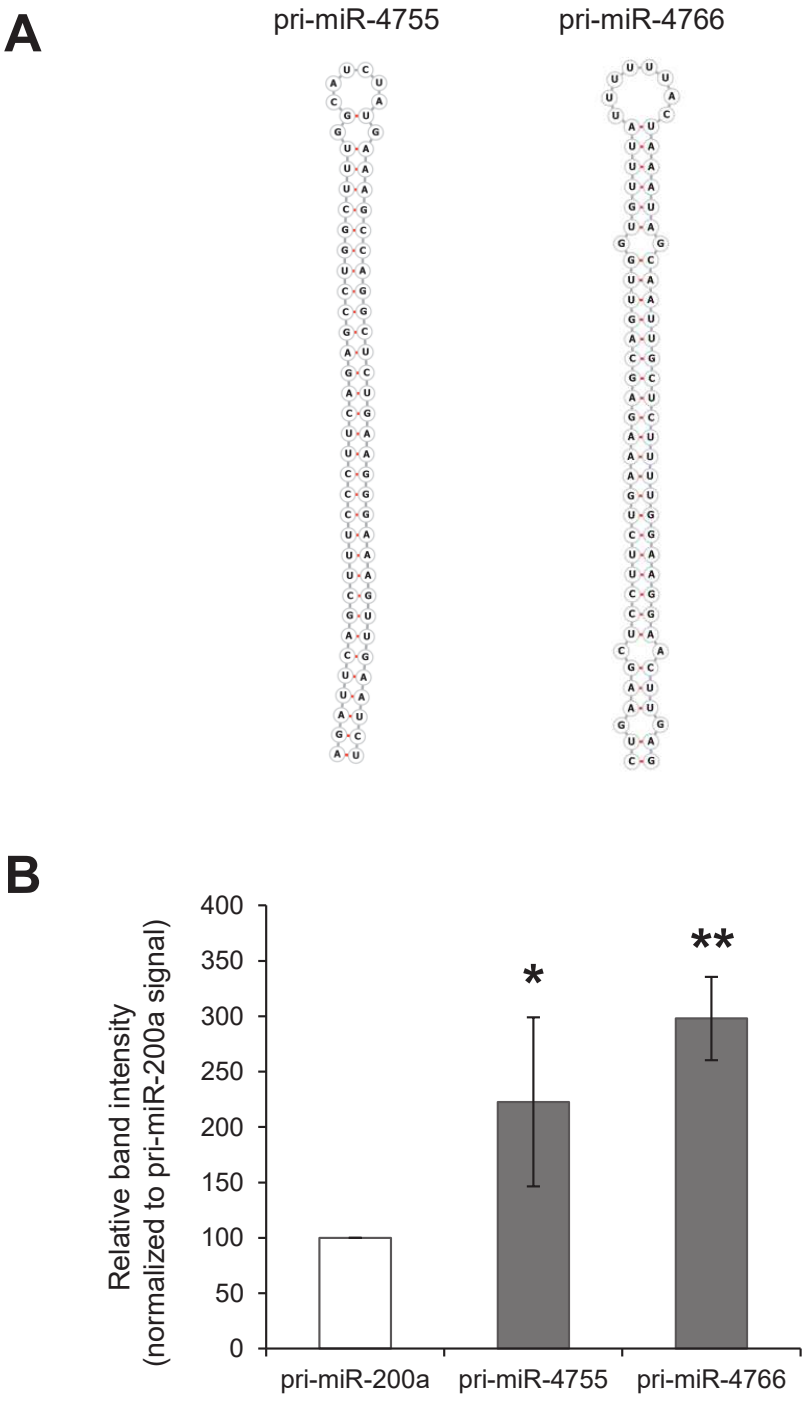


Figure S7. (A) NF90-modulated pri-miRNAs are highly stable. Predicted folding of examples of pri-miRNAs whose mature products are upregulated following loss of NF90. RNA structures were predicted using FORNA. **(B)** RNA EMSA shown in Figure 6C performed using recombinant NF90 and probed with radiolabeled pri-miR-200a, pri-miR-4755 and pri-miR-4766 was carried out in three independent experiments. The graph shows the mean ± SD of the relative band intensity normalized to pri-miR-200a *(*P < 0.05, **P < 0.05, independent Student's t test).*

Supplementary Table S1. Double stranded siRNAs used in this study.

siRNA	Sequence (5' to 3')
Scr	gcgcgcuuuguaggauucg(dTdT)
Scr#2	ucugcaagguuaggcgucu(dTdT)
NF90	ccaaggaacucuaucacaa(dTdT)
NF90#2	gaguugaaguauugauaac(dTdT)
Drosha	cgaguaggcuucgugacuu(dTdT)
DGCR8	caucggacaagagugugau(dTdT)
NF45	guggugauacucaagauucugccaa(dTdT)

Supplementary Table S2. Primary antibodies used in this study.

Antibody	Reference	Supplier
NF90	A303-651A	Bethyl Laboratories
GDF-15	sc-377195	SCBT
ZRANB2	sc-514200	SCBT
TIAM2	sc-514090	SCBT
NDUFS8	sc-515527	SCBT
TUBULIN	DM1A clone, T6199	Sigma-Aldrich
TBP	sc-421	SCBT
Drosha (IP)	Ab-12286	Abcam
Drosha (WB)	sc-33778	SCBT
DGCR8	Ab-82876	Abcam
Rabbit IgG	p120-101	Bethyl Laboratories
DICER1	sc-136981	SCBT
ZFR	A14281	ABClonal
NF45	A303-147A	Bethyl Laboratories

Supplementary Table S3. Primers used in this study.

Primer	Forward (5' to 3')	Reverse (5' to 3')
Spiced GAPDH	cac atc gct cag aca cca t	gag gtc aat gaa ggg gtc at
U6	ctc gct tcg gca gca cat ata c	gga acg ctt cac gaa ttt gcg tg
pri-miR-1273C (EMSA)	ctt ggg aag ctg agg tag gc	act tgg tac tga ggc gga gg
pri-miR-186 (EMSA)	aca gaa cac cca tca tat tc	gtt gac att cac atg ctt c
pri-miR-200a (EMSA)	ctg gct gct cac cgc tcc	gat gtg cct cgg tgg tgt cc
pri-miR-3173 (EMSA)	cat tgg agg tct agg gct ta	gtt ctt cct cgg cac aag
pri-miR-3189 (EMSA)	agc agc ccc cat atc taa tc	ctg gca tcc ctg tac ctc
pri-miR-4755 (EMSA)	aga gat gag gaa ggt tat ggc t	tgg ccc aaa cct cat aga c
pri-miR-4766 (EMSA)	ccc ttc tac ctt tct gaa gct c	cac aca ggt ggc act caa c
5'RLM-RACE pri-miR-186 (outer)	gct gat ggc gat gaa tga aca ctg (adapter)	aaa cca ggt ata tgg cac agc aac
5'RLM-RACE pri-miR-186 (inner)	cgc gga tcc gaa cac tgc gtt tgc tgg ctt tga tg (adapter)	tgt tga cat tca cat gct tca ggt
5'RLM-RACE pri-miR-3189 (outer)	gct gat ggc gat gaa tga aca ctg (adapter)	acc aca ccc cca ttg ttt ctct
5'RLM-RACE pri-miR-3189 (inner)	cgc gga tcc gaa cac tgc gtt tgc tgg ctt tga tg (adapter)	acc aca ccc cca ttg ttt ctct
ZRANB2 e3-i3 (RIP)	gag ccg agg cct att tag tg	aag gtt acc ctg gct tgt ca
ZRANB2 pri-miR-186 (RIP)	cct gaa gca tgt gaa tgt caa	cca ggt ata tgg cac agc aa
DICER i22-e23 (RIP)	ggc cat gat ttt aaa gtt gc	tcc tcc tcc tcg taa tcc tc
DICER pri-miR-3173 (RIP)	aac aga acc tgg aca ctg ag	aga cac caa cct cac tca ag
TIAM2 i12-e13 (RIP)	ggt ttg agt ttg cag cct tc	aga aaa cag ggc ctc cat ct
TIAM2 pri-miR-1273c (RIP)	ctg aaa tgc tgt ccc cat ct	tgc cca gtc tct tct cgt tt
GDF-15 i2-e3 (RIP)	ctc cca aag tgc tgg gat ta	aga gat acg cag gtg cag gt
GDF-15 pri-miR-3189 (RIP)	acg cta cga gga cct gct aa	tta gat atg ggg gct gct tg
MTUS1 e13-i13 (RIP)	aga aag cct gaa agc tgt gtt	aat gca ggg ctc aat ttc ac
MTUS1 pri-miR-548v (RIP)	tct cag cgt ggc tac tag gaa	tgt acg gct aca gca tct gg
ZFR e4-i4 (RIP)	tca gcc ttc tgt tgc tga aa	ggg ctg aaa gtc cag aaa tg
ZFR pri-miR579 (RIP)	ttt tgt gtc tgg cat cgt tc	gga aac aag ttg cat gtc ca
DALRD3 e6-i6 (RIP)	atc tgt ggc cct gtg aaa gt	gta tgc cgg aac ctg tgt tt
DALRD3 pri-miR-425 (RIP)	agg gct gca atg gta gtg ac	aag gtg cat gac ctg gag ac
NDUFS8 miRNA unspliced	ctt agc cgg agt cca gga g	aag ccg cag tag atg cac tt
NDUFS8 miRNA spliced	cca cca tca act acc cgt tc	aag ccg cag tag atg cac tt
NDUFS8 unspliced	caa tgg cag cgt cct aca gt	caa cgg gga cac cac act
NDUFS8 spliced	atg gca gcg tcc tac agt g	agc agc ata ggc gtg gtg
TIAM2 miRNA unspliced	aac caa ggt ttt gcg tga ag	acc cag gta gct gaa gac ga
TIAM2 miRNA spliced	gct cag cca cca cct ata cc	acc cag gta gct gaa gac ga
TIAM2 unspliced	gga tgc ttt gga tag ccg ta	cga atg tgt gga ttc act tc
TIAM2 spliced	atc agt gac tgg acg gga ag	tcg cat gtg tgg att cac tt
ZFR miRNA unspliced	gat gca agt tct tgg gct gt	cct gga gga cca tga gga ta
ZFR miRNA spliced	gga gta ctg gcg aag acg ag	cct gga gga cca tga gga ta
ZFR unspliced	tcc ttc cac ttg tct cat agc a	ttt cag caa cag aag gct ga
ZFR spliced	tta tgg agg cta ccc cac tg	cag cag ttg ctg ttg gtt gt
DICER1 miRNA unspliced	gga aga gtt tga atg gct ca	ggg ctt ttc att cat cca gtg
DICER1 miRNA spliced	gtc cga tgg ttc tcg aag	gtt cta gca cag ctt act g
DICER1 unspliced	aat tgc att ctc act act gca	gtc cac aat cca cca caa tc
DICER1 spliced	att gtc cat cat gtc ctc gc	gtc cac aat cca cca caa tc

Supplementary Table S4. Wild-type and mutant pri-miRNAs sequences used for RNA EMSA. The pri-miRNA sequence is shown in red and flanking sequence is shown in black.

	pri-miR-200a	pri-miR-3173	pri-miR-186
Wild-type	ctggctgctcaccgctccggttcttccctgggct tccacagcagcccctgcctgcctggcgggac cccacgtccctcccgggcccctgtgagcatct taccggacagtgctggatttcccagcttgactc taacactgtctggtaacgatgttcaaaggtga cccgccgctcgccggggacaccaccgagg cacatc	cattggaggtctagggcttattttccagat agaattgagtctttgttggtcttgggccag cttccctgccctgcctgttttctcctttgtgatt ttatgagaacaaaggaggaaataggca ggccagggaaacgatctctctccctctctt gtgccgaggaagaact	acagaacacccatcatattcttcccaaacatttttttcat tgcttgtaactttccaaagaattctccttttgggctttctg gttttattttaagcccaaaggtgaatttttttgggaagttt gagctaaattccttcaaccaaaatatacaagtgaag aaaaaaaatttgtatttaaacatttgcacatttacttct acctgaagcatgtgaatgtcaac
Mutant #1	ctggctgctcaccgctccggttcttccctgggct tccacagcagcccctgcctgcctggcgggac cccacgtccctcccgggcccctgtgagcatct taccggacagtgctggatttcccagcctgtctg gtaacgatgttcaaaggtgacccgccgctcg ccggggacaccaccgaggcacatc	cattggaggtctagggcttattttccagat agaattgagtctttgttggtcttgggccag cttccctgcgacgcctgcctgttttctccttt gtgattttatgagaacaaaggaggaaag cgctaggcaggccagggaaacgatctc tctccctctcttgtgccgaggaagaact	acagaacacccatcatattcttcccaaacatttttttcat tgcttgtaactttccaactaaagaattctccttttgggct ttctggttttattttaagcccacctctaaggtgaatttttttg ggaagtttgagctaaattccttcaaccaaaatataca agtgaagaaaaaaaaatttgtatttaaacatttgcaca tttacttctacctgaagcatgtgaatgtcaac
Mutant #2		cattggaggtctagggcttattttccagat agaattgagtctttgttggtcttgggccag cttccctgcaagtcctgttttctcctttgtgatt ttatgagaacaaaggagaccctaggca ggccagggaaacgatctctctccctctctt gtgccgaggaagaact	acagaacacccatcatattcttcccaaacatttttttcat tgcttgtccagttccaaagaattctccttttgggctttct ggttttattttaagcccaaaggtgaaccacgtgggaa gtttgagctaaattccttcaaccaaaatatacaagtg aagaaaaaaaatttgtatttaaacatttgcacatttac ttctacctgaagcatgtgaatgtcaac

Supplementary Table S5. NF90-associated pri-miRNAs, as determined by eCLIP analysis.

hsa-mir-1273c	hsa-mir-4485	hsa-mir-548d-1
hsa-mir-1290	hsa-mir-4635	hsa-mir-548u
hsa-mir-15b	hsa-mir-4659a	hsa-mir-548v
hsa-mir-186	hsa-mir-4687	hsa-mir-5581
hsa-mir-1914	hsa-mir-4712	hsa-mir-570
hsa-mir-3140	hsa-mir-4714	hsa-mir-578
hsa-mir-3145	hsa-mir-4730	hsa-mir-579
hsa-mir-3173	hsa-mir-4762	hsa-mir-606
hsa-mir-3189	hsa-mir-4775	hsa-mir-624
hsa-mir-3646	hsa-mir-4779	hsa-mir-6751
hsa-mir-3648-1	hsa-mir-4782	hsa-mir-6839
hsa-mir-3680-1	hsa-mir-548aq	hsa-mir-7-1
hsa-mir-3939	hsa-mir-548ar	

Supplementary Table S6. 'Double positive' miRNAs whose abundance increased following loss of NF90 and that were positive for NF90 association by eCLIP, and their host gene.

miRNA	Small RNA-seq		Host Gene
	Fold Change (log2)	p value (-Log10)	
hsa-mir-1273c	3.55	33.94	TIAM2
hsa-mir-186	1.01	4.55	ZRANB2
hsa-mir-3140	3.77	37.16	FBXW7
hsa-mir-3145	2.56	9.78	NHSL1
hsa-mir-3173	3.03	25.36	DICER1
hsa-mir-3189	2.16	19.36	GDF15
hsa-mir-3646	3.76	4.88	HNF4A
hsa-mir-3939	1.27	7.59	RP1-167A14.2
hsa-mir-4659a	2.21	10.93	AGPAT5
hsa-mir-4714	3.36	23.31	IGF1R
hsa-mir-4762	2.14	5.66	ATXN10
hsa-mir-4775	1.37	3.72	CCNYL1
hsa-mir-4779	4.03	10.9	IMMT
hsa-mir-4782	3.94	3.87	SLC35F5
hsa-mir-548ar	3.24	6.79	CDC16
hsa-mir-548u	2.24	2.38	PRIM2
hsa-mir-548v	2.01	7.81	MTUS1
hsa-mir-5581	2.69	19.23	MEAF6
hsa-mir-578	2.17	3.18	CPE
hsa-mir-579	4	73.29	ZFR
hsa-mir-624	2.68	25.49	STRN3
hsa-mir-7-1	0.99	5.76	HNRNPK

Supplementary Table S7. MiRNAs downregulated in abundance following loss of NF90 and that are associated with NF90 by eCLIP.

miRNA	**Fold Change (log2)**	**p value (-Log10)**	**Host Gene**
hsa-mir-1914	-0.86	2.88	UCKL1
hsa-mir-6751	-1.52	3.73	SYVN1

Chapter 3

NF90 Interacts with Components of RISC and Modulates Association of Ago2 with mRNA

Giuseppa Grasso[1], Celine Franckhauser[1], Rima Nait-Saidi[1], Maxime Bello[1], Jérôme Barbier[1] and Rosemary Kiernan[1]*

[1] UMR9002 CNRS-UM, Institut de Génétique Humaine-Université de Montpellier, Gene Regulation lab, Montpellier, 34396, France

* To whom correspondence should be addressed. Tel: 33 4 34359939; Fax: 33 4 34359901; Email: Rosemary.Kiernan@igh.cnrs.fr

Abstract

Nuclear Factor 90 (NF90) is a double-stranded RNA-binding protein involved in a multitude of different cellular mechanisms such as transcription, translation, viral infection and mRNA stability. Recent data suggest that NF90 might influence the abundance of target mRNAs in the cytoplasm through miRNA- and Argonaute 2 (Ago2)-dependent activity. Identification of the NF90 interactome in the cytoplasm revealed several components of the RNA-induced silencing complex (RISC) and associated factors. Co-immunoprecipitation analysis confirmed interaction of NF90 with the RISC-associated RNA helicase, Moloney leukemia virus 10 (MOV10), and other proteins involved in RISC-mediated silencing, including Ago2. Furthermore,

NF90 association with MOV10 and Ago2 was found to be RNA-dependent. Glycerol gradient sedimentation indicates that these proteins occur in the same complex. Using target RNAs predicted to bind both NF90 and MOV10 in their 3' UTRs, NF90 association to the selected target mRNAs was increased upon loss of MOV10 and vice versa, suggesting the two proteins may compete for the binding of common target mRNAs. Interestingly, loss of NF90 led to an increase in association of Ago2 as well as the abundance of the target mRNA.These findings suggest a role for NF90 in the regulation of RISC-mediated silencing which stabilizes target mRNAs.

Introduction

Nuclear Factor 90 (NF90) is a double-stranded RNA-binding protein (RBP) that is involved in a plethora of different cellular processes and pathways, such as transcription, splicing, translation and mRNA stability or degradation (Castella et al., 2015). More recently, NF90 was also linked to microRNA (miRNA) biogenesis and circular RNAs (circRNAs) stability (X. Li et al., 2017; Sakamoto et al., 2009). NF90 is an ubiquitous and generally abundant protein that has been shown to shuttle from the nucleus to the cytoplasm depending on its phosphorylation status and as a result of several stimuli, such as viral infection or hypoxia (Harashima et al., 2010; W. Zhang et al., 2018). During viral infection, cytoplasmic NF90 can bind viral RNAs supporting or inhibiting viral replication, depending on the type of virus (Patiño et al., 2015). Besides viral RNAs, NF90 can also associate with cellular mRNAs, increasing their stability or influencing their translation (D. Song et al., 2017; Vumbaca et al., 2008).

NF90 contains two tandem double-stranded RNA-binding motifs (dsRBMs) that were shown to contribute to the binding of the same RNA molecule simultaneously (Jayachandran et al., 2016). However, the precise RNA binding mode of NF90 is still under debate. It was shown that NF90 is able to recognize specific RNAs structures, such as minihelix, and that this structure is sufficient for NF90 binding (Grasso et al., 2020; Gwizdek et al., 2004). However, a sequence motif has also been

shown to positively influence its binding to short mRNAs (Jayachandran et al., 2016). Nevertheless, NF90 RNA binding activity is strongly influenced by the heterodimerization with its protein partner Nuclear Factor 45 (NF45) which leads to thermodynamic stabilization of the complex and enhanced affinity for RNA substrates (Schmidt et al., 2017).

Recent findings implicate NF90 in mRNA stability and translation through miRNAs (Idda et al., 2018), which could suggest an involvement in RISC-mediated gene silencing. RISC-mediated gene silencing is a well-known posttranscriptional gene regulation mechanism that, in human, was shown to promote translational inhibition and degradation of target mRNAs mainly by recruiting Argonaute 2 (AGO2) protein (Flores et al., 2014). Ago2, guided by the sequence complementarity of a miRNA, is able to diffuse along the 3′ UTR of target mRNAs recognizing the miRNAs recognition elements (MRE) and recruiting effector proteins, such as deadenylases and 5′-to-3′ exonucleases, in order to inhibit translation and/or trigger mRNA degradation (Bartel, 2009). Interestingly, characterization of the Ago2 interactome identified NF90/NF45 heterodimer as an interactant of Ago2 in the cytoplasm (Höck et al., 2007), which could suggest a role for NF90/NF45 is RISC-mediated activities.

For Ago2 lateral diffusion to be efficient, it needs the activity of a helicase to disrupt occlusive secondary RNA structures that could interfere with its binding. Moloney leukemia virus 10 (MOV10) is an ATP-dependent helicase that belongs to the Up frameshift (UPF)-like helicase superfamily 1 (SF1). MOV10 binds ssRNA and translocates 5′-to-3′ along the target RNA (Meister et al., 2005). It was initially described to be involved in the inhibition of viral replication for HIV-1 and Hepatitis C viruses but also as inhibitor of LINE-1 retrotrasposition (Meister et al., 2005). In addition, MOV10 was found to co-localize in P-bodies together with Ago2 and other factors involved in RISC, identifying a role of MOV10 in miRNA-mediated regulation (Meister et al., 2005). MOV10 binds in close proximity to UPF1 binding sites, to resolve structures and displace RBPs from the 3′ UTR of the target mRNA, exposing the miRNA recognition element (MRE) for Ago2

binding (Gregersen et al., 2014).

MOV10 frequently binds to 3′ UTRs within mRNAs, specifically at regions with low conservation and upstream of local secondary structures, consistent with its 5′-to-3′ directional unwinding activity (Gregersen et al., 2014). Moreover, for efficient loading and unwinding, MOV10, like many RNA helicases, needs a single-stranded region adjacent to a duplex (Q. Yang & Jankowsky, 2006). Consistent with its role in RISC-mediated silencing, MOV10 can regulate the abundance of its bound mRNAs. In particular, it was shown that KD of MOV10 inhibits translational suppression leading to a global stabilization of its target mRNAs (Gregersen et al., 2014). However, in contrast with this observation, MOV10 was also found to increase the expression of a limited subset of mRNAs by inhibiting Ago2 binding, in the presence of fragile X mental retardation protein 1 (FMRP1). Therefore, the concomitant binding of MOV10 and FMRP on the same mRNA can inhibit the canonical role of MOV10 in RISC (Kenny et al., 2020; Kenny et al., 2014).

Here, we show that cytoplasmic NF90 interacts with proteins involved in translational repression, RNA stability, degradation and viral replication. We determined that NF90 interacts with MOV10 in an RNA-dependent fashion and both factors can be found in the same complex with Ago2, upon glycerol gradient sedimentation. Using published CLIP data of MOV10 and NF90, we showed that both proteins can bind the same target mRNAs using RNA immunoprecipitation (RIP) analysis. Upon loss of MOV10, association of NF90 with the targets increased. Similarly, after loss of NF90/NF45, we detected an increase in the association of MOV10 to the selected target mRNAs. To determine whether NF90 binding might impact association of RISC with the mRNA targets, we performed RIP analysis for Ago2. Association of Ago2 with target mRNAs increased following loss of NF90. The abundance of the target mRNAs was also increased. These data suggest that NF90 may be involved in RISC-mediated gene silencing by regulating MOV10 and Ago2 association with target mRNAs.

Material and methods

Cell culture, stable cell line production and cellular treatments

Human HEK293T cell line was grown in Dulbecco's Modified Eagle's medium – high glucose with HEPES modification (Sigma-Aldrich®, D6171) supplemented with 10% fetal bovine serum (PAN Biotech, 8500-P131704), 1% penicillin-streptomycin (v/v) (Sigma Aldrich®, P4333) and 1% L-glutamine (v/v) (Sigma Aldrich®, G7513). Cells were cultured at 37°C in a humidified atmosphere containing 5% CO_2 and seeded at 3×10^6 cells in 100mm culture dishes the day of siRNA transfection.

Plasmid encoding pOZ-NF90-FLAG-HA (pOZ-NF90-FH) were cloned using pOZ-N-FH vector, as previously explained (Contreras et al., 2018). Briefly, lentiviral particles expressing NF90 were produced in HEK293T cells by transfecting plasmids using calcium-phosphate. HEK293T were transduced using Polybrene infection/transfection reagent (Sigma-Aldrich®, TR-1003), according to the manufacturer's instructions. After 7 days, selection of transduced cells was carried out by magnetic affinity sorting with antibody against IL2 to achieve a pure population. HEK293T stably expressing NF90 (pOZ-NF90-Flag-HA HEK293T) was grown in the same conditions as HEK293T. pOZ-NF90-FH HEK293T were seeded at 1×10^7 in 150mm culture dishes the day prior to protein extraction.

Transfection of small interfering RNAs

Double-stranded RNA oligonucleotides used for RNAi were purchased from Eurofins MWG Operon or Integrated DNA Technologies. Sequences of small interfering RNAs (siRNAs) used in this study have been described previously (Grasso et al., 2020) and are shown in Supplementary Table S1. HEK-293T cells were transfected with siRNA (30 nM final concentration) using INTERFERin® siRNA transfection reagent (Polyplus Transfection) according to the manufacturer's instructions. The transfection was carried out the day of seeding and cells were collected for protein extraction approximately 65 h after transfection.

Immunoblot

HEK293T and pOZ-NF90-FH HEK293T were lysed using RIPA buffer (50 mM Tris-HCl pH=7.5, 150 mM NaCl, 1% NP40, 0.5% Sodium Deoxycholate, 0.1% SDS, Halt™ Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific)), unless otherwise indicated. Protein extracts were immunoblotted using the indicated primary antibodies (Table S2) and anti-mouse, anti-rabbit or anti-rat IgG-linked HRP secondary antibodies (GE Healthcare) followed by ECL (Advansta).

Cytoplasmic extracts and co-IPs

HEK293T were seeded in 150mm culture dishes the day prior to protein extraction. Cytoplasmic proteins were extracted using a mild lysis buffer (10mM Hepes pH 7.9, 10mM KCl, 0.1mM EDTA pH 8.0, 2mM $MgCl_2$, 1mM DTT, EDTA-free protease and phosphatase inhibitor). The cell pellet was incubated for 10 min on ice, adding 0.07% of NP-40 and incubating for additional 10 min on ice. After centrifugation, 1 mg of lysates were incubated at 4°C overnight with 2g of antibodies recognizing NF90, AGO2 and IgG controls and protein A/G PLUS-Agarose beads (Santa Cruz, sc-2003). Beads were then washed twice with IP buffer (150 mM KCl, 20 mM Tris pH 7.5, 0.05% of NP-40, 0.1% Tween, 10% glycerol, 5 mM $MgCl_2$, 1 mM DTT and EDTA-free protease and phosphatase inhibitor). Samples were treated with RNAse A/T1 mix (ThermoFisher Scientific, EN0551) for 30 min at room temperature, incubating on a rotating wheel. After incubation, beads were washed three times with IP buffer as aforementioned and 2X Laemmli buffer was added directly to the beads.

Tandem Immunoprecipitation and mass spectrometry

For mass spectrometry, cytoplasmic extracts were obtained as aforementioned. Tandem immunoprecipitation (Flag and HA) was carried out using 10 mg of cytoplasmic extract. Flag IP was performed using EZview™ Red ANTI-Flag® M2 Affinity gel (SigmaAldrich, F2426), following the manufacturer's instructions. Washes were carried out 3 times as aforementioned and protein complexes were eluted by competition performing 2 consecutive elutions using Flag elution buffer (250ng/µl FLAG® Peptide (SigmaAldrich, F3290), diluted in IP buffer) incubating

for 1 h at 4°C on a rotating wheel. HA IP was performed using the elutions obtained from the first IP incubated with Pierce™ Anti-HA Agarose beads (ThermoScientific, 26181) for 2 h at 4°C on a rotating wheel. Washes were carried out 5 times as aforementioned and elutions were performed using HA elution buffer (400 ng/ µl HA peptide (ThermoScientific, 26184), diluted in IP buffer) for 1 h at 4°C on a rotating wheel. Following elution, beads were removed using Pierce™ Centrifuge Columns (ThermoScientific, 11894131), as specified by manufacturer's instructions. Silver-staining was performed according to the manufacturer's instruction (Silverquest, Invitrogen). Mass spectrometry was performed at Taplin facility, Harvard University, Boston, MA.

Glycerol gradient sedimentation

Glycerol gradient sedimentation was performed as described previously (Contreras et al., 2018). Briefly, NF90-associated proteins were purified by performing FLAG IP on pOZ-NF90-FH HEK293T cytoplasmic fraction, as aforementioned. One ml layers of glycerol (final concentration 15 to 35%–20 mM Tris pH 7.5, 0.15 M KCl, 2.5 mM $MgCl_2$, 0.05% NP-40, 0.1% Tween) were layered into centrifugation tubes (13 x 51 mm Ultra-Clear Tubes, Beckman). A linear gradient was obtained after 12 h of diffusion at 4°C. Flag elution from pOZ-NF90-FH HEK293T immunoprecipitation was loaded on top of the glycerol gradient. Complexes were fractionated by ultracentrifugation in an SW 55Ti rotor (Beckman) at 30,000 rpm for 18 h at 4°C. 25 fractions of 200 L were collected from top of the gradient. An equal volume of fractions was resolved by SDS-PAGE and immunoblotted with indicated antibodies.

RNA immunoprepicipation (RIP)

RIP was performed as previously described (Grasso et al., 2020). Briefly, HEK293T were seeded in 100 mm culture dishes and transfected with siRNAs the day of seeding, as aforementioned. Cells were harvested 65 h after the treatment and lysed for 10 min in RIP buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 2.5 mM $MgCl_2 \bullet 6H_2O$, 250 mM sucrose, 0.05% (v/v) NP-40 and 0.5% (v/v) Triton X-100)

containing 20 U ml[1] of RNasin (Promega), 1 mM DTT, 0.1 mM PMSF and EDTA-free protease and phosphatase inhibitor. After centrifugation, lysates were incubated for 4 h at 4^{circ}C with 2 g of antibodies recognizing NF90, MOV10, AGO2 and IgG control and then incubated for 1 h at 4^{circ}C with Dynabeads™ Protein A or G (ThermoFisher Scientific). After incubation, beads were washed five times with RIP buffer for 5 min at 4^{circ}C and RNA was extracted using TRIzol (Thermo Fisher Scientific) according to the manufacturer's instructions. RNA was treated with DNAse I (Promega) and RT was performed using SuperScript™ III Reverse Transcriptase (ThermoFisher Scientific) according to the manufacturer's instructions. cDNA was treated with RNAse H (ThermoFisher Scientific) and the samples were used to perform qPCRs using LightCycler™ 480 SYBR Green I Master mix (Roche), according to the manufacturer's instructions, using the primers shown in Supplementary Table S3.

Bioinformatic analyses

Enhanced UV crosslinking followed by immunoprecipitation (eCLIP) data for NF90 were obtained from Nussbacher and Yeo (Nussbacher & Yeo, 2018) and retrieved from the NCBI database (NF90 eCLIP: ENCSR786TSC). Individual-nucleotide-resolution UV crosslinking and (iCLIP) data for MOV10 were obtained and retrieved from the NCBI database (MOV10 iCLIP: GSE51443). MOV10 iCLIP data was lifted to hg38 genome annotation using UCSC liftOver tool. Peaks were filtered based on Fold Change (FC 1.5) and p-value (Bonferroni-Adj P-val 0.05). Bigwig files from different replicates were merged using bigWigMerge v2. Statistical analyses were performed using RStudio v1.4. Gene ontology was performed using DAVID Functional Annotation Tool database version 6.8 (https://david.ncifcrf.gov).

Results

NF90 interacts with proteins involved in translational repression and RNA processing

In order to understand the role of NF90 in the cytoplasm, we determined its cytoplasmic interactome by performing mass spectrometry on a HEK293T cell line stably overexpressing NF90-FLAG-HA (Figure 1A), after tandem affinity purification of the cytoplasmic protein fraction (Figure 1B). Excluding the proteins for which one peptide or more was found in the mock, 319 proteins were detected associated with NF90 in the cytoplasm (Table S4). As expected, the most abundant protein identified was NF90, followed by the 5'-to-3' helicase MOV10 and ILF2 (NF45), a well-known NF90 protein partner. Ago2, the main component of RISC, was also associated with NF90, as reported previously (Höck et al., 2007). Gene ontology of the significantly enriched NF90 interactants identified a number of pathways, such as translation, translational initiation and its regulation, mRNA processing and regulation of mRNA stability (Figure 1C). These findings are consistent with previous observations suggesting the implication of NF90 in the regulation of mRNA stability and mRNA translation for specific target mRNAs (Idda et al., 2018; W. Zhang et al., 2018).

NF90 interacts with MOV10 and Ago2 in an RNA-dependent manner

Since NF90 interactants are well-known RNA-binding proteins (Supplementary Figure S1), we wondered if their interaction with NF90 was RNA dependent. In order to validate the results of mass spectrometry and to determine if the binding of NF90 is RNA dependent, we performed co-IP in HEK293T cell line, with and without RNAse A/T treatment. The results showed that the binding of NF90 to PACT, MOV10 and DICER is RNA-dependent (Figure 2A). On the other hand, the binding of NF90 to NF45 was RNA-independent (Figure 2A), as previously reported (Guan et al., 2008; Wolkowicz & Cook, 2012). Although the binding of Ago2 to NF90 was previously reported to be RNA-independent (Höck et al., 2007), under the conditions used in this study, NF90 binding to Ago2 was RNA-dependent, in agreement with a recent study (Idda et al., 2018). The

interaction of NF90 with MOV10, DICER or PACT has not been reported previously.

To determine if these interactants can be found in the same complex with NF90, we performed glycerol gradient sedimentation of immunopurified NF90-Flag-HA complexes (Figure 2B and S2). NF90 was detected in the top 10 fractions (data not shown) where it co-sedimented with ADAR, MOV10, AGO2, DDX6, PACT and NF45 (Figure 2B). In agreement with these results, NF90 was already found to co-sediment with Ago2, NF45 and other helicases (DDX47, DDX36, DDX30) on a sucrose gradient (Höck et al., 2007). These findings suggest that, in the cytoplasm, NF90 can bind Ago2, MOV10, PACT and DICER through RNA. In addition, these proteins can be found in the same complex with ADAR, DDX6 and NF45.

NF90 and MOV10 can bind the same target mRNAs

To further understand the interaction between NF90 and MOV10, we took advantage of an enhanced UV crosslinking followed by immunoprecipitation of NF90 (eCLIP) (Nussbacher & Yeo, 2018) and an individual-nucleotide resolution UV crosslinking and IP of MOV10 (iCLIP) (Kenny et al., 2014) datasets. Analysis of MOV10 iCLIP identified 1103 mRNAs significantly bound by MOV10. Approximately half of the bound mRNAs (542 mRNAs) showed a MOV10 peak in the 3'UTR (Figure 3A), consistent with previous findings suggesting that MOV10 mainly binds 3'UTRs of mRNAs (Gregersen et al., 2014). The remaining MOV10-bound mRNAs contained peaks in introns (279 mRNAs), exons (255 mRNAs) and 5' UTRs (27 mRNAs) (Figure 3A).

On the other hand, analyses of NF90 eCLIP suggest that the majority of mRNAs significantly associated with NF90 were bound in their introns (3217 out of 3942 NF90-bound mRNAs). However, some mRNAs were also bound in their in exons, 5' UTRs and 3' UTRs (301, 70 and 354, respectively) (Figure 3B).

Since the interaction between NF90 and MOV10 was found to be RNA-dependent (Figure 2A), we wondered if these two proteins could bind the same mRNAs.

Intersection of the mRNAs bearing at least one peak of NF90 and MOV10 identified 456 mRNAs that could be potentially bound by both proteins (Figure 3C). Interestingly, this corresponds to around 41% of all mRNAs bound by MOV10 and it is significantly enriched (P= 1.7e-48, Fisher's exact test). Next, since the helicase activity of MOV10 implicated in RISC-mediated silencing occurs in 3' UTRs, we wondered whether NF90 and MOV10 were associated with the 3'UTR of common mRNAs. We identified 52 mRNAs that bear at least one peak of both NF90 and MOV10 in their 3' UTRs (Figure 3D), which is significantly enriched (P=3.21e-21, Fisher's exact test). These findings suggest that NF90 and MOV10 can bind the 3' UTR of a common set of target mRNAs. MOV10-associated mRNAs usually showed one peak of MOV10, while NF90-bound mRNAs often bore more than one peak of NF90 (data not shown). Importantly, MOV10 is known to bind single-stranded regions adjacent to duplex regions, like other RNA helicases (Gregersen et al., 2014; Gross & Shuman, 1996), while NF90 has been shown to bind structured RNAs (Gwizdek et al., 2004). This observation might potentially explain the difference in the number of MOV10 peaks found in MOV10-associated mRNAs, usually one, and the number of NF90 peaks found in NF90-associated mRNAs, often more than one. Target mRNAs to further analyze in this study were chosen from the mRNAs that host at least one peak of both NF90 and MOV10 in their 3'UTRs. In order to avoid possible artefacts due to multiple NF90 binding sites in one mRNA, we chose target mRNAs bearing only one peak of both NF90 and MOV10 in their 3'UTRs (Figure S3).

NF90 and MOV10 influence each other's binding to target mRNAs

MOV10 has been shown to facilitate RISC-mediated silencing (Meister et al., 2005) while, on the contrary, NF90 was found to increase specific target mRNAs stability (Vumbaca et al., 2008). We therefore wondered if NF90 could interfere with the binding of MOV10 to the target mRNAs and vice versa. Therefore, in order to understand the function of NF90 and MOV10 binding to the same target mRNAs, we performed RNA immunoprecipitation (RIP) in HEK293T cell line after treatment with a non-targeting control siRNA (siScr), an siRNA targeting NF90 (siNF90) or MOV10 (siMOV10), followed by quantitative PCR (qPCR), on target

mRNAs selected on the basis of MOV10 iCLIP and NF90 eCLIP analyses (Figure 3D).

We performed RIP of NF90, MOV10 or IgG control after downregulation of MOV10 (Figure S4A). Loss of MOV10 did not significantly affect the total level of the target mRNAs (Figure S4B). RIP results revealed that MOV10 association with the selected target mRNAs was significantly decreased after downregulation of MOV10, as expected (Figure 4). On the other hand, NF90 binding to the same target mRNAs was significantly increased after downregulation of MOV10, while its binding to the negative control, H2BC1, did not significantly change (Figure 4).

In order to further investigate this mechanism, we performed RIP of NF90, MOV10 or IgG control after downregulation of NF90 and its protein partner NF45 (Figure S5A). Interestingly, the loss of NF90/NF45 significantly decreased the total level of the selected target mRNAs (Figure S5B), consistent with its role in increasing mRNA stability (Vumbaca et al., 2008). As expected, NF90 association with the target mRNAs was significantly decreased after NF90/NF45 downregulation (Figure 5). On the other hand, RIP results revealed that NF90/NF45 downregulation lead to a significant increase in MOV10 association with the selected target mRNAs while its binding to the negative control, H2BC1, did not significantly change (Figure 5).
These results suggest that the binding of NF90 and MOV10 at common target mRNAs is mutually influenced by the presence of the other factor.

Downregulation of NF90/NF45 complex increases Ago2 binding to the target mRNAs

MOV10 is known to promote Ago2 association to mRNAs, enhancing RISC-mediated silencing. Since NF90 and MOV10 influence each other's binding to common target mRNAs, we wondered if alteration of the association NF90/NF45 complex could have similar effects on Ago2 binding. To this end, we performed RIP of Ago2 or IgG control after downregulation of NF90/NF45. Depletion of NF90/NF45 had no effect of Ago2 expression in extracts or immunoprecipitates

(Figure S6A, B). However, loss of NF90/NF45 significantly increased Ago2 binding to the target mRNAs tested, while the negative control mRNA, H2BC1, which was poorly associated with Ago2, was not significantly increased relative to the IgG control (Figure 6).

Discussion

The dsRNA binding protein, NF90, has been implicated in a number of cellular pathways, including regulation of translation and RNA stability. However, the precise mechanisms involved in its different functions are not entirely clear. To better understand how NF90 performs these roles, we identified the interactome of cytoplasmic NF90. Identification of its partners may shed light on mechanisms by which NF90 is implicated in different pathways. Gene ontology analysis of the interactants revealed, not surprisingly, that almost all NF90 partners are involved in pathways in the processing, stability or translation of cellular RNA, consistent with the known functions of NF90.

Interestingly, pathways associated with viral transcription and viral translation (Figure 1C) were among the significantly enriched gene ontology terms. This is notable since it has been shown that NF90 translocates from the nucleus to the cytoplasm as a consequence of viral infection and, following its nuclear export, NF90 was shown to bind viral mRNAs, playing a role in the antiviral immune response (X. Li et al., 2017). We previously showed that nuclear NF90 is able to inhibit the biogenesis of several miRNAs involved in viral replication and antiviral response (Grasso et al., 2020), such as miR-4753 and miR-3145. Consistently, these latter were shown to be overexpressed in response to viral infection, inhibiting viral transcription and replication (Khongnomnan et al., 2015). Thus, the identification of cytoplasmic NF90 partners involved in viral transcription and translation suggests that NF90 may play an important role in the cellular response to viral infection.

NF90, has been implicated in the regulation of translation and mRNA stability. Moreover, a recent study put forward the hypothesis that this might be due to its involvement in miRNA-mediated gene silencing (Idda et al., 2018). However, a

direct role for NF90 in RISC-mediated silencing had not so far been demonstrated. The interactome of cytoplasmic NF90 identified interactions with several components of RISC. These include DHX30, UPF1, DDX6 as well as the RISC-associated RNA helicase, MOV10. The effector of RISC-mediated silencing, Ago2, was also identified among NF90 interactants, which confirms previous reports identifying an interaction between NF90 and Ago2 (Höck et al., 2007; Idda et al., 2018) . We furthermore determined that the association of NF90 with MOV10 and Ago2 occurs through RNA. NF90 likely exists in a cytosolic complex with RISC and RISC-associated factors since these factors co-sediment in a glycerol gradient. These data suggest that NF90 may be linked to RISC-mediated silencing through its interaction with RISC-associated proteins.

Analysis of NF90 and MOV10 CLIP data suggested that both proteins can associate with the 3′ UTR of a subset of target mRNAs. However, NF90 was shown to increase mRNA stability while MOV10 enhances RISC-mediated silencing. Since these effects appear contradictory, we wondered whether NF90 and MOV10 could compete for the binding of selected target mRNAs. Upon downregulation of MOV10, an increase in the association of NF90 to the target mRNAs was detected. Likewise, upon loss of the heterodimer NF90/NF45, an increase in MOV10 binding to the same mRNAs was measured. These findings suggest that NF90 and MOV10 might influence or interfere with the ability of the other factor to associate with its target RNAs. This interference is probably unlikely to occur through direct competition for the same binding sites, as reported for NF90 and Microprocessor binding to pri-miRNAs in the nucleus (Grasso et al., 2020; Sakamoto et al., 2009). Indeed, the binding preferences for NF90 and MOV10 differ significantly. MOV10 binds ssRNA upstream of a structured region while NF90 appears to bind highly stable hairpin structures (Grasso et al., 2020; Gwizdek et al., 2004). The binding of one factor may modify RNA structure to disfavor binding of the other. Alternatively, one factor may recruit additional proteins that may influence the binding of the other factor. In any case, NF90 and MOV10 mutually interfere with each other's binding to a subset of common target mRNAs.

MOV10 helicase activity is known to resolve mRNA structures in order to reveal obscured MREs within 3' UTRs. This facilitates the binding of Ago2, which favors RISC-mediated silencing. Interestingly, depletion of NF90 significantly increased the association of Ago2 with target mRNAs. Consistent with increased Ago2 association, downregulation of the total level of the selected target mRNAs upon loss of NF90 was detected. These data suggest that NF90 enhances the stability of certain mRNAs by modulating the ability of Ago2 to associate with its target site and induce RISC-mediate silencing. It is interesting to note that NF90 has been identified as a subunit of P-bodies (Hubstenberger et al., 2017). P-bodies are a site of mRNA storage as well as RISC-mediated mRNA degradation. It is tempting to speculate that NF90 within P-bodies may be implicated in the control of mRNA stability versus degradation by modulating the binding of RISC. It would also be interesting to determine whether the destabilization of mRNAs observed upon loss of NF90 occurs within P-bodies.

Since NF90 is known to translocate to the cytoplasm during viral infection or cancer-induced hypoxia, it would be tempting to speculate that, under these conditions, NF90 might displace Ago2, leading to inhibition of RISC-mediated silencing for some specific target mRNAs. Therefore, a deeper understanding of how NF90 is implicated in RISC-mediated silencing could potentially elucidate its effect on the fate of mRNAs involved in the antiviral immune response or during hypoxia induced in solid tumors.

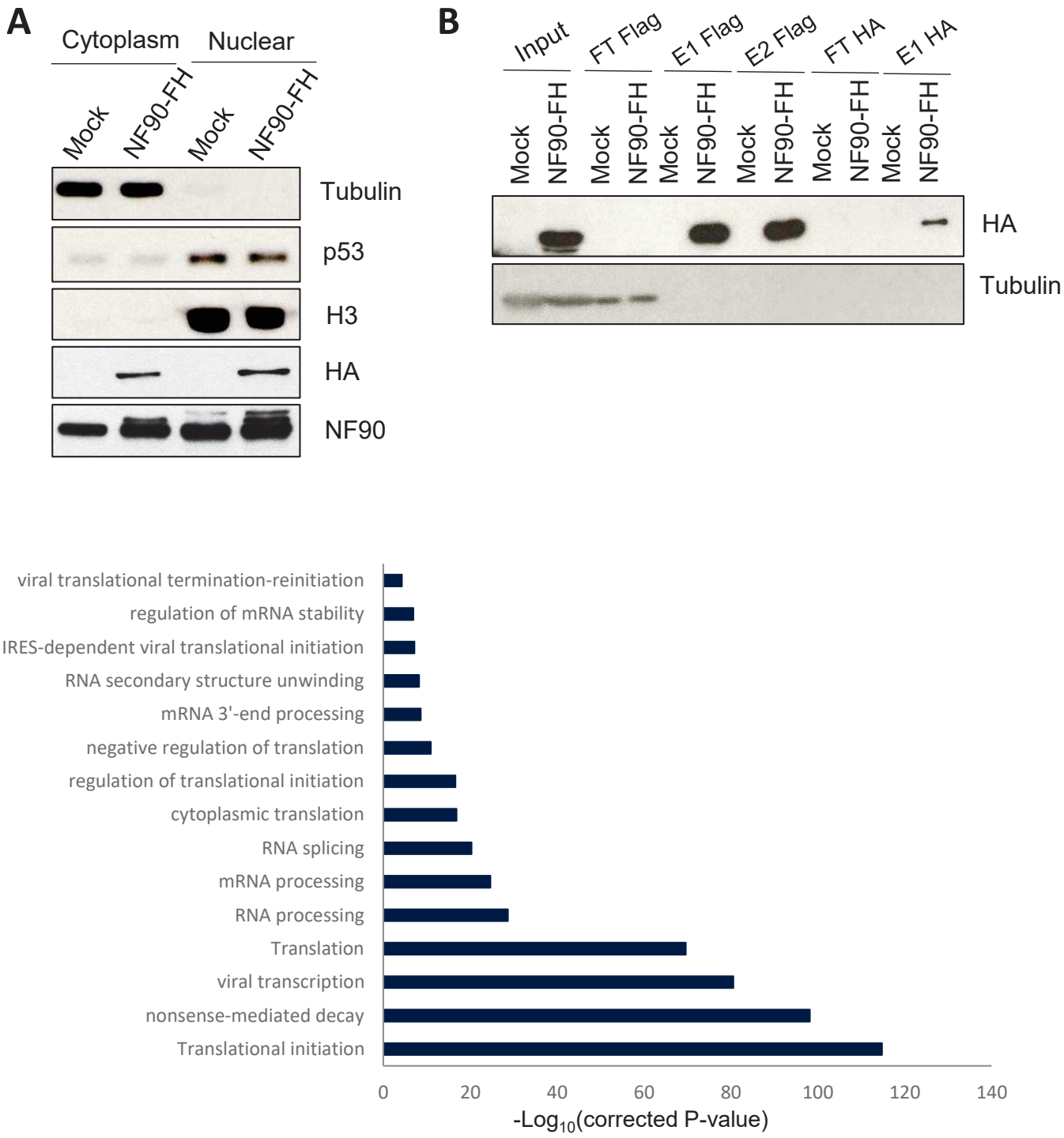


Figure 1. NF90 interacts with proteins involved in translational repression and RNA processing. **(A)** Cytoplasmic and Nuclear extracts of WT and NF90-FH stably overexpressing HEK-293T cells were analyzed by Western blot using the indicated antibodies. **(B)** Samples described in A underwent tandem affinity purification using FLAG and HA antibodies. Results were analyzed by western blot using the indicated antibodies. **(C)** Gene ontology analyses (cellular pathways) of NF90-associated proteins, found by mass spectrometry.

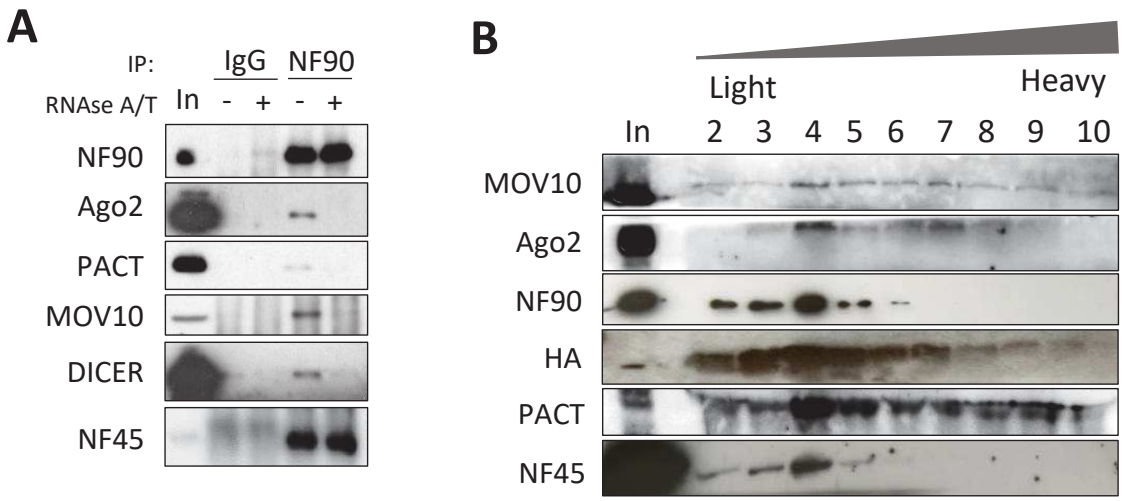


Figure 2. NF90 complexes with MOV10 and Ago2 in an RNA-dependent manner. **(A)** Cytoplasmic extracts from HEK-293T cells were used for immunoprecipitation using an anti-IgG control or anti-NF90 antibody followed by RNAse A/T treatment or mock treatment, as indicated. Immunoprecipitates and an aliquot of extract (In) were analyzed by western blot, using the indicated antibodies. **(B)** FLAG immunoprecipitates of cytoplasmic extracts of NF90-FH overexpressing HEK293T cells were separated by glycerol gradient sedimentation. NF90-containing fractions and an aliquot of extract (In) were analyzed by western blot using the antibodies indicated.

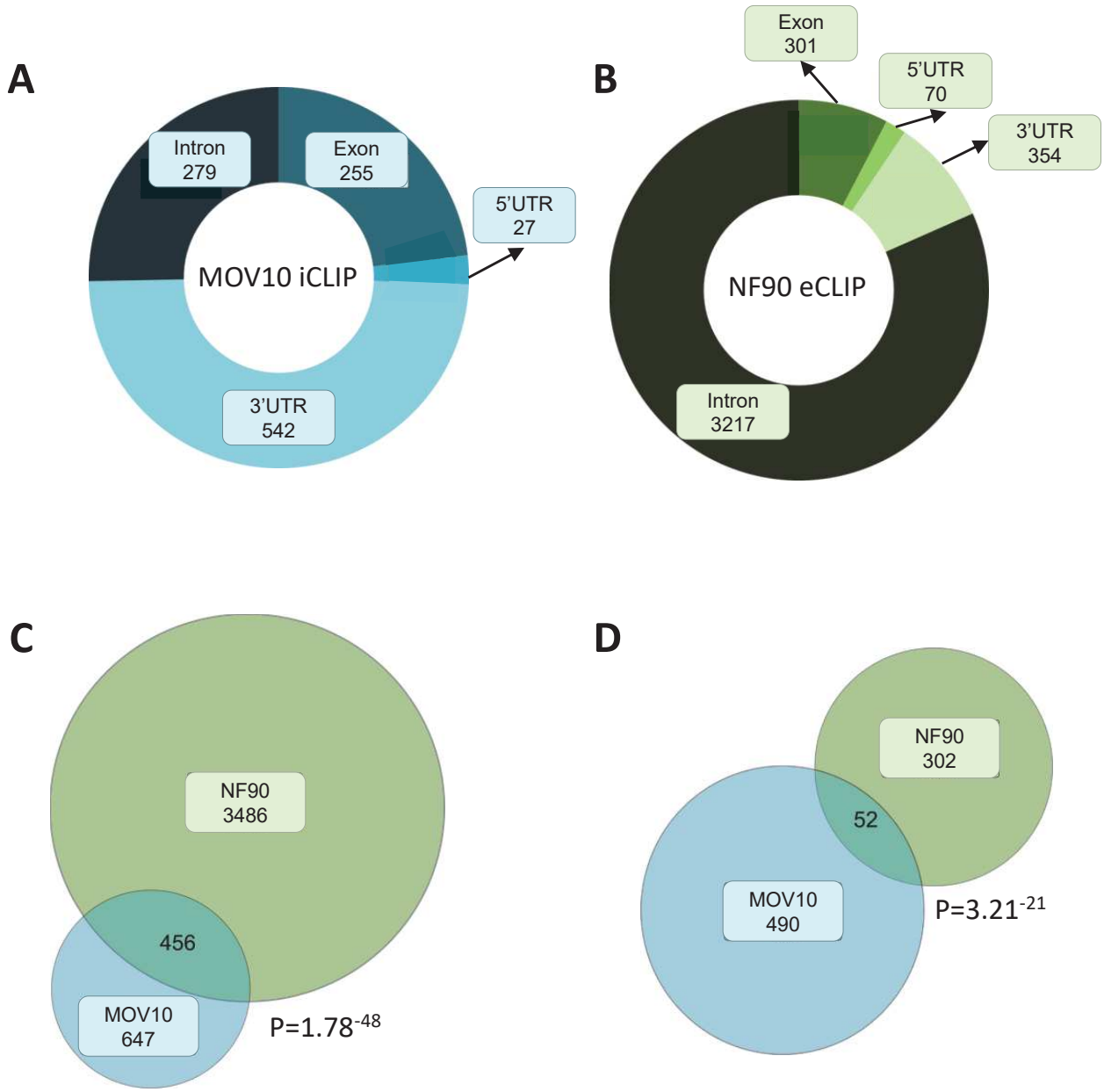


Figure 3. NF90 and MOV10 can bind common target mRNAs. **(A)** MOV10 iCLIP data were analyzed to show the distribution of MOV10 binding along MOV10-associated mRNAs. **(B)** NF90 eCLIP data were analyzed to show the distribution of NF90 binding along NF90-associated mRNAs. **(C)** MOV10 iCLIP and NF90 eCLIP were correlated to detect target mRNAs that are bound by both factors. Statistical analysis were carried out using Fisher's exact test. **(D)** MOV10 iCLIP and NF90 eCLIP were correlated to detect target mRNAs that are bound by both factors at their 3' UTRs. Statistical analysis were carried out using Fisher's exact test.

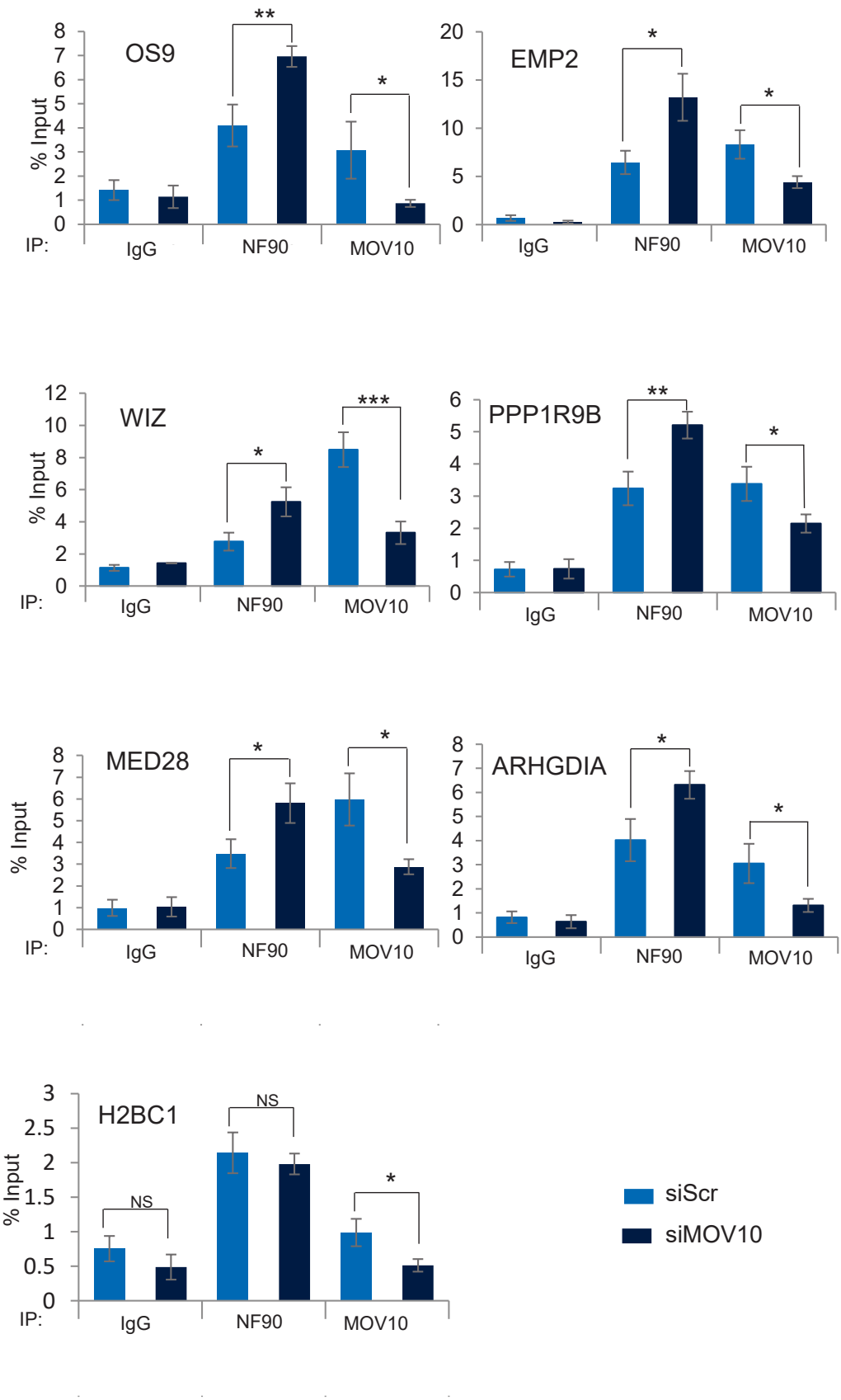


Figure 4. MOV10 modulates NF90 association with common target mRNAs. RIP analysis of HEK293T cells transfected with MOV10-targeting siRNA or a non-targeting control (Scr), as indicated. RIPs were performed using anti-NF90, anti-MOV10 or a control IgG antibody. Immunoprecipitates were analyzed using RT-qPCR. Data represent mean ± SEM obtained from 4 independent experiments *(*P < 0.05, **P < 0.01, ***P < 0.001*, independent Student's *t* test).

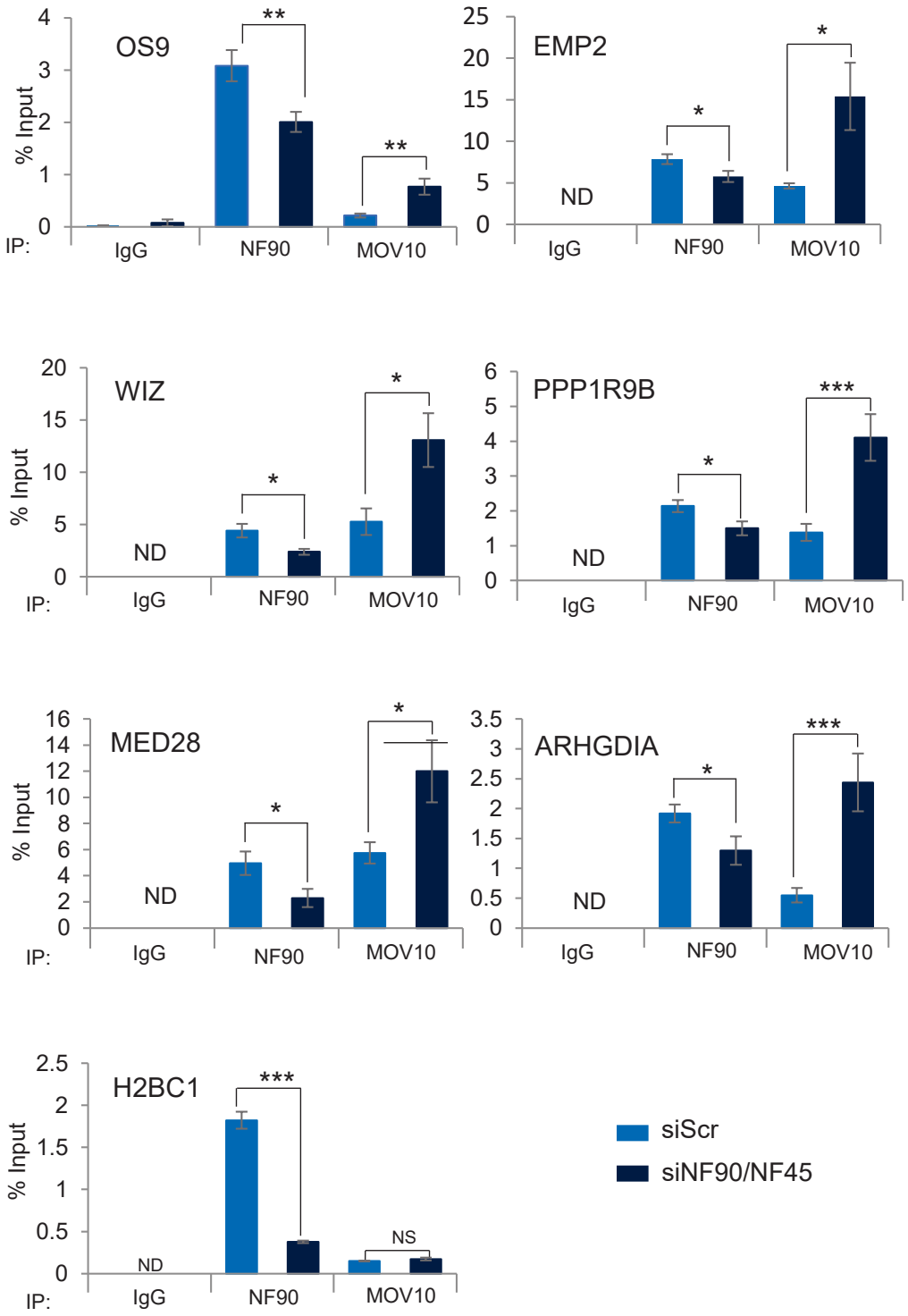


Figure 5. MOV10 modulates NF90 association with common target mRNAs. RIP analysis of HEK293T cells transfected with NF90/NF45-targeting siRNAs or a non-targeting control (Scr), as indicated. RIPs were performed using anti-NF90, anti-MOV10 or a control IgG antibody. Immunoprecipitates were analyzed using RT-qPCR. ND indicates 'Not Detected'. Data represent mean ± SEM obtained from 4 independent experiments *(*P < 0.05, **P < 0.01, ***P < 0.001*, independent Student's *t* test).

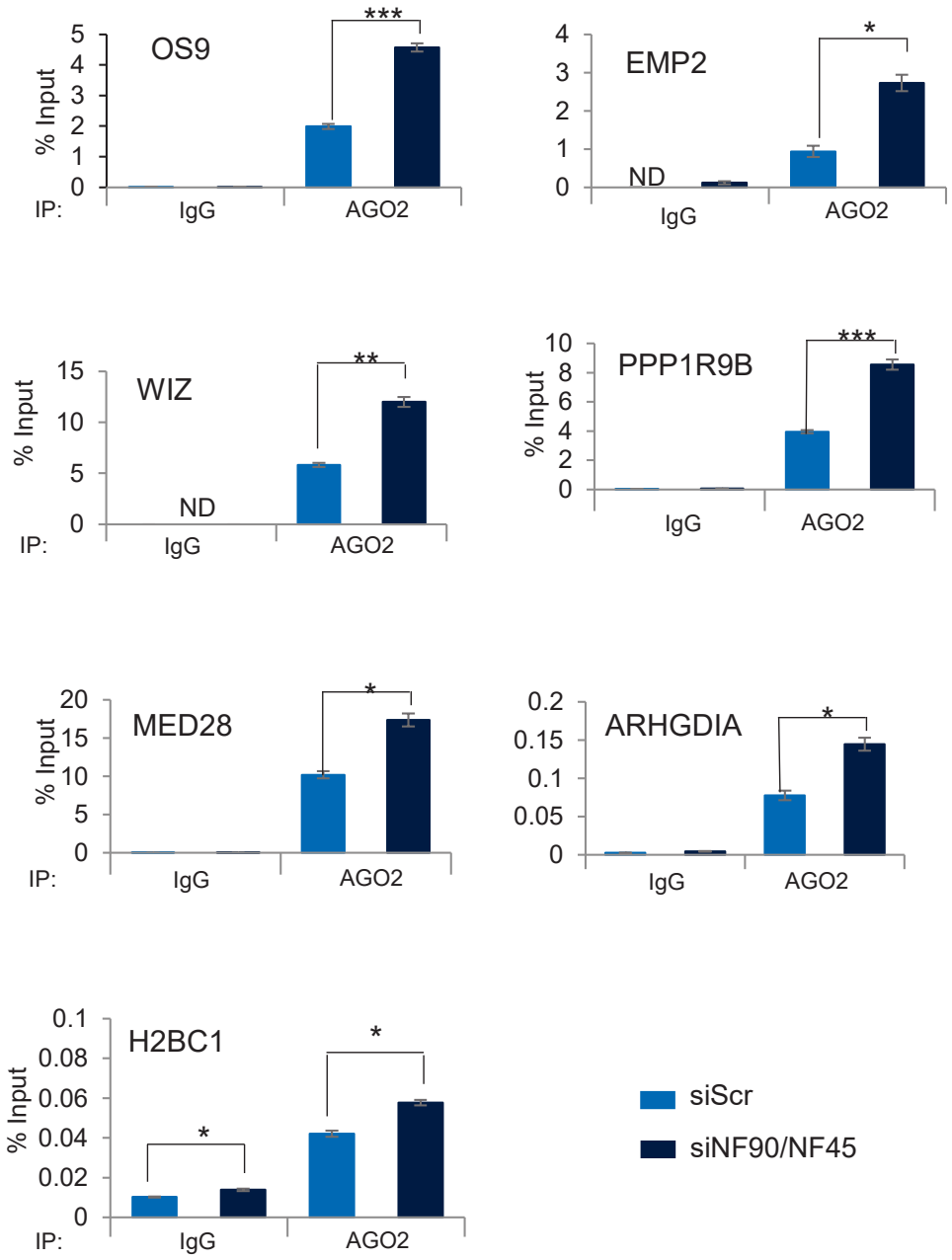


Figure 6. Downregulation of NF90/NF45 complex increases Ago2 binding to selected target mRNAs. RIP analysis of HEK293T cells transfected with NF90/NF45 targeting siRNAs or a non-targeting control (Scr), as indicated. RIPs were performed using anti-Ago2 or a control IgG antibody. Immunoprecipitates were analyzed using RT-qPCR. ND indicates 'Not Detected'. Data represent mean ± SEM obtained from 3 independent experiments *(*P < 0.05, **P < 0.01, ***P < 0.001*, independent Student's *t* test).

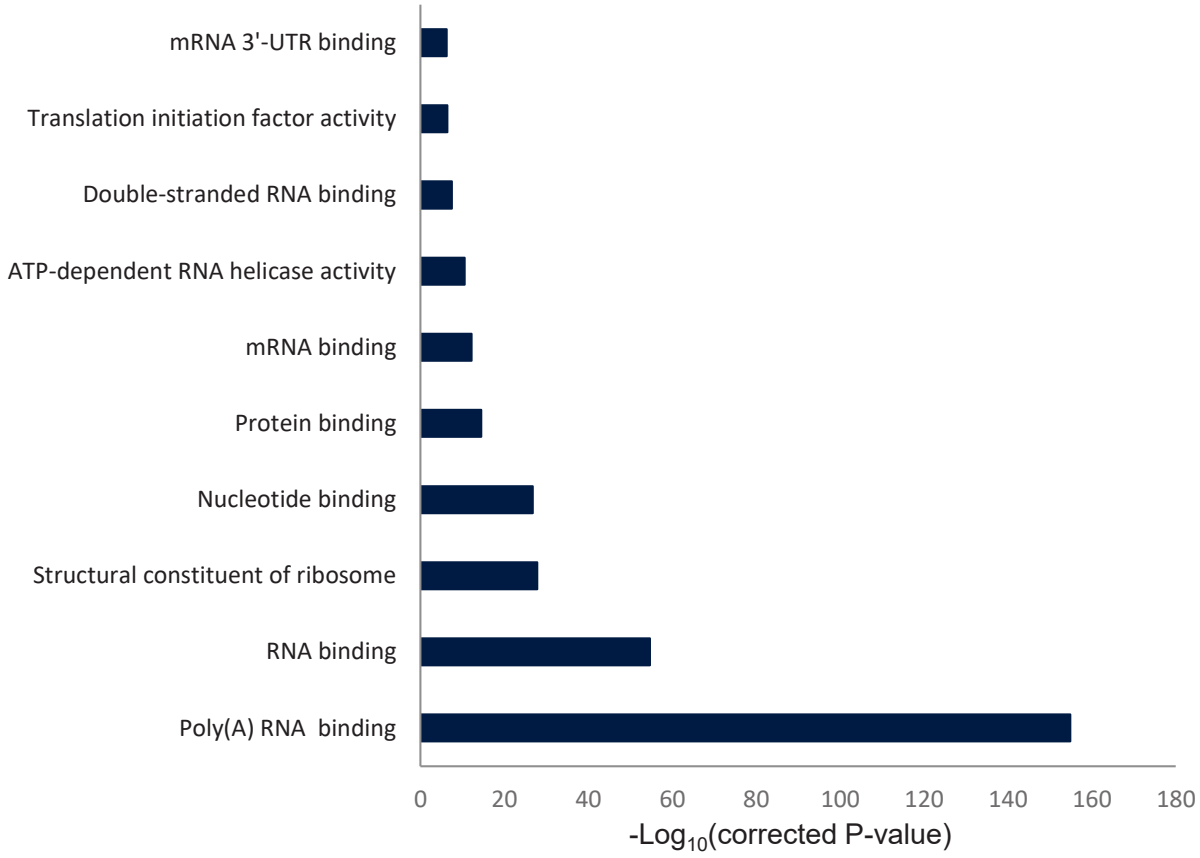


Figure S1. NF90 interacts with proteins involved in translational repression and RNA processing. Molecular functions of NF90-associated proteins, found by mass spectrometry, were analyzed using gene ontology.

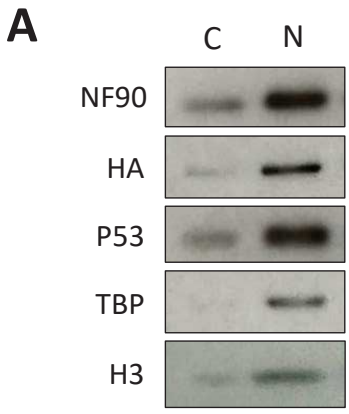


Figure S2. NF90 complexes with MOV10 and Ago2 in an RNA-dependent manner. **(A)** Cytoplasmic (C) and nuclear (N) extracts from NF90-FH overexpressing HEK-293T cells were analyzed by western blot, using the antibodies indicated

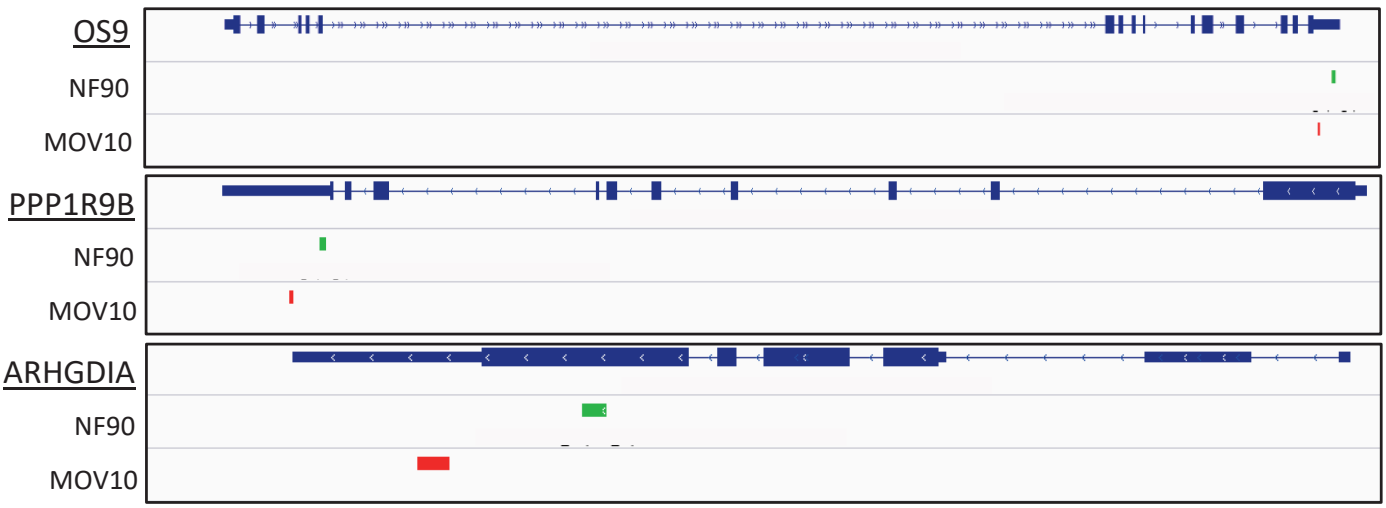


Figure S3. NF90 and MOV10 can bind the same target mRNAs. Screenshots of NF90 and MOV10 eCLIP and iCLIP, respectively, in some of the selected target mRNAs.

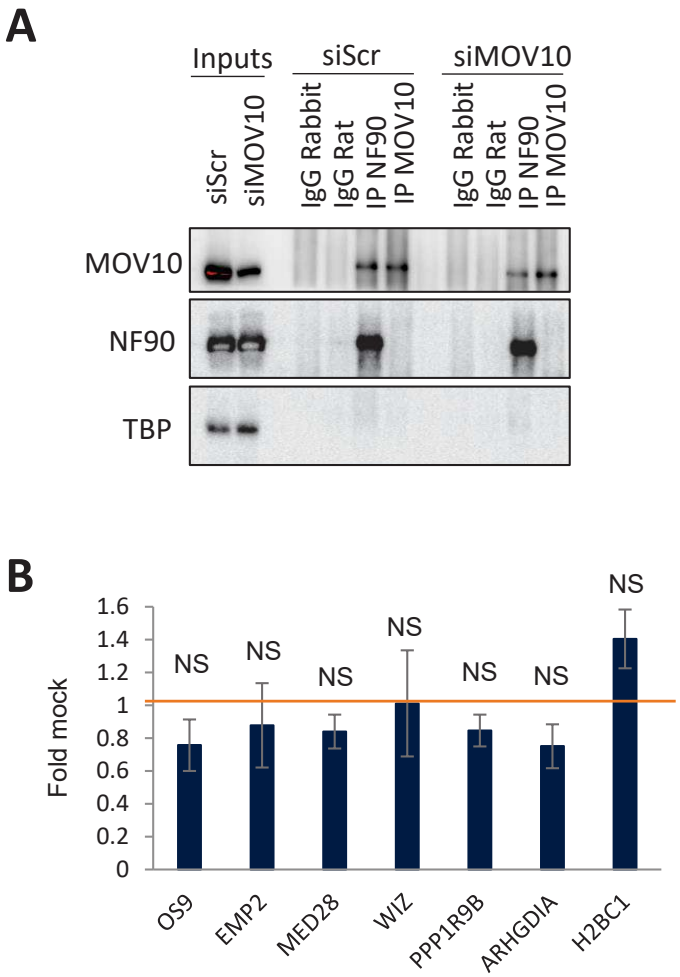


Figure S4. MOV10 modulates NF90 association with common target mRNAs. **(A)** Extracts of HEK293T cells transfected with siRNAs targeting MOV10 or a non-targeting control (Scr) and immunoprecipitates obtained using antibodies anti-NF90, anti-MOV10 or control IgG were analyzed by Western blot using the indicated antibodies. **(B)** Total RNA obtained from HEK293T transfected with siRNAs targeting MOV10 or a non-targeting control (Scr) was analyzed by RT-qPCR. NS indicates 'Not Significant'. Data represent Fold Mock (IgG) relative to the control samples (siScr), which was attributed a value of 1 (red line), obtained from 4 independent experiments *(*P < 0.05, **P < 0.01, ***P < 0.001*, independent Student's *t* test).

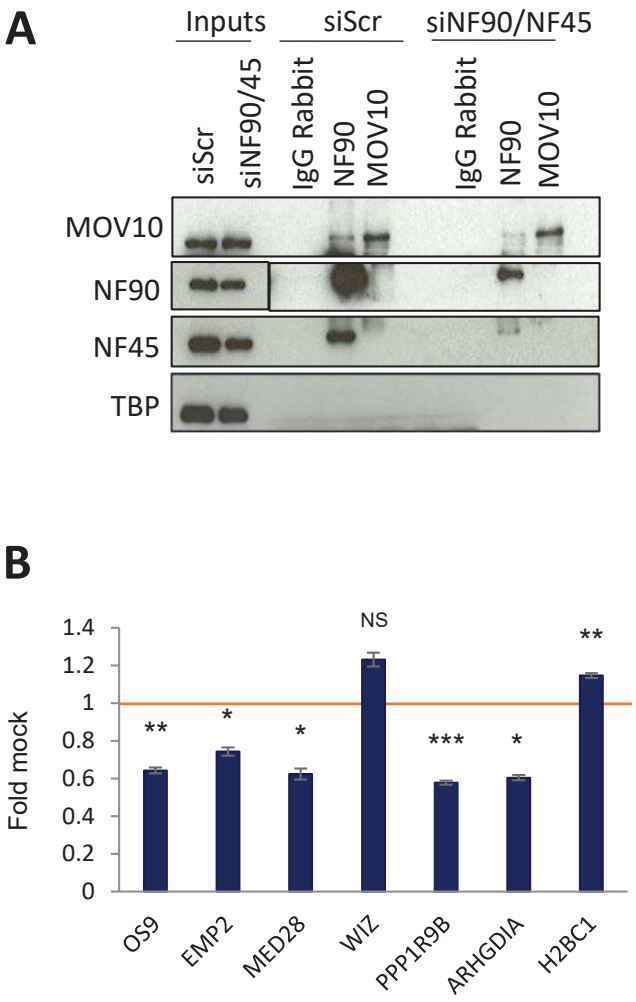


Figure S5. NF90 modulates MOV10 association with common target mRNAs. **(A)** Extracts of HEK293T cells transfected with siRNAs targeting NF90 and NF45 or a non-targeting control (Scr) and immunoprecipitates obtained using antibodies anti-NF90, anti-MOV10 or control IgG were analyzed by Western blot using the indicated antibodies. **(B)** Total RNA obtained from HEK293T transfected with siRNAs targeting NF90 and NF45 or a non-targeting control (Scr) was analyzed by RT-qPCR. NS indicates 'Not Significant'. Data represent Fold Mock (IgG) relative to the control samples (siScr), which was attributed a value of 1 (red line), obtained from 4 independent experiments *(*P < 0.05, **P < 0.01, ***P < 0.001*, independent Student's *t* test).

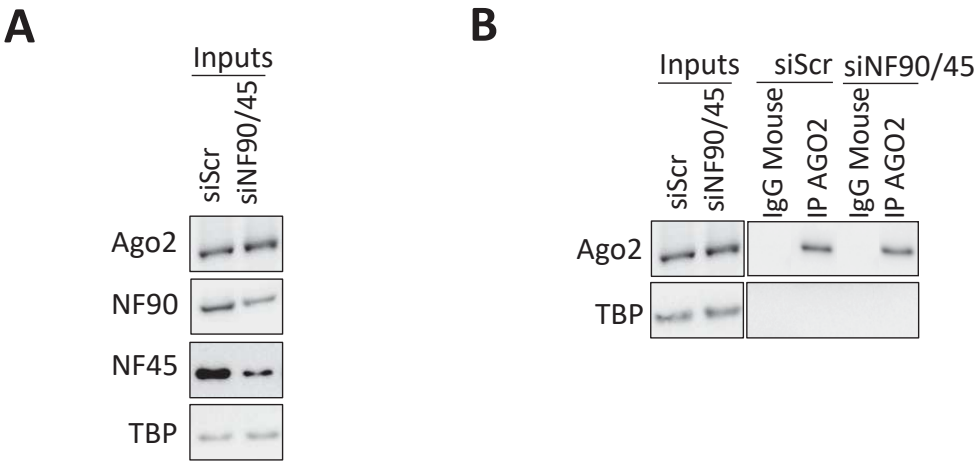


Figure S6. Downregulation of NF90/NF45 complex does not affect the expression of Ago2 **(A)** Extracts of HEK293T cells transfected with siRNAs targeting NF90/NF45 or a non-targeting control (Scr) were analyzed by Western blot using the indicated antibodies. **(B) I**mmunoprecipitates obtained using anti-Ago2 or control IgG antibodies from extracts of HEK293T cells transfected with siRNAs targeting NF90/NF45 or a non-targeting control (Scr) were analyzed by Western blot using the indicated antibodies.

Supplementary Table S1. Double stranded siRNAs used in this study.

siRNA	Sequence (5’ to 3’)
Scr	gcgcgcuuuguaggauucg(dTdT)
NF45	guggugauacucaagauucugccaa(dTdT)
NF90	ccaaggaacucuaucacaa(dTdT)
MOV10	guggaauuggaccgugucaagcuga(dTdT)

Supplementary Table S2. Primary antibodies used in this study.

Antibody	Reference	Supplier
NF90	A303-651A	Bethyl Laboratories
NF45	A303-147A	Bethyl Laboratories
MOV10	A301-571A	Bethyl Laboratories
Ago2 WB	MABE253, clone 119A	Sigma-Aldrich
Ago2 RIP	03-110	Sigma-Aldrich
TUBULIN	DM1A clone, T6199	Sigma-Aldrich
TBP	sc-421	SCBT

Supplementary Table S3. Primers used in this study.

Primer	Forward (5’ to 3’)	Reverse (5’ to 3’)
OS9	CTGCCCTTAGTGATGTTTGGA	ATTCTACCATTTCCTCAGCAACA
EMP2	GACCCATCCACCATTCATTC	CCACTGTACCCAGCCAGTTT
MED28	GAGCAGGCAGTAGGATGAGG	CCCCTCCACAGGGATATTTT
WIZ	TTGGCTGCTCCTTCTTGTTT	GCCTGTTAATCCCTCCTTCC
PPP1R9B	GCCCTGCACTGATTTCTCAT	TATTTGGCACCTGGAAGAGG
ARHGDIA	TGCCTCTGCCTTTTCTGTCT	GCACTTGGTCCCTTGTTTGT
H2BC1	GTGCTAAAGCAGGTCCATCC	GCATGTTTAGCCAGCTCTCC

Supplementary Table S4. Proteins associated with NF90 in the cytoplasm, detected by tandem affinity purification followed by tandem mass spectrometry. Proteins detected in the mock sample were excluded from the list.

Reference	Gene Symbol	Number of peptides	Reference	Gene Symbol	Number of peptides
Q12906_ILF3_HUMAN	ILF3	40	P62277_RS13_HUMAN	RPS13	6
Q9HCE1_MOV10_HUMAN	MOV10	25	P26599_PTBP1_HUMAN	PTBP1	6
Q12905_ILF2_HUMAN	ILF2	21	Q9Y2P8_RCL1_HUMAN	RCL1	6
Q6PKG0_LARP1_HUMAN	LARP1	21	P62280_RS11_HUMAN	RPS11	6
Q7L2E3_DHX30_HUMAN	DHX30	21	Q99459_CDC5L_HUMAN	CDC5L	6
O60506_HNRPQ_HUMAN	SYNCRIP	18	P43243_MATR3_HUMAN	MATR3	6
O43390_HNRPR_HUMAN	HNRNPR	17	Q92615_LAR4B_HUMAN	LARP4B	6
Q92900_RENT1_HUMAN	UPF1	17	Q00577_PURA_HUMAN	PURA	6
Q9Y6M1_IF2B2_HUMAN	IGF2BP2	15	Q12926_ELAV2_HUMAN	ELAVL2	6
P55265_DSRAD_HUMAN	ADAR	14	Q04637_IF4G1_HUMAN	EIF4G1	6
P78347_GTF2I_HUMAN	GTF2I	13	Q4G0J3_LARP7_HUMAN	LARP7	6
O75643_U520_HUMAN	SNRNP200	12	Q9UNX3_RL26L_HUMAN	RPL26L1	5
P39023_RL3_HUMAN	RPL3	11	P62906_RL10A_HUMAN	RPL10A	5
Q9NUL3_STAU2_HUMAN	STAU2	11	O75569_PRKRA_HUMAN	PRKRA	5
P05455_LA_HUMAN	SSB	10	Q96I25_SPF45_HUMAN	RBM17	5
P14866_HNRPL_HUMAN	HNRNPL	10	Q7L2H7_EIF3M_HUMAN	EIF3M	5
Q9NZB2_F120A_HUMAN	FAM120A	10	Q8N163_CCAR2_HUMAN	CCAR2	5
Q99873_ANM1_HUMAN	PRMT1	10	O00303_EIF3F_HUMAN	EIF3F	5
Q9NUD5_ZCHC3_HUMAN	ZCCHC3	10	Q96I24_FUBP3_HUMAN	FUBP3	5
O95793_STAU1_HUMAN	STAU1	10	P50454_SERPH_HUMAN	SERPINH1	5
P62750_RL23A_HUMAN	RPL23A	9	Q8IWX8_CHERP_HUMAN	CHERP	5
Q14157_UBP2L_HUMAN	UBAP2L	8	Q15029_U5S1_HUMAN	EFTUD2	5
P46777_RL5_HUMAN	RPL5	8	P60228_EIF3E_HUMAN	EIF3E	5
P48634_PRC2A_HUMAN	PRRC2A	8	P53621_COPA_HUMAN	COPA	5
Q9H2U1_DHX36_HUMAN	DHX36	8	Q9NVI7_ATD3A_HUMAN	ATAD3A	5
Q9BVP2_GNL3_HUMAN	GNL3	8	Q8NHW5_RLA0L_HUMAN	RPLP0P6	4
Q9NR30_DDX21_HUMAN	DDX21	8	P06748_NPM_HUMAN	NPM1	4
P07910_HNRPC_HUMAN	HNRNPC	7	P63173_RL38_HUMAN	RPL38	4
P18124_RL7_HUMAN	RPL7	7	P05387_RLA2_HUMAN	RPLP2	4
P62829_RL23_HUMAN	RPL23	7	Q15020_SART3_HUMAN	SART3	4
P32969_RL9_HUMAN	RPL9	7	Q9UMS4_PRP19_HUMAN	PRPF19	4
P52292_IMA1_HUMAN	KPNA2	7	P16989_YBOX3_HUMAN	YBX3	4
Q14692_BMS1_HUMAN	BMS1	7	P62888_RL30_HUMAN	RPL30	4
P26196_DDX6_HUMAN	DDX6	7	Q2NL82_TSR1_HUMAN	TSR1	4
Q13263_TIF1B_HUMAN	TRIM28	7	Q9NW13_RBM28_HUMAN	RBM28	4
P38919_IF4A3_HUMAN	EIF4A3	7	P46778_RL21_HUMAN	RPL21	4
Q6P2Q9_PRP8_HUMAN	PRPF8	7	P62316_SMD2_HUMAN	SNRPD2	4
Q9UQ35_SRRM2_HUMAN	SRRM2	7	Q92945_FUBP2_HUMAN	KHSRP	4
P26373_RL13_HUMAN	RPL13	6	Q13283_G3BP1_HUMAN	G3BP1	4
Q08945_SSRP1_HUMAN	SSRP1	6	Q7Z478_DHX29_HUMAN	DHX29	4

Reference	Gene Symbol	Number of peptides
Q09161_NCBP1_HUMAN	NCBP1	4
Q9UN81_LORF1_HUMAN	L1RE1	4
Q96N67_DOCK7_HUMAN	DOCK7	4
O00442_RTCA_HUMAN	RTCA	4
Q9UKM9_RALY_HUMAN	RALY	4
Q96CT7_CC124_HUMAN	CCDC124	4
P12532_KCRU_HUMAN	CKMT1A	4
O15234_CASC3_HUMAN	CASC3	4
Q9Y4E8_UBP15_HUMAN	USP15	4
Q6P158_DHX57_HUMAN	DHX57	4
P62917_RL8_HUMAN	RPL8	3
P27635_RL10_HUMAN	RPL10	3
P62857_RS28_HUMAN	RPS28	3
P05388_RLA0_HUMAN	RPLP0	3
Q96DH6_MSI2H_HUMAN	MSI2	3
Q9Y5B9_SP16H_HUMAN	SUPT16H	3
Q6Y7W6_PERQ2_HUMAN	GIGYF2	3
Q8NC51_PAIRB_HUMAN	SERBP1	3
P62910_RL32_HUMAN	RPL32	3
Q9BRJ6_CG050_HUMAN	C7orf50	3
Q7Z2W4_ZCCHV_HUMAN	ZC3HAV1	3
Q1KMD3_HNRL2_HUMAN	HNRNPUL2	3
Q92974_ARHG2_HUMAN	ARHGEF2	3
P52597_HNRPF_HUMAN	HNRNPF	3
Q6NZY4_ZCHC8_HUMAN	ZCCHC8	3
Q86V81_THOC4_HUMAN	ALYREF	3
Q9Y3B4_SF3B6_HUMAN	SF3B6	3
P04406_G3P_HUMAN	GAPDH	3
Q659C4_LAR1B_HUMAN	LARP1B	3
Q06787_FMR1_HUMAN	FMR1	3
Q71RC2_LARP4_HUMAN	LARP4	3
Q92947_GCDH_HUMAN	GCDH	3
Q07021_C1QBP_HUMAN	C1QBP	3
P56537_IF6_HUMAN	EIF6	3
Q8WXF1_PSPC1_HUMAN	PSPC1	3
P62304_RUXE_HUMAN	SNRPE	3
O76021_RL1D1_HUMAN	RSL1D1	3
Q9UK59_DBR1_HUMAN	DBR1	3
Q04837_SSBP_HUMAN	SSBP1	3
P04040_CATA_HUMAN	CAT	3
Q15477_SKIV2_HUMAN	SKIV2L	3
Q8N3C0_ASCC3_HUMAN	ASCC3	3

Reference	Gene Symbol	Number of peptides
O43175_SERA_HUMAN	PHGDH	3
Q96AE4_FUBP1_HUMAN	FUBP1	3
Q8IUX4_ABC3F_HUMAN	APOBEC3F	3
P35606_COPB2_HUMAN	COPB2	3
P42166_LAP2A_HUMAN	TMPO	3
Q96A72_MGN2_HUMAN	MAGOHB	3
Q14008_CKAP5_HUMAN	CKAP5	3
P09661_RU2A_HUMAN	SNRPA1	3
P42285_SK2L2_HUMAN	SKIV2L2	3
Q53EL6_PDCD4_HUMAN	PDCD4	3
Q9P1Y5_CAMP3_HUMAN	CAMSAP3	3
Q06830_PRDX1_HUMAN	PRDX1	3
P25311_ZA2G_HUMAN	AZGP1	3
Q6PGP7_TTC37_HUMAN	TTC37	3
P46779_RL28_HUMAN	RPL28	2
P84098_RL19_HUMAN	RPL19	2
P47914_RL29_HUMAN	RPL29	2
P18077_RL35A_HUMAN	RPL35A	2
P49207_RL34_HUMAN	RPL34	2
B4DY08_B4DY08_HUMAN	HNRNPC	2
Q9Y5S9_RBM8A_HUMAN	RBM8A	2
P60866_RS20_HUMAN	RPS20	2
O60256_KPRB_HUMAN	PRPSAP2	2
Q9BQ67_GRWD1_HUMAN	GRWD1	2
P35244_RFA3_HUMAN	RPA3	2
Q9Y3Y2_CHTOP_HUMAN	CHTOP	2
P84103_SRSF3_HUMAN	SRSF3	2
P42766_RL35_HUMAN	RPL35	2
Q9Y3U8_RL36_HUMAN	RPL36	2
Q02543_RL18A_HUMAN	RPL18A	2
Q15427_SF3B4_HUMAN	SF3B4	2
P62995_TRA2B_HUMAN	TRA2B	2
O15226_NKRF_HUMAN	NKRF	2
O75531_BAF_HUMAN	BANF1	2
P62841_RS15_HUMAN	RPS15	2
Q16875_F263_HUMAN	PFKFB3	2
Q92522_H1X_HUMAN	H1FX	2
Q86U42_PABP2_HUMAN	PABPN1	2
O14654_IRS4_HUMAN	IRS4	2
Q9H6S0_YTDC2_HUMAN	YTHDC2	2
O00267_SPT5H_HUMAN	SUPT5H	2
Q9Y5A9_YTHD2_HUMAN	YTHDF2	2

Reference	Gene Symbol	Number of peptides
P63220_RS21_HUMAN	RPS21	2
Q15459_SF3A1_HUMAN	SF3A1	2
Q9UL40_ZN346_HUMAN	ZNF346	2
Q13148_TADBP_HUMAN	TARDBP	2
Q8N954_GPT11_HUMAN	GPATCH11	2
Q9UNQ2_DIM1_HUMAN	DIMT1	2
P14618_KPYM_HUMAN	PKM	2
Q99848_EBP2_HUMAN	EBNA1BP2	2
Q13595_TRA2A_HUMAN	TRA2A	2
P16403_H12_HUMAN	HIST1H1C	2
P82650_RT22_HUMAN	MRPS22	2
O75175_CNOT3_HUMAN	CNOT3	2
Q13573_SNW1_HUMAN	SNW1	2
P78344_IF4G2_HUMAN	EIF4G2	2
Q96KR1_ZFR_HUMAN	ZFR	2
Q96PU8_QKI_HUMAN	QKI	2
P42696_RBM34_HUMAN	RBM34	2
Q96QR8_PURB_HUMAN	PURB	2
Q9NX24_NHP2_HUMAN	NHP2	2
Q9NX05_F120C_HUMAN	FAM120C	2
Q9ULX6_AKP8L_HUMAN	AKAP8L	2
Q99575_POP1_HUMAN	POP1	2
Q9BYJ9_YTHD1_HUMAN	YTHDF1	2
P16383_GCFC2_HUMAN	GCFC2	2
Q8IX01_SUGP2_HUMAN	SUGP2	2
Q9UL18_AGO1_HUMAN	AGO1	2
Q8IZH2_XRN1_HUMAN	XRN1	2
Q9Y265_RUVB1_HUMAN	RUVBL1	2
P33993_MCM7_HUMAN	MCM7	2
IGH1M_MOUSE	Ighg1	2
Q9UKV3_ACINU_HUMAN	ACIN1	2
O75822_EIF3J_HUMAN	EIF3J	2
Q9H307_PININ_HUMAN	PNN	2
P42357_HUTH_HUMAN	HAL	2
O43242_PSMD3_HUMAN	PSMD3	2
P13639_EF2_HUMAN	EEF2	2
P35251_RFC1_HUMAN	RFC1	2
P08621_RU17_HUMAN	SNRNP70	2
P0CB38_PAB4L_HUMAN	PABPC4L	1
Q02539_H11_HUMAN	HIST1H1A	1
P42677_RS27_HUMAN	RPS27	1
Q5JNZ5_RS26L_HUMAN	RPS26P11	1
Q16695_H31T_HUMAN	HIST3H3	1
P13995_MTDC_HUMAN	MTHFD2	1
P05386_RLA1_HUMAN	RPLP1	1
Q01804_OTUD4_HUMAN	OTUD4	1
Q6EMK4_VASN_HUMAN	VASN	1
O43709_WBS22_HUMAN	WBSCR22	1
O43347_MSI1H_HUMAN	MSI1	1
Q15365_PCBP1_HUMAN	PCBP1	1
Q15287_RNPS1_HUMAN	RNPS1	1
P61513_RL37A_HUMAN	RPL37A	1
Q7Z2T5_TRM1L_HUMAN	TRMT1L	1
P09012_SNRPA_HUMAN	SNRPA	1
P29558_RBMS1_HUMAN	RBMS1	1
Q9BRZ2_TRI56_HUMAN	TRIM56	1
P19525_E2AK2_HUMAN	EIF2AK2	1
O75934_SPF27_HUMAN	BCAS2	1
P55209_NP1L1_HUMAN	NAP1L1	1
P62314_SMD1_HUMAN	SNRPD1	1
Q3MHD2_LSM12_HUMAN	LSM12	1
Q6ZN17_LN28B_HUMAN	LIN28B	1
Q58A45_PAN3_HUMAN	PAN3	1
Q9ULR0_ISY1_HUMAN	ISY1	1
Q8N4Q0_ZADH2_HUMAN	ZADH2	1
Q14498_RBM39_HUMAN	RBM39	1
Q07666_KHDR1_HUMAN	KHDRBS1	1
P24534_EF1B_HUMAN	EEF1B2	1
Q9Y383_LC7L2_HUMAN	LUC7L2	1
P50991_TCPD_HUMAN	CCT4	1
P08238_HS90B_HUMAN	HSP90AB1	1
O75494_SRS10_HUMAN	SRSF10	1
P60709_ACTB_HUMAN	ACTB	1
Q15366_PCBP2_HUMAN	PCBP2	1
P04844_RPN2_HUMAN	RPN2	1
P25705_ATPA_HUMAN	ATP5A1	1
P60891_PRPS1_HUMAN	PRPS1	1
Q9BYD3_RM04_HUMAN	MRPL4	1
O43167_ZBT24_HUMAN	ZBTB24	1
Q9UHB9_SRP68_HUMAN	SRP68	1
K7ER90_K7ER90_HUMAN	EIF3G	1
Q9BUF5_TBB6_HUMAN	TUBB6	1
P11908_PRPS2_HUMAN	PRPS2	1
Q9NQ92_COPRS_HUMAN	COPRS	1

Reference	Gene Symbol	Number of peptides
P32119_PRDX2_HUMAN	PRDX2	1
Q02809_PLOD1_HUMAN	PLOD1	1
Q5BKZ1_ZN326_HUMAN	ZNF326	1
Q15434_RBMS2_HUMAN	RBMS2	1
Q14558_KPRA_HUMAN	PRPSAP1	1
O14818_PSA7_HUMAN	PSMA7	1
Q58FF8_H90B2_HUMAN	HSP90AB2P	1
O43734_CIKS_HUMAN	TRAF3IP2	1
Q9BWU0_NADAP_HUMAN	SLC4A1AP	1
Q96EC8_YIPF6_HUMAN	YIPF6	1
P20618_PSB1_HUMAN	PSMB1	1
P98179_RBM3_HUMAN	RBM3	1
Q13243_SRSF5_HUMAN	SRSF5	1
Q8N5C8_TAB3_HUMAN	TAB3	1
Q5VYS8_TUT7_HUMAN	ZCCHC6	1
J3QR62_J3QR62_HUMAN	DDX5	1
P05141_ADT2_HUMAN	SLC25A5	1
Q07955_SRSF1_HUMAN	SRSF1	1
Q14671_PUM1_HUMAN	PUM1	1
Q9NZN8_CNOT2_HUMAN	CNOT2	1
B7Z645_B7Z645_HUMAN	SYNCRIP	1
Q15084_PDIA6_HUMAN	PDIA6	1
P62273_RS29_HUMAN	RPS29	1
Q9NP73_ALG13_HUMAN	ALG13	1
Q9NY12_GAR1_HUMAN	GAR1	1
O00567_NOP56_HUMAN	NOP56	1
Q92600_RCD1_HUMAN	RQCD1	1
Q9BTZ2_DHRS4_HUMAN	DHRS4	1
P49327_FAS_HUMAN	FASN	1
Q14694_UBP10_HUMAN	USP10	1
Q16629_SRSF7_HUMAN	SRSF7	1
C9JUF0_C9JUF0_HUMAN	EIF4A2	1
Q9BZI7_REN3B_HUMAN	UPF3B	1
Q9HCS7_SYF1_HUMAN	XAB2	1
Q9H7E9_CH033_HUMAN	C8orf33	1
Q9NWU5_RM22_HUMAN	MRPL22	1
Q9Y3D9_RT23_HUMAN	MRPS23	1
P26368_U2AF2_HUMAN	U2AF2	1
Q96HS1_PGAM5_HUMAN	PGAM5	1
O75940_SPF30_HUMAN	SMNDC1	1
P42704_LPPRC_HUMAN	LRPPRC	1
O43172_PRP4_HUMAN	PRPF4	1
C9JQR9_C9JQR9_HUMAN	RPSAP58	1
Q96L21_RL10L_HUMAN	RPL10L	1
Q08378_GOGA3_HUMAN	GOLGA3	1
Q3KQU3_MA7D1_HUMAN	MAP7D1	1
E9PRG8_CK098_HUMAN	C11orf98	1
O14980_XPO1_HUMAN	XPO1	1
Q5LJB1_Q5LJB1_HUMAN	UCHL5	1
P49736_MCM2_HUMAN	MCM2	1
Q8N8E3_CE112_HUMAN	CEP112	1
O75152_ZC11A_HUMAN	ZC3H11A	1
Q8WXX5_DNJC9_HUMAN	DNAJC9	1
Q9NVU7_SDA1_HUMAN	SDAD1	1
O00505_IMA4_HUMAN	KPNA3	1
O00458_IFRD1_HUMAN	IFRD1	1
P51659_DHB4_HUMAN	HSD17B4	1
Q96PX6_CC85A_HUMAN	CCDC85A	1
P08579_RU2B_HUMAN	SNRPB2	1
Q5QJ74_TBCEL_HUMAN	TBCEL	1
Q9UKV8_AGO2_HUMAN	AGO2	1
O15294_OGT1_HUMAN	OGT	1
P62807_H2B1C_HUMAN	HIST1H2BC	1
Q9BQ39_DDX50_HUMAN	DDX50	1
O75962_TRIO_HUMAN	TRIO	1
P25789_PSA4_HUMAN	PSMA4	1
O95721_SNP29_HUMAN	SNAP29	1
Q9NRW3_ABC3C_HUMAN	APOBEC3C	1
O60437_PEPL_HUMAN	PPL	1
P30872_SSR1_HUMAN	SSTR1	1
Q5TZA2_CROCC_HUMAN	CROCC	1

Chapter 4

Discussion

NF90 is a double-stranded RNA binding protein that was shown to have a role in a multitude of cellular mechanisms, such as transcription, viral infection, translation, mRNA stability and miRNA biogenesis (Castella et al., 2015; X. Li et al., 2017; Sakamoto et al., 2009). NF90 structure displays two dsRBDs that have been shown to cooperatively contribute to the binding of the same RNA molecule. However, its binding mode is not yet definitively characterized. Data suggest that NF90 might be able to recognize specific RNA secondary structures, such as minihelix-like structures, and these features would be sufficient for efficient NF90 binding (Gwizdek et al., 2004). However, recent findings propose that a particular sequence preference might also influence its binding to RNAs (Jayachandran et al., 2016). Although NF90 alone is able to bind dsRNAs, its binding activity is strongly affected by the heterodimerization with its protein partner, NF45, which leads to thermodynamic stabilization of the complex and enhanced affinity for RNA substrates (Schmidt et al., 2017).

NF90 is a very abundant and ubiquitous protein that, by its binding to RNAs, carries out a plethora of different effects in both the nucleus and the cytoplasm. In fact, NF90 has been shown to shuttle between these two cellular compartments according to its phosphorylation status or as a consequence of several stimuli, such as viral infection and cancer-induced hypoxia (Harashima et al., 2010; Patiño et al., 2015; W. Zhang et al., 2018). Therefore, the level of nuclear versus cytoplasmic NF90 is extremely important and its export is strictly regulated.

During my PhD I investigated the role of NF90 in miRNA biogenesis pathway in the nucleus and its implication in the RISC-mediated silencing in the cytoplasm.

In the nucleus, NF90 was previously shown to modulate the processing of a subset of human miRNA precursors (Barbier et al., 2018; Sakamoto et al., 2009). However, its global impact on the biogenesis of miRNAs was not clearly defined.
During my thesis, we exploited genome-wide approaches in order to understand how widespread the effect of NF90 is on the biogenesis of human miRNAs in HepG2 cell line (Grasso et al., 2020).
Our data indicated that NF90 is able to directly bind and modulate the processing of a specific subset of human miRNA precursors that are weakly bound by the Microprocessor. This suggests that NF90 and the Microprocessor might be in competition for the binding of pri-miRNAs. Analyzing the features of the NF90-bound and modulated pri-miRNAs, we found that they display significant longer duplexes with fewer and shorter bulges, compared to all human pri-miRNAs. Since these features might influence RNA stability, we compared the free energy of NF90-associated pri-miRNAs to all human pri-miRNAs, concluding that they are, in fact, more stable than other human pri-miRNAs. Moreover, we were able to diminish or increase NF90 association to selected pri-miRNAs by introducing destabilizing or stabilizing mutations, respectively. These findings are consistent with previous reports suggesting that NF90 is able to bind the adenovirus-expressed VA1 RNA, in a manner that is largely dependent on a minihelix-like structure (Gwizdek et al., 2004).
Finally, transcriptomic analysis showed that NF90 association with pri-miRNAs may modulate the expression of their host genes, as described previously (Barbier et al., 2018). For instance, the expression of TIAM2, hosting pri-miR-1273C, is down-regulated upon loss of NF90. TIAM2 is a well-known oncogene and metastasis factor in hepatocarcinoma (HCC) and its overexpression was shown to promote proliferation and invasion in liver cancer (J.-S. Chen et al., 2012).
Interestingly, NF90 is upregulated in HCC. Therefore, it would be interesting to determine whether NF90-dependent modulation of TIAM2 might contribute to HCC pathogenesis.

On the same lines, some of NF90 bound and modulated pri-miRNAs, such as miR-34a, are tumour-suppressor miRNAs, which have been shown to be downregulated during cancer development and progression (Long et al., 2018). NF90 binding to these pri-miRNAs would protect them from Microprocessor cleavage, leading to cancer progression. In fact, similarly to mir-34a, we identified other miRNAs with known tumor suppressor activities, such as miR-16, miR-128 and miR-145, whose biogenesis is repressed by NF90 (Grasso et al., 2020; Lu et al., 2005). Interestingly, in bladder cancer, miR-145 was shown to be part of a dysregulated signalling axis that is responsible for drug resistance, involving lncRNA-Low Expression in Tumor (lncRNA-LET) and NF90. In particular, low levels of lncRNA-LET, a known repressor of NF90 expression, leads to high levels of NF90 which, in turn, can bind the tumor suppressor miR-145 suppressing its biogenesis (Zhuang et al., 2017).

Since low expression of lncRNA-LET is linked to metastasis and poor prognosis in many different tumors (F. Liu et al., 2016), we could speculate that the miRNAs found upregulated after loss of NF90 in HepG2 cell line, such as miR-1273c, might also be part of this dysregulated signalling axis, which might therefore have a more widespread effect than currently thought.

Contrarily, in ovarian cancer, NF90 is a good prognosis factor that, binding to pri-miR-3173, embedded in DICER pre-mRNA, leads to DICER overexpression, which has tumour-suppressor activities (Barbier et al., 2018). In fact, it seems that the resulting effect of NF90 on cancer development and progression might be dictated not only by the effect of the miRNAs it modulates, but also by the role of the host-gene that contains the pri-miRNA. Therefore, the outcome of NF90 in cancer depends on a complex mechanism that needs to be further investigated.

Given the widespread implication of NF90 in cancer, through its ability to bind pri-miRNAs, it would be tempting to envision the application of this mechanism for the development of therapeutics designed in order to inhibit or promote NF90 binding to selected pri-miRNAs. This would potentially have an effect not only on the miRNAs and their mRNA targets, but also on the mRNA hosting the

pri-miRNAs, which can be tumour suppressors, such as Dicer, or oncogenes, such as TIAM2.

Although our results were not conclusive on the mechanism by which the loss of NF90 could reduce the abundance of the host gene mRNA, several hypotheses could be envisioned. For instance, it is possible that the cleavage of the intronic pri-miRNA, upon loss of NF90, produces a break in the intron that might potentially lead to rapid degradation of the mRNA by exonucleases or to a deleterious splicing defect if the two flanking exons were not previously tethered together. However, it could also be possible that, upon loss of NF90, pri-miRNA processing by the Microprocessor might directly interfere with the spliceosome machinery, leading to splicing defects and, ultimately, downregulation of the host-gene protein. Needless to say, more evidence need to be gathered to elucidate this mechanism.

On a different note, our GO analyses, on the validated targets of NF90-bound and modulated miRNAs, revealed a significant enrichment for pathways implicated in infection by viruses such as Epstein Barr Virus (EBV), hepatitis B virus (HBV) and human T lymphoma virus type 1 (HTLV1) and in viral carcinogenesis. Interestingly, as previously mentioned, viral infection was shown to induce the translocation of NF90 from the nucleus to the cytoplasm (X. Li et al., 2017). As a consequence, cytoplasmic NF90 was found to bind viral RNAs, cellular RNAs and circRNAs, modulating their abundance and playing a role in the response to viral infection. It's important to note that the role of NF90 in viral infection response is controversial. In fact, the consequence of NF90 binding to viral RNA or DNA can vary, supporting or inhibiting viral replication and viral genome expression, depending on the type of virus (Patiño et al., 2015). It is possible that the contribution of NF90 to antiviral immunity might be carried out through different mechanisms and pathways, depending on the cell type and the infecting virus. Therefore, more evidence is needed to fully elucidate the role of NF90 in viral infection.

As previously mentioned, NF90 is a highly abundant and ubiquitously expressed protein. As expected, NF90-associated pri-miRNAs are poorly processed in most cells and, therefore, lowly expressed, if expressed at all. Moreover, NF90-bound pri-miRNAs are often poorly conserved between species and most of them are even primate-specific. Taken together, this data, combined with our results showing a minihelix-like structure preference for NF90 binding, would suggest a role for NF90 in protecting emerging and unpredictable miRNAs whose expression might lead to deleterious side effects. In fact, it is known that the human genome encodes for millions of hairpin structures, especially in transcripts (Bentwich et al., 2005). According to hypotheses of miRNA evolution, these structures might potentially become functional miRNAs upon the acquisition of a series of mismatches (Felippes et al., 2008). However, prior to this maturation step, hairpins might still be recognized and cleaved by the Microprocessor leading to spurious products with unforeseeable sides effects (B. Kim et al., 2017). To avoid this, proteins, such as NF90, might be used to avoid the cleavage of immature hairpins to produce unintended miRNA products. In agreement with this hypothesis, more conserved and expressed miRNAs, which have a higher number of mismatches, are often poorly bound by NF90.

It is possible that many of the pri-miRNAs that we found to be bound and modulated by NF90 are recent but already functional miRNAs whose expression needs to be restricted to specific conditions, such as viral infection or hypoxia, or in different cell types. In this case, modulation of the abundance or the localization of NF90 might contribute to an additional level of regulation for the expression of these miRNAs.

Other proteins have been shown to have a similar function, binding to pri-miRNAs to prevent microprocessing. For instance, pri-miR-7-1, enriched in brain and pancreatic tissue, has a conserved terminal loop that was shown to be bound by MSI2 and HuR proteins (N. R. Choudhury et al., 2013). Interestingly, their binding to pri-miR-7-1 terminal loop in non-neuronal cells leads to a more stable and rigid stem and, consequently, to the inhibition of microprocessor activity (N. R. Choudhury et al., 2013). We showed that NF90 is able to bind pri-miR-7-1 in HepG2 cells, repressing its cleavage. Therefore, it would be interesting to

determine whether MSI2/HuR and NF90 can bind the same pri-miRNAs in the same cell type, cooperating to control the processing of pri-miRNAs, or they might exert similar functions in different cell types.

In the cytoplasm, NF90 was already shown to have a role in translation regulation and mRNA stability (Castella et al., 2015). It was recently suggested that the role of NF90 in controlling translation might involve miRNAs and miRNA recognition sites on target mRNAs. In fact, NF90 is able to diminish the abundance and the translation of a B-cell specific mRNA, BAFF, and this function seems to be dependent on miR-15a and its MRE within BAFF 3′ UTR (Idda et al., 2018). These recent finding would suggest a role for NF90 in miRNA-guided RISC-mediated silencing. However, the contribution of NF90 to this pathway was unclear. During the second part of my PhD, we investigated the effect of NF90 in the RISC-mediated silencing.

In order to understand how NF90 is involved in translation regulation and mRNA stability, we identified the interactome of cytoplasmic NF90 by performing tandem affinity purification followed by mass spectrometry, in HEK293T cells expressing Flag.HA tagged-NF90. Our data indicated that cytoplasmic NF90 interacts with proteins involved in translation regulation, viral replication and mRNA stability and degradation. Among NF90 interactants in the cytoplasm, the most abundant was MOV10, a 5′-to-3′ helicase involved in RISC-mediated silencing (Meister et al., 2005). MOV10 is able to unwind structured RNA to reveal MREs and facilitate AGO2 binding (Gregersen et al., 2014). Interestingly, the effector of RISC-mediated silencing, AGO2, was also found in the interactome of NF90. Furthermore, we determined that the association of NF90 to MOV10 and AGO2 occurs through RNA. Using glycerol gradient sedimentation of NF90 immune complexes, we found that NF90, MOV10, AGO2 and other proteins involved in RISC-mediated silencing can be found in the same fractions, suggesting that they can exist in the same cytosolic complex.

In addition, we took advantage of published eCLIP and iCLIP data of NF90 and MOV10 (Kenny et al., 2014; Nussbacher & Yeo, 2018), respectively, and found that NF90 and MOV10 can be associated with the 3' UTRs of the same target mRNAs. Interestingly, NF90 binding occurs on a high proportion of MOV10-bound mRNAs, around 40%. Although the techniques and the cell lines used in the two CLIPs are different, this high overlap could suggest that MOV10-bound mRNAs might often be bound by NF90.

Since NF90 was shown to increase mRNA stability, while MOV10 enhances RISC-mediated silencing, which consequently leads to mRNA degradation, we wondered if they could compete for the binding of selected target mRNAs. Indeed, upon downregulation of MOV10, we could detect an increase in the association of NF90 to the target mRNAs. Likewise, loss of the heterodimer NF90/NF45 increased MOV10 binding while decreasing the abundance of the target mRNAs. These findings suggest that NF90 might interfere with the binding of MOV10 on target mRNAs and *vice versa*, which could subsequently modify the stability and abundance of the mRNAs.

This interference is unlikely to occur through direct competition between NF90 and MOV10. In fact, the binding preferences of NF90 and MOV10 differ significantly (Gregersen et al., 2014; Gwizdek et al., 2004). While NF90 can bind stable hairpin structures, MOV10 seems to bind ssRNA upstream of a structured region. However, it is possible that the binding of one factor might modify the RNA structure in the proximity of the binding site, disfavoring the association of the other. Another possibility is that NF90 or MOV10 might recruit additional factors which ultimately interfere with the binding of the other.

As previously mentioned in this chapter, MOV10 is a helicase known to resolve structured mRNAs in order to reveal obscured MREs and facilitate the binding of AGO2, favoring RISC-mediated silencing.

Therefore, in order to understand if the binding of NF90 to the target mRNAs ultimately influences AGO2 binding, we performed RIP of AGO2 following loss of the heterodimer NF90/NF45. Interestingly, we found that loss of NF90/NF45

significantly enhances the binding of AGO2 to the target mRNAs while decreasing their abundance. These results suggest that NF90 might enhance the stability of a subset of mRNAs by modulating their association with AGO2.

Consistent with this hypothesis, NF90 was recently identified as a subunit of P-bodies suggesting that it might be able to control the stability and translation of mRNAs by modulating the association of RISC (Hubstenberger et al., 2017). However, the mRNAs identified in our study seem to be depleted in P-bodies extracted from HEK293, according to a recent study (Hubstenberger et al., 2017). It is therefore possible that, although NF90 was co-purified with other RBPs that localize to P-bodies, the binding of NF90 to the identified target mRNAs occurs outside of P-bodies. In fact, it was recently shown that the localization of mRNAs in P-bodies and their stability is dependent on their GC content (Courel et al., 2019). In particular, AU-rich mRNAs localize in P-bodies where they are stored and their translation is inhibited while GC rich mRNAs are depleted in P-bodies and they can be targetted for mRNA decay. This is also consistent with our finding that loss of NF90 negatively affected the abundance of the identified target mRNAs, which are mainly GC-rich. We could speculate that NF90 might be a factor involved in the determination of translational repression (in P-bodies) over decay and *vice versa*, by binding to mRNAs in P-bodies or interacting with proteins involved in mRNA degradation. Needless to say, more evidence is needed to shed light on the role of NF90 in P-bodies, RISC-mediated silencing and decay.

Since NF90 is known to translocate to the cytoplasm during viral infection or cancer-induced hypoxia (X. Li et al., 2017; W. Zhang et al., 2018), it would be tempting to speculate that, under these conditions, NF90 might interfere with MOV10 and AGO2 binding, leading to inhibition of RISC-mediated silencing for some specific target mRNAs. Therefore, understanding exactly how NF90 is implicated in RISC-mediated silencing might potentially clarify its role in mRNA translation and degradation during viral infection or cancer progression.

Ultimately, the roles of NF90 in the nucleus and in the cytoplasm might be strictly connected in order to achieve a common and integrated response to external stimuli. For instance, it has been shown that during viral infection, NF90 in the

nucleus suppresses the biogenesis of several miRNAs, such as miR-4753 and miR-3145, that inhibit H5N1 and H3N2 transcription and replication (Grasso et al., 2020; Zheng et al., 2020). On the other hand, we found that interactors of cytoplasmic NF90 are also involved in viral transcription and viral translation. Moreover, NF90 itself was found to enhance HIV viral infection in the cytoplasm (Y. Li & Belshan, 2016). Therefore, we could speculate that during viral infection the effects of NF90 in the nucleus and in the cytoplasm might be coordinated to either promote or inhibit viral replication, depending on the infecting virus.

In conclusion, NF90 is a polyvalent factor involved in a number of cellular pathways which may be interconnected to generate a concerted response under specific physiological conditions. However, its important role in viral response and cancer progression remains largely unknown and warrants further investigation.

Appendix A

NF90 modulates processing of a subset of human pri-miRNAs

6874–6888 *Nucleic Acids Research, 2020, Vol. 48, No. 12* *Published online 19 May 2020*
doi: 10.1093/nar/gkaa386

NF90 modulates processing of a subset of human pri-miRNAs

Giuseppa Grasso[1], Takuma Higuchi[2], Victor Mac[1], Jérôme Barbier[1], Marion Helsmoortel[1], Claudio Lorenzi[3], Gabriel Sanchez[1], Maxime Bello[1], William Ritchie[3], Shuji Sakamoto[2] and Rosemary Kiernan[1,*]

[1]UMR9002 CNRS-UM, Institut de Génétique Humaine-Université de Montpellier, Gene Regulation lab, Montpellier 34396, France, [2]Laboratory of Molecular Biology, Science Research Centre, Kochi Medical School, Kochi University, Kochi 783-8505, Japan and [3]UMR9002 CNRS-UM, Institut de Génétique Humaine-Université de Montpellier, Artificial Intelligence and Gene Regulation lab, Montpellier 34396, France

Received July 16, 2019; Revised April 24, 2020; Editorial Decision April 30, 2020; Accepted May 01, 2020

ABSTRACT

MicroRNAs (miRNAs) are predicted to regulate the expression of >60% of mammalian genes and play fundamental roles in most biological processes. Deregulation of miRNA expression is a hallmark of most cancers and further investigation of mechanisms controlling miRNA biogenesis is needed. The double stranded RNA-binding protein, NF90 has been shown to act as a competitor of Microprocessor for a limited number of primary miRNAs (pri-miRNAs). Here, we show that NF90 has a more widespread effect on pri-miRNA biogenesis than previously thought. Genome-wide approaches revealed that NF90 is associated with the stem region of 38 pri-miRNAs, in a manner that is largely exclusive of Microprocessor. Following loss of NF90, 22 NF90-bound pri-miRNAs showed increased abundance of mature miRNA products. NF90-targeted pri-miRNAs are highly stable, having a lower free energy and fewer mismatches compared to all pri-miRNAs. Mutations leading to less stable structures reduced NF90 binding while increasing pri-miRNA stability led to acquisition of NF90 association, as determined by RNA electrophoretic mobility shift assay (EMSA). NF90-bound and downregulated pri-miRNAs are embedded in introns of host genes and expression of several host genes is concomitantly reduced. These data suggest that NF90 controls the processing of a subset of highly stable, intronic miRNAs.

INTRODUCTION

MicroRNAs (miRNAs) are short non-coding RNAs that negatively regulate the expression of a large proportion of cellular mRNAs, thus affecting a multitude of cellular and developmental pathways (1,2). The canonical miRNA biogenesis pathway involves two sequential processing events catalysed by RNase III enzymes. In the nucleus, the microprocessor complex, comprising the RNase III enzyme Drosha, the double-stranded RNA-binding protein, DGCR8 and additional proteins carries out the first processing event, which results in the production of precursor miRNAs (pre-miRNAs) (3,4). These are exported to the cytoplasm, where a second processing event is carried out by another RNase III enzyme, DICER, leading to the production of miRNA duplexes. The duplexes are loaded into the RISC complex and the release of the 'passenger' strands leads to the formation of mature miRNAs and mature RISC complexes (5).

Due to the central role of miRNAs in the control of gene expression, their levels must be tightly controlled. Indeed, deregulation of miRNA expression is associated with aberrant gene expression and leads to human disease (6–9). Consequently, miRNA biogenesis is tightly regulated at multiple steps, both transcriptional and post-transcriptional. Increasing evidence suggests that RNA binding proteins (RBPs) act as post-transcriptional regulators of miRNA processing. Many RBPs modulate the processing efficiency of Microprocessor, either positively or negatively, by binding to regions of the pri-miRNA. A number of RBPs have been shown to bind the terminal loop, which can either facilitate or inhibit cropping by Microprocessor. For example, LIN28B binds the terminal loop of pri-let-7, which prevents its processing by Microprocessor (10). Binding of hnRNP A1 to the terminal loop has been shown to exert either positive or negative effects on Microprocessor activity, depending on the pri-miRNA target. It promotes cropping of pri-miR-18A while it inhibits processing of pri-let-7. KSRP is another terminal loop-binding RBP that facilitates Microprocessor cleavage of several pri-miRNA targets, including pri-let-7 where it acts as a competitor of hn-

*To whom correspondence should be addressed. Tel: +33 4 34359939; Fax: +33 4 34359901; Email: Rosemary.Kiernan@igh.cnrs.fr

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

RNP A1 (11,12). Several other RBPs, including SMAD, TPD-43, SRSF1 and RBFOX, have been shown to bind pri-miRNA terminal loops to influence Microprocessor activity (see (13) for review). In most cases, they have been shown to bind specific pri-miRNAs, such as pri-let-7, or a limited subset of pri-miRNAs. To date, only NF90/NF45 heterodimer and ADAR1,2 have been shown to bind the double stranded stem region of pri-miRNAs (14–17). Both factors negatively affect Microprocessor activity. Indeed, NF90 has been shown to bind double stranded RNA in a mode similar to that of ADAR2 (18). Like terminal loop binding RBPs, binding of NF90/NF45 or ADAR1,2 has thus far been demonstrated for a very limited number of pri-miRNAs. NF90 has been shown to associate with pri-miR-7-1, pri-let-7A and pri-miR-3173 in human cells (14,15,19).

We have previously shown that NF90 associates with pri-miR-3173, which is located in the first intron of Dicer pre-mRNA (19). Binding of NF90 prevented cropping of pri-miR-3173 by Microprocessor and promoted splicing of the intron, thereby facilitating expression of DICER. By modulating DICER expression, NF90 was found to be an independent prognostic marker of ovarian carcinoma progression (19). Levels of NF90 are known to be elevated in hepatocellular carcinoma (HCC) and the effect of NF90 on processing of pri-miR-7-1 contributes to cellular proliferation in HCC models (14,20). Here, we have used genome-wide approaches to identify pri-miRNAs that are associated with and modulated by NF90 in HepG2 model of HCC. We identified 38 pri-miRNAs that are associated with NF90, in a manner that is for the most part exclusive of Microprocessor. Of these, 22 showed increased abundance of mature miRNAs products upon loss of NF90. NF90-targeted pri-miRNAs appear to be highly stable, having a lower free energy and fewer mismatches compared to all pri-miRNAs. Destabilization of the structures by mutation reduced NF90 association as determined by RNA EMSA. Of the 22 NF90-modulated pri-miRNAs, 20 are embedded exclusively in introns of host genes. Transcriptomic analysis revealed that the expression of the host gene is concomitantly downregulated for several, including an oncogene implicated in metastasis of hepatocellular carcinoma, TIAM2. These data suggest that NF90 controls the processing of a subset of intronic miRNAs, which in some cases affects the expression of the host gene.

MATERIALS AND METHODS

Cell culture

Human HepG2 cell line was grown in Dulbecco's modified Eagle's medium—high glucose (Sigma-Aldrich®, D6429) supplemented with 10% fetal bovine serum (PAN Biotech, 8500-P131704), 1% penicillin–streptomicin (v/v) (Sigma Aldrich®, P4333) and 1% L-glutamine (v/v) (Sigma Aldrich®, G7513). Human HEK-293T cells were grown in Dulbecco's modified Eagle's high glucose medium with HEPES (Sigma-Aldrich®, D6171) supplemented with 10% fetal bovine serum, 1% penicillin–streptomicin and 1% L-glutamine. Cells were cultured at 37°C in a humidified atmosphere containing 5% CO_2. To perform small RNA-seq and RNA-seq, HepG2 were seeded at 1.5×10^6 cells in six-well plates the day of siRNA transfection while HEK-293T were seeded at 6×10^5 cells in six-well plates.

To perform RNA Immunoprecipitation, HepG2 were seeded at 8×10^6 cells in 100 mM culture dishes the day of siRNA transfection.

Transfection of small interfering RNAs

Double-stranded RNA oligonucleotides used for RNAi were purchased from Eurofins MWG Operon or Integrated DNA Technologies. Sequences of small interfering RNAs (siRNAs) used in this study have been described previously (19) and are shown in Supplementary Table S1.

HepG2 or HEK-293T cells were transfected with siRNA (30 nM final concentration) using INTERFERin® siRNA transfection reagent (Polyplus Transfection) according to the manufacturer's instructions. To perform small RNA-seq and RNA-seq, two rounds of transfection were performed. The first transfection was carried out the day of seeding; on the fourth day cells were passaged and a second round of transfection was performed. Cells were collected for RNA extraction or protein purification ∼65 h after the second transfection. To perform RNA Immunoprecipitation, one round of siRNA transfection was carried out, as explained, the day of seeding. Cells were collected ∼65 h after siRNA transfection.

Immunoblot

HepG2 were lysed using RIPA buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1% NP40, 0.5% Sodium Deoxycholate, 0.1% SDS, Halt™ Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific)). Protein extracts (30 μg for NDUFS8, 50 μg for TIAM2 and 5 μg for all other proteins) were immunoblotted using the indicated primary antibodies (Supplementary Table S2) and anti-mouse, anti-rabbit or anti-rat IgG-linked HRP secondary antibodies (GE Healthcare) followed by ECL (Advansta).

Small RNA-seq and RNA-seq

Total RNA was extracted using TRIzol (Thermo Fisher Scientific) according to the manufacturer's instructions. Small RNA-seq (single end, 50 bp) was carried out by BGI Genomic Services (HepG2) or Fasteris (HEK-293T) in triplicate samples. Raw data were processed using the Subread package (version 1.6.0) as previously described (21) and the reference annotation was obtained from miRBase release 22.1 database (22). Statistical analysis was performed using DESeq2 (version 2.11.40.2). RNA-seq (paired-end, 125 bp) was carried out by BGI Genomic Services in triplicates. Raw data were processed using HISAT2 (version 2.1.0) and featureCounts (version 1.6.3), statistical analysis was performed using DESeq2. Reference annotation was obtained from ENSEMBL (GRCh38.96).

RT-qPCR, modified 5′ RLM RACE and RNA EMSA

Total RNA was extracted from HepG2 cells using TRIzol reagent (Thermo Fisher Scientific) and RNA was treated with DNAse I (Promega) according to the manufacturer's

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

instructions. RNA was used for RT-PCR and modified 5′ RLM-RACE as described previously (19).

For RT-qPCR, RT was performed using TaqMan™ Reverse Transcription Reagent or TaqMan™ Advanced miRNA cDNA Synthesis Kit (Thermo Fisher). qPCRs were performed using GoTaq® Probe qPCR Master Mix (Promega) or TaqMan® Fast Advanced Master Mix (Thermo Fisher).

Modified 5′ RLM RACE was performed according to the manufacturer's instructions (FirstChoice™ RLM-RACE kit, ThermoFisher Scientific). In order to detect premature miRNAs, the step using calf intestine alkaline phosphatase was omitted. Sequences of the primers used for PCR amplification are shown in Supplementary Table S3.

RNA EMSA was performed as described previously (15) using recombinant NF90 and recombinant DGCR8 dsRBDs (amino acids 484–773) in at least three replicates. The pri-miRNA probes were amplified by PCR using the primers shown in Supplementary Table S3. Sequences of mutant pri-miRNAs are shown in Supplementary Table S4.

RNA immunoprepicipation (RIP)

RIP was performed as previously described (23). HepG2 were seeded in 100 mm culture dishes and transfected with siRNAs the day of seeding as aforementioned. Cells were harvested ∼65 h after the treatment and lysed for 15 min in RIP buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 2.5 mM $MgCl_2 \bullet 6H_2O$, 250 mM sucrose, 0.05% (v/v) NP-40 and 0.5% (v/v) Triton X-100) containing 20 U ml^{-1} of RNasin (Promega), 1 mM DTT, 0.1 mM PMSF and EDTA-free protease and phosphatase inhibitor. After centrifugation, lysates were incubated for 4 h at 4°C with 2 μg of antibodies recognizing NF90, Drosha and IgG control and then incubated for 1 h at 4°C with Dynabeads™ Protein A (ThermoFisher Scientific). After incubation, beads were washed five times with RIP buffer for 5 min at 4°C and RNA was extracted as previously explained. RNA was treated with DNAse I (Promega) and RT was performed using SuperScript™ III Reverse Transcriptase (ThermoFisher Scientific) according to the manufacturer's instructions. cDNA was treated with RNAse H (ThermoFisher Scientific) and the samples were used to perform qPCRs using QuantiTect SYBR® Green PCR Kit (Qiagen) according to the manufacturer's instructions.

Splicing analysis

Splicing analyses were carried out as previously described (19). HepG2 were seeded in six-well plates and transfected with siRNAs, as aforementioned. Approximately 65 h after the second transfection, RNA was extracted using TRIzol reagent (ThermoFisher Scientific) and treated with DNAse I (Promega) according to the manufacturer's instructions. RT was performed using SuperScript™ III Reverse Transcriptase (ThermoFisher Scientific) and cDNA was treated with RNAse H (ThermoFisher Scientific). qPCRs were performed using QuantiTect SYBR® Green PCR Kit (Qiagen) using primers overlapping exon–intron boundaries to detect unspliced pre-mRNAs or primers amplifying exon-exon boundaries to detect the spliced mRNA.

Bioinformatic analyses

Enhanced UV crosslinking followed by immunoprecipitation (eCLIP) data for NF90, DGCR8 and DROSHA obtained in HepG2 cells by Nussbacher and Yeo (24) were retrieved from the NCBI database (NF90 eCLIP: ENCSR786TSC; DGCR8 eCLIP: ENCSR061SZV; DROSHA eCLIP: ENCSR834YLD). Peaks were filtered based on Fold Change (FC ≥ 1.5) and *P*-value (Bonferroni-Adj *P*-val ≤ 0.05). Distribution of eCLIP reads along the miRNAs was evaluated using deeptools software (version 3.1.3). Bigwig files from different replicates were merged using bigWigMerge v2. The base pair probability at each position of miRNA hairpins was calculated using RNAfold software (version 2.4.7).

Free energy analysis was performed using RNAfold software, version 2.4.7. Statistical analysis was performed using R (version 3.5.1).

Validated targets of the double positive miRNAs were extracted from MirTarBase database, release 7.0 (25). Gene ontology was performed on the expressed validated target using DAVID Functional Annotation Tool database version 6.8 (https://david.ncifcrf.gov) (26). Motif search was performed using MEME (version 5.0.5).

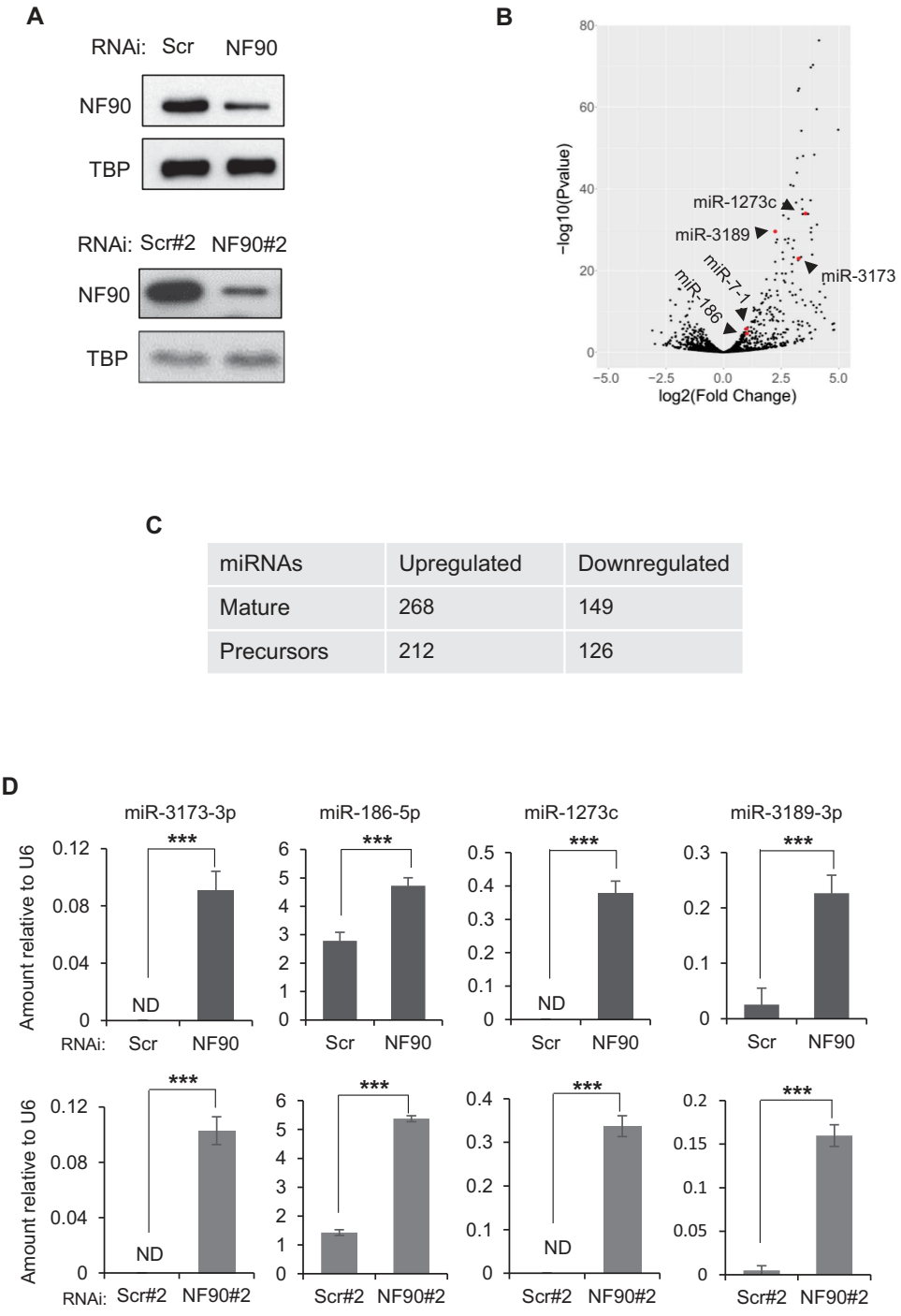
RESULTS

NF90 affects the abundance of a subset of human miRNAs

To determine the effect of NF90 on the abundance of miRNAs, we performed small RNA-seq of biological triplicate samples obtained from HepG2 cells that had been transfected with a non-targeting control siRNA (siScr) or an siRNA targeting NF90 (siNF90) (Figure 1A, top panel). Of 1917 miRNA precursors annotated in miRBase, 1105, which corresponds to 1661 mature 5p and 3p miRNA products, were found to be expressed in HepG2 cells. Following loss of NF90, differential expression analysis (fold change ≥ 1.5 or ≤ 0.667; Adj*P*-value ≤ 0.05) showed that 268 mature miRNAs, corresponding to 212 precursor miRNAs, were upregulated while 149, corresponding to 126 precursor miRNAs, were downregulated (Figure 1B). The number of upregulated and downregulated miRNAs in HepG2 cells after loss of NF90 is summarized in Figure 1C. MiRNAs that have previously been shown to be repressed by NF90, miR-7-1 (14) and miR3173 (19), were found to be upregulated in HepG2 cells following loss of NF90 (Figure 1B, red dots).

The effect of NF90 on the abundance of miRNAs observed by miRNA profiling were validated by RT-qPCR analysis of selected miRNAs, miR-3173-3p, miR-186-5p, miR-1273c and miR-3189-3p, from biological triplicate samples. The results obtained confirmed the effects observed by miRNA profiling (Figure 1B, D). In addition, RNA was extracted from cells transfected with an independent non-targeting siRNA (Scr#2) and an NF90-targeting siRNA (NF90#2) that has been described previously (19) (Figure 1A, lower panel). Quantification of miRNAs 3173-3p, -186-5p, -1273c and -3189-3p in biological triplicate samples (Figure 1D, lower panels) showed similar results to those obtained in Figure 1D upper panel, and also validated the results obtained by small RNA-seq. While we cannot

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020



miRNAs	Upregulated	Downregulated
Mature	268	149
Precursors	212	126

Figure 1. NF90 modulates the expression level of a subset of miRNAs in HepG2 cells. (**A**) Extracts of HepG2 cells transfected with non-targeting control siRNAs (Scr, Scr#2) or siRNA targeting NF90 (NF90, NF90#2) as indicated were analyzed by immunoblot using the antibodies indicated. (**B**) Total RNA extracted from cells transfected with siScr or siNF90 were analyzed by small RNA-seq. Results are shown as $\log_2$ fold change versus $-\log_{10}$ *P*-value. (**C**) Table summarizing the number of mature miRNAs and pri-miRNAs modulated in HepG2 cell line upon loss of NF90, according to small-RNA seq. (**D**) Total RNA extracted from cells described in (A) were analyzed by Taqman RT-qPCR as indicated. Results were normalized by those obtained for U6 abundance in the same samples. ND indicates 'not detected'. Data represent mean ± SEM obtained from three independent experiments (***$P < 0.001$, independent Student's *t* test).

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

exclude the possibility that a proportion of the small RNA-seq results could be due to off-target effects of the siRNAs, since only a single control and NF90-targeting siRNA were used, validation of a subset of the results using additional control and NF90-targeting siRNA suggests that the data are, to some extent, robust.

To evaluate whether the effect of NF90 on miRNA abundance might be cell type specific, we performed small RNA-seq in biological triplicate in HEK-293T cells transfected with control or NF90-targeting siRNA (Supplementary Figure S1A). Of 1917 annotated miRNA precursors, 1121, corresponding to 1647 mature miRNAs, were expressed in HEK-293T. Differential expression analysis (fold change $\geq$ 1.5 or $\leq$ 0.667; Adj*P*-value $\leq$ 0.05) revealed that 278 mature miRNAs, corresponding to 217 miRNA precursors, were upregulated following loss of NF90 while 84 mature miRNAs, corresponding to 77 precursors, were downregulated (Supplementary Figures S1B, C). Comparing upregulated miRNAs in the two cell types, we found 139 miRNAs that were upregulated in both cell lines after NF90 knockdown (Supplementary Figure S1D). This represents >65% of miRNAs upregulated in HepG2 and 64% of those upregulated in HEK-293T. Thus, NF90 appears to regulate a common subset of miRNAs.

NF90 associates with a subset of pri-miRNAs

To determine which of the miRNAs upregulated upon loss of NF90 (Figure 1B) are direct targets of NF90, that is, pri-miRNAs that are bound by NF90, we took advantage of enhanced UV crosslinking followed by immunoprecipitation (eCLIP) dataset obtained in HepG2 cells (24). Analysis of HepG2 eCLIP data revealed 38 pri-miRNAs for which eCLIP peaks overlapped annotated pri-miRNA localizations +/− 25 nt of flanking region (FC $\geq$ 1.5 and Bonferroni Adj*P* $\leq$ 0.05), as depicted in Figure 2A and Supplementary Table S5. Pri-miR-3173 and pri-miR-7-1 were among the 38 NF90-associated pri-miRNAs (Figure 2A, red dots).

We next analysed eCLIP read coverage across the pri-miRNA hairpin ±200 bp for the 38 NF90-associated miRNAs compared to all pri-miRNAs (Figure 2B). As expected, analysis of all pri-miRNAs did not show significant read coverage for NF90 association. In contrast, NF90-associated miRNAs showed highest read coverage over the region having the strongest base pair probability and therefore likely corresponding to the double stranded pri-miRNA stem (Figure 2B). The region corresponding to the terminal loop, which has a low base pair probability, was not significantly bound by NF90. Interestingly, NF90 also appeared to bind to the pri-miRNA flanking region. Browser shots showing NF90 association with pri-miR-7-1, pri-miR-186 and pri-miR-1273c by eCLIP are shown in Supplementary Figure S2A.

To validate NF90 association with pri-miRNAs identified by eCLIP analysis (Figure 2A), we performed RNA EMSA using pri-miR-186, pri-miR-3173, pri-miR-1273c and pri-miR-3189 as radiolabeled probes together with recombinant NF90 (Supplementary Figure S2B), as described previously for pri-miR-7-1 and pri-miR-3173 (14,19). RNA EMSA, performed in triplicate, confirmed NF90 association with pri-miR-186, pri-miR-3173, pri-miR-1273c and pri-miR-3189 (Figure 2C and Supplementary Figure S2C). Similarly, RNA EMSA confirmed that NF90 was not highly associated with pri-miR-200a, as indicated by eCLIP (Figure 2C). NF90 association with the pri-miRNAs identified by eCLIP analysis was also validated for several endogenous pri-miRNAs by performing RNA immunoprecipitation (RIP). RIP confirmed the association of NF90 with region proximal to the endogenous pri-miRNA (Figure 2D and Supplementary Figure S2D), while negative controls, pri-miR-200a and DALRD3, were not significantly associated with NF90. In contrast, pri-miR-200a was significantly bound by Drosha (Figure 2D and Supplementary Figure S2D). While not all NF90-bound pri-miRNAs identified by eCLIP have been tested, RIP analysis confirmed the association with NF90 *in vivo* for at least several.

Previous studies have indicated that NF90 may act as a competitor of Microprocessor for binding to pri-miRNAs (13–15,19). We therefore analysed eCLIP data for DGCR8 and Drosha performed in HepG2 cells (24). Association of DGCR8 was detected at 203 pri-miRNAs, while 147 pri-miRNAs were positive for Drosha binding (Figure 3A). Not surprisingly, there was a significant overlap between pri-miRNAs that were bound by both subunits of Microprocessor (Figure 3A). Indeed, 125 pri-miRNAs were associated with both factors, which represents approximately 60% and 85% of pri-miRNAs positive for DGCR8 and Drosha, respectively. Interestingly, only 10 pri-miRNAs bound by NF90 overlapped with those bound by either DGCR8 or Drosha, which represents approximately 24% overlap with DGCR8 and 13% overlap with Drosha (Figure 3A). This result indicates that NF90-associated pri-miRNAs are not highly associated with Microprocessor. Analysis of eCLIP reads showed association of DGCR8 with both apical and stem regions of pri-miRNAs (Figure 3B), as expected (27).

We further analysed eCLIP read coverage over the pri-miRNAs that were found to be associated with both NF90 and DGCR8. While the profile for DGCR8 was similar to that for all DGCR8 positive pri-miRNAs (compare Figure 3C, top panel to Figure 3B), the profile for NF90 read coverage was somewhat different to that for all NF90-positive pri-miRNAs (compare Figure 3C, lower panel, to Figure 2B). Interestingly, for pri-miRNAs that are bound by both DGCR8 and NF90, the profiles appear to be complementary (Figure 3C, compare top and lower panels). Plot profiles of DROSHA and DGCR8 eCLIP data suggest that pri-miRNAs common with NF90 (shown with red dots) are not among the most enriched for Microprocessor binding (Figure 3D and Supplementary Figure S3).

To further explore the competition between NF90 and the Microprocessor for the binding of pri-miRNAs, we performed RNA EMSA on pri-miR-3189 and pri-miR-1273c using recombinant NF90 and the dsRNA-binding domains of DGCR8 (Supplementary Figures S2B and S4A). Upon addition of rNF90, a shift corresponding to the formation of NF90-pri-miRNA complex and a reduction in the intensity of the band corresponding to DGCR8-pri-miRNA complex could be detected (Figure 4A). These results indicate that NF90 competes with Microprocessor for binding to certain pri-miRNAs, at least *in vitro.* Further analysis will be required to determine whether this competition also occurs *in vivo*.

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

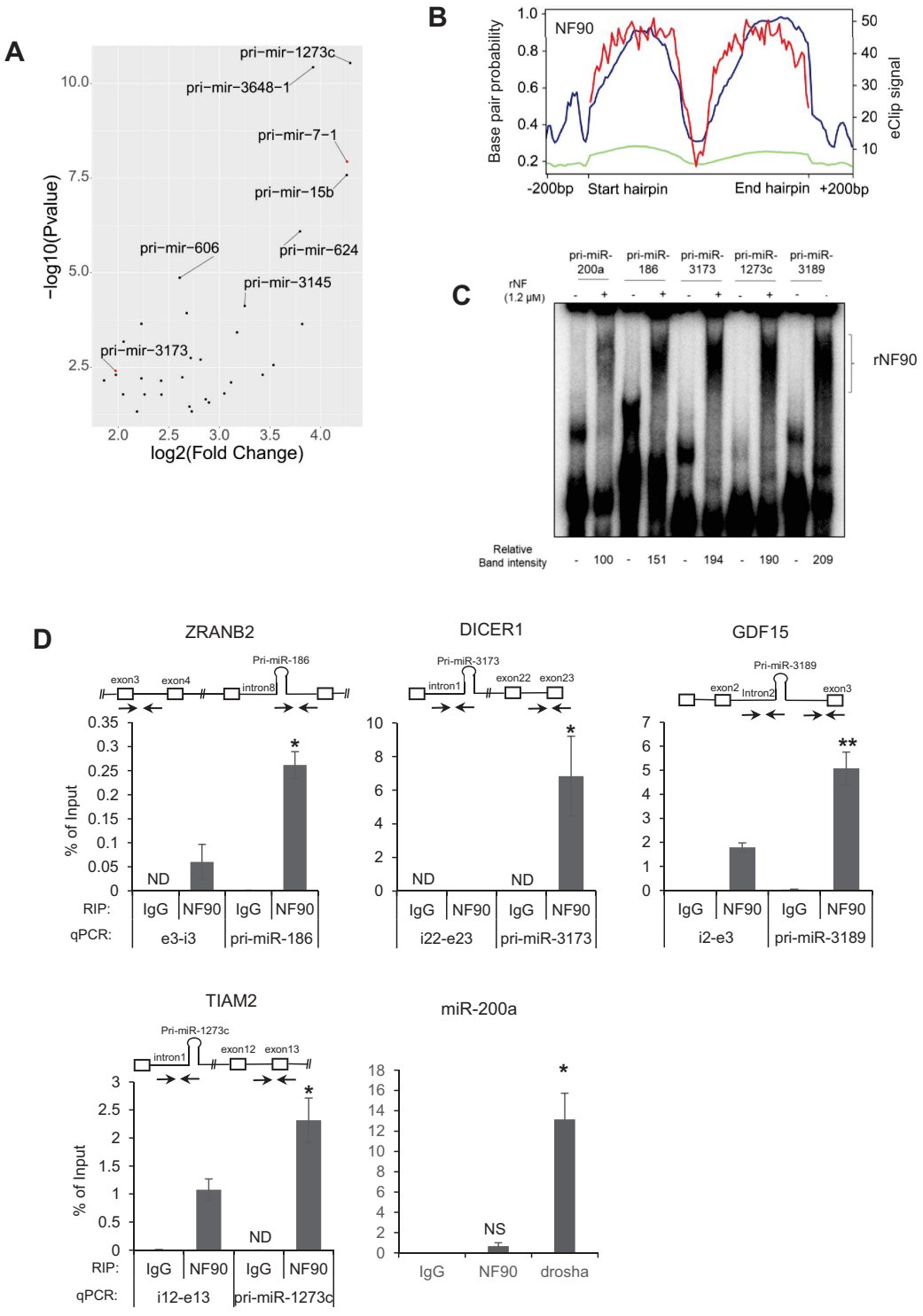


Figure 2. NF90 is associated with a subset of pri-miRNAs in HepG2 cells. (**A**) Dot plot representation of eCLIP data showing the 38 pri-miRNAs significantly associated with NF90 in HepG2 cells. Graph shows log2 fold change versus $-\log_{10}$ *P*-value. (**B**) Distribution of NF90 eCLIP reads along the region ±200 bp of NF90-associated pri-miRNAs (blue) or all miRNAs (green) and base pair probability of NF90-associated hairpins (red). (**C**) RNA EMSA performed using recombinant NF90 was probed with radiolabelled pri-miRNAs as indicated. rNF90-pri-miRNA complexes are indicated on the figure. (**D**) RIP analysis of HepG2 cells transfected with NF90 targeting siRNA or a non-targeting control (Scr), as indicated using anti-NF90, anti-Drosha or a control IgG antibody. Immunoprecipitates were analyzed by RT-qPCR amplifying a region proximal or distal to the miRNAs. ND indicates 'Not Detected'. NS indicates 'Not Significant'. Data represent mean ± SEM obtained from 3 independent experiments (*$P < 0.05$, **$P < 0.01$, independent Student's *t* test).

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

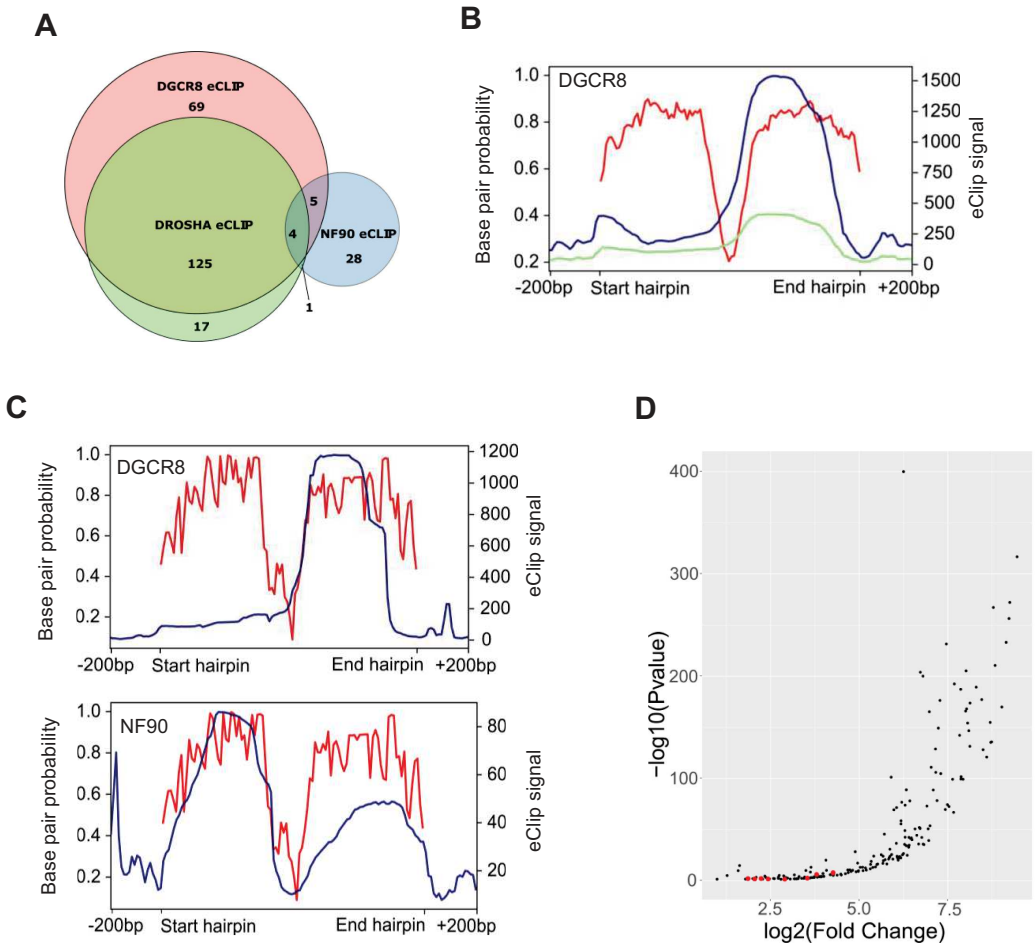


Figure 3. NF90-associated pri-miRNAs are poorly associated with Microprocessor. (**A**) Venn diagram showing the number of pri-miRNAs associated with DGCR8, Drosha or NF90 detected by eCLIP, as indicated. (**B**) Distribution of DGCR8 eCLIP reads along the region ±200 bp of DGCR8-associated pri-miRNAs (blue) or all miRNAs (green) and base pair probability of DGCR8-associated hairpins (red). (**C**) Distribution of eCLIP reads along the region ±200 bp of pri-miRNAs associated with both DGCR8 and NF90 (blue) and base pair probability of the hairpins (red). Left panel shows DGCR8 eCLIP reads, right panel shows NF90 eCLIP reads in blue. (**D**) Dot plot representation of eCLIP data showing 203 pri-miRNAs significantly associated with DGCR8 in HepG2 cells. Graph shows log2 fold change versus $-\log_{10}$ *P*-value. Red dots indicate pri-miRNAs that are also significantly associated with NF90.

We next tested whether loss of NF90/NF45 or Drosha/DGCR8 complexes could affect the binding of the complexes to endogenous pri-miRNAs *in vivo*. We performed RIP of NF90, Drosha or IgG control after downregulation of either NF90/NF45 or Drosha/DGCR8. Drosha association with the region surrounding the target pri-miRNAs was significantly enhanced after downregulation of NF90/NF45, while NF90 association was significantly enhanced after downregulation of Drosha/DGCR8 only for pri-miR-1273c (Figure 4B and Supplementary Figure S4B). This could be explained considering that these miRNAs are already poorly bound by the Microprocessor. To test this hypothesis, we analysed the association of NF90 to two pri-miRNAs poorly bound by NF90, pri-miR-200a and pri-miR-425. Notably, NF90 association with these miRNAs was significantly increased after loss of Drosha/DGCR8 complex (Supplementary Figure S4C). On the other hand, downregulation of NF90/NF45 complex did not significantly affect the association of pri-miR-200a and pri-miR-425 with Drosha (Supplementary Figure S4C), possibly because these miRNAs are poorly bound by NF90/NF45 under control conditions. Taken together, these results suggest that target pri-miRNAs may have binding preferences for either NF90/NF45 or Microprocessor under wild-type conditions, but that the relative abundance of these complexes can also influence the observed binding to specific pri-miRNAs.

Pri-miRNAs that are bound and downregulated by NF90 are highly stable

We next asked whether NF90 association with pri-miRNAs might affect their cropping by Microprocessor. If so, loss of NF90 would be predicted to increase the abundance of the mature miRNA products, as observed previously (14,15,19). MiRNA profiling revealed that of the 38 NF90-associated pri-miRNAs, 22 showed an increase in mature miRNA products, representing more than 57% of NF90-

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

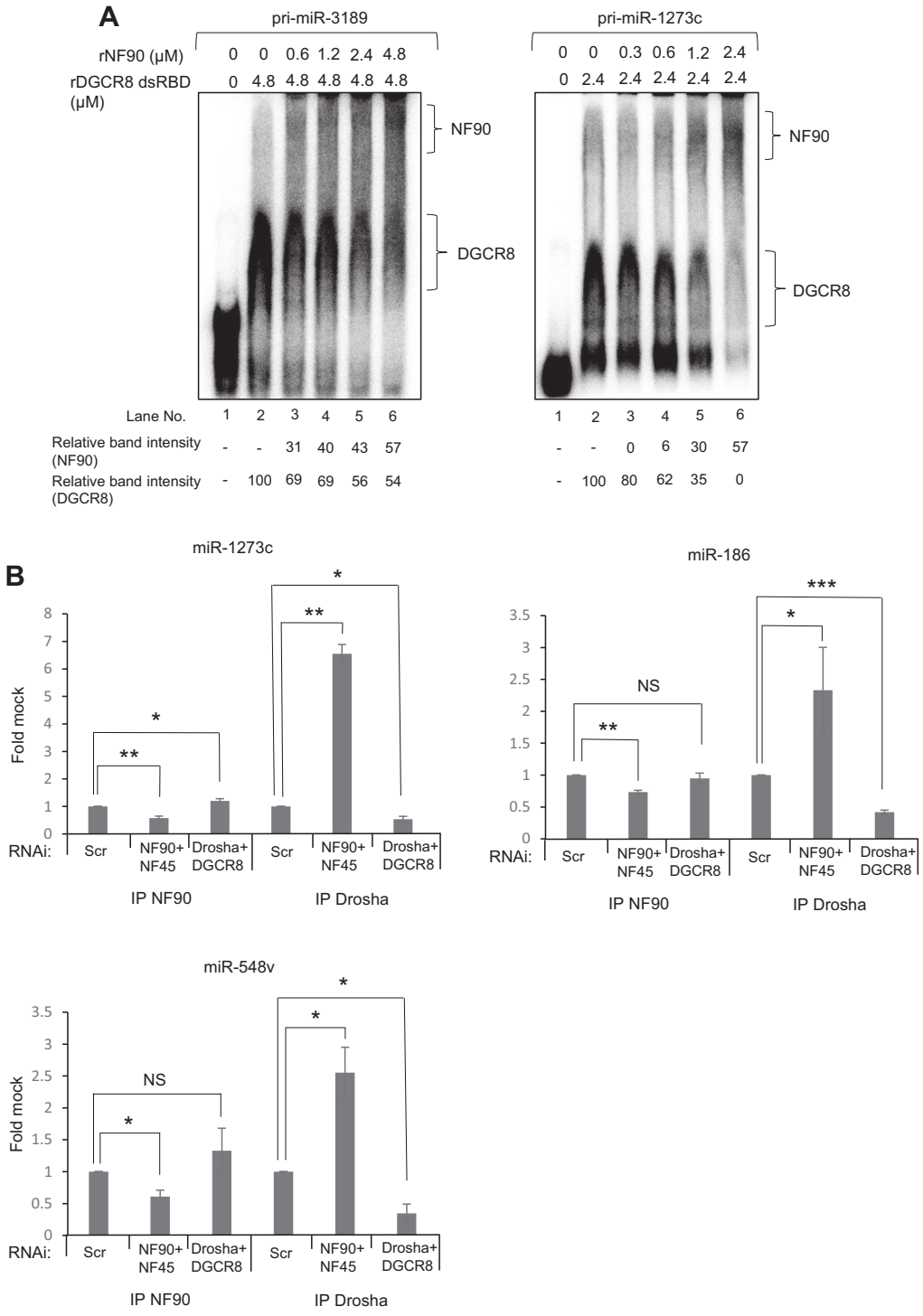


Figure 4. NF90 competes with the Microprocessor for the binding to pri-miRNAs. (**A**) RNA EMSA carried out using rDGCR8 dsRBD either alone or together with increasing amounts of rNF90 and probed with radio-labelled pri-miR-3189 or pri-miR-1273c. (**B**) Immunoprecipitates obtained using anti-NF90, anti-Drosha or a control antibody were analyzed by RT-qPCR amplifying a region proximal to the pri-miRNAs. The fold change relative to the control antibody sample was calculated and results are presented relative to the control sample (siScr), which was attributed a value of 1. Data represent mean ± SEM obtained from three independent experiments (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, NS 'Not Significant'. independent Student's *t* test).

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

associated pri-miRNAs, while only two were decreased (Supplementary Tables S6 and S7). Thus, we identified a subset of 22 pri-miRNAs that are bound by NF90 and whose abundance is increased following loss of NF90, which we named 'double-positive' pri-miRNAs. Both pri-miR-7-1 and pri-miR-3173 were identified within the double positive subset. Thus, NF90 downregulates the expression of most of its target pri-miRNAs.

Gene ontology of validated mRNA targets of double positive miRNAs revealed an implication particularly in cancer and infection by viruses, such as Epstein Barr Virus (EBV), hepatitis B virus (HBV), and human T lymphoma virus type 1 (HTLV1), as well as viral carcinogenesis (Supplementary Figure S5). This result is interesting given that NF90 translocates from the nucleus to the cytoplasm following viral infection of cells (28). Thus, viral infection could result in the coordinated processing of the NF90-modulated subset of pri-miRNAs, whose target mRNAs are implicated in viral replication. Interestingly, several miRNAs upregulated following loss of NF90 in this study have been shown to target RNAs expressed by influenza A virus subtypes. For instance, miR-3682 is involved in viral replication by targeting the NS gene of pH1N1 and H3N2 subtypes (29). Similarly, miR-4753 and miR-3145, which target PS and PB1 genes of H5N1 and H3N2 subtypes, are overexpressed in response to viral infection and inhibit viral transcription and replication (30).

We wondered whether pri-miRNAs that are associated with NF90 and downregulated upon its loss might share a common characteristic that would make them targets for NF90 binding. A MEME search did not reveal a simple binding motif common to the 22 pri-miRNA sequences. Compared to all human pri-miRNAs, the subset of 22 double-positive pri-miRNAs did not show any significant difference in their overall length (mean = 82.5 nt compared to 81.88 nt) or in the size of the terminal loop (mean = 7.87 nt compared to 7.92 nt) (Figure 5A). In contrast, however, the minimal stretch containing a mismatch ≤1 nt was significantly longer for double-positive pri-miRNAs compared to all pri-miRNAs, with a mean of 27.68 nt for double-positive pri-miRNAs compared to 21.11 nt for all pri-miRNAs (Figure 5A). This analysis suggests that double-positive pri-miRNAs might be more stable, having a longer duplex and less bulges compared to all human pri-miRNAs. To further investigate this possibility, we compared the free energy of the 22 double-positive pri-miRNAs compared to all pri-miRNAs. The 22 double-positive pri-miRNAs had a lower free energy (mean = −42.26) compared to all pri-miRNAs (mean = −38.19), as shown in Figure 5B. Taken together, these data suggest that double positive pri-miRNAs are more stable and have less mismatches than all pri-miRNAs. Predicted folding of double-positive pri-miRNA sequences also revealed highly stable structures with very few bulges, compared to pri-miR-200a, which is not highly associated with NF90 (Supplementary Figure S6A).

To test the idea that NF90 can bind to pri-miRNAs that have a stable structure with few bulges, we designed mutations within NF90-binding pri-miRNAs predicted to reduce stability and form bulge-like regions that might disrupt NF90 association. For each of the NF90-associated pri-miRNAs tested, we designed two mutant structures that

A

Feature (nt)	All pri-miRNAs	Double positive pri-miRNAs
Length	81.88	82.5
Terminal loop size	7.92	7.87
Longest duplex	21.11	27.68

B

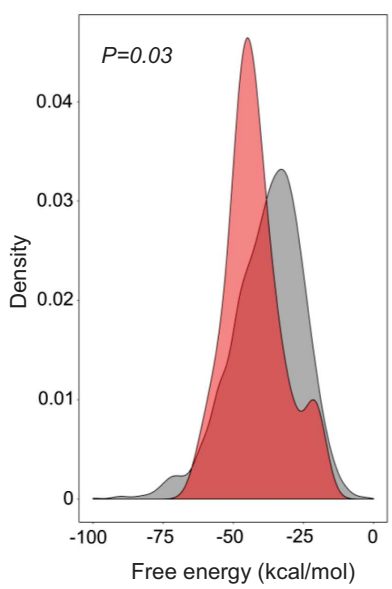


Figure 5. NF90 associates with a subset of highly stable pri-miRNAs. (**A**) Structural characteristics of all human pri-miRNAs and NF90 double positive pri-miRNAs. (**B**) Graph showing the free energy of all pri-miRNAs (grey) and NF90 double positive pri-miRNAs (red).

would be less stable than wild-type structures. (Figure 6A). WT and mutated pri-miRNAs were tested for NF90 association by RNA EMSA. As shown in Figure 6B and Supplementary Figure S6B, mutation of pri-miR-3173 or pri-miR-186 to less stable structures diminished NF90 binding. On the other hand, mutation of pri-miR-200a to a more stable structure enhanced NF90 binding. These data suggest that NF90 shows a preference for association with stable pri-miRNA hairpin structures having few bulge regions.

We then wondered whether pri-miRNAs whose mature products increased following loss of NF90, but were not considered eCLIP-positive using the applied cut-offs, might share the characteristics identified for double-positive pri-miRNAs. We therefore calculated the longest duplex length, allowing a mismatch of 1 nt, for the group of 181 upregulated but eCLIP negative pri-miRNAs, and 124 downregulated pri-miRNAs, as well as for those falling outside these groups (other) (Figure 7A). Interestingly, pri-miRNAs upregulated after loss of NF90 and eCLIP negative have a significantly longer duplex than all pri-miRNAs or other pri-miRNAs. Indeed, the duplex length is similar to that observed for the double positive group. In contrast, pri-

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

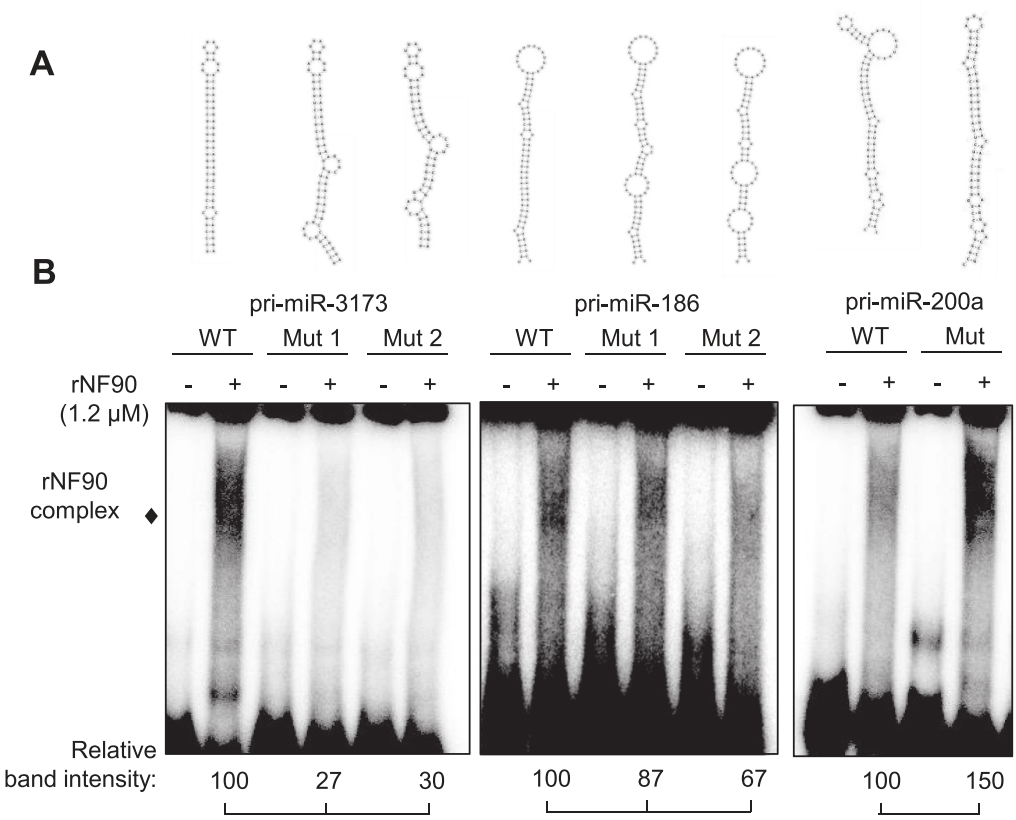


Figure 6. Modification of pri-miRNA structure alters NF90 binding. (**A**) Representations of wt or mutant pri-miRNAs sequences, as indicated. (**B**) RNA EMSA performed using recombinant NF90 and probed with radiolabelled pri-miRNAs as indicated. rNF90-pri-miRNA complexes are indicated on the figure. Relative band intensities (normalized to signal for wt) are shown below.

miRNAs downregulated upon loss of NF90 have a shorter duplex compared to all pri-miRNAs or other pri-miRNAs. We then calculated the mean free energy for the upregulated, eCLIP-negative group and the downregulated group of pri-miRNAs (Figure 7B). Similarly, when compared to all pri-miRNAs, the upregulated, eCLIP-negative group of pri-miRNAs had a significantly lower free energy. Free energy of the downregulated group was similar to that of all pri-miRNAs. In contrast, terminal loop size was comparable between the two groups; 7.86 nt (downregulated group) compared with 8.64 nt (upregulated eCLIP-negative group). Of note, total pri-miRNA length was higher for the upregulated eCLIP-negative group (87.01 nt) compared to the downregulated group (77.79 nt). These analyses suggest that upregulated, eCLIP-negative pri-miRNAs share some characteristics with double-positive pri-miRNAs. It is feasible that some NF90-associated pri-miRNAs were not detected by eCLIP analysis or did not pass the selection criteria used to identify eCLIP-positive pri-miRNAs. To test this idea, we selected two pri-miRNAs, pri-miR-4755 and pri-miR-4766, from the upregulated, eCLIP-negative group whose structure corresponds to the defined criteria for NF90 association, that is, having low free energy and few mismatches (Supplementary Figure S7A). NF90 binding to the pri-miRNAs was tested by RNA EMSA (Figure 7C and Supplementary Figure S7B). Indeed, both pri-miR-4755 and pri-miR-4766 were found to be significantly associated with NF90.

NF90 modulates the expression of a subset of genes hosting NF90-associated pri-miRNAs

Approximately 70% of human miRNAs are located in an intron of a host gene. Out of 22 double-positive pri-miRNAs, 20 are exclusively intronic. Two double-positive pri-miRNAs are found in either the 3′ UTR or an intron depending on transcript usage (Supplementary Table S6).

To determine whether loss of NF90 also affected the expression or splicing efficiency of the host genes, we performed RNA-seq in HepG2 cells transfected with control siRNA or siRNA targeting NF90. Loss of NF90 significantly diminished expression of three genes containing NF90-associated pri-miRNA; growth differentiation factor 15 (GDF15) hosting pri-miR-3189, 1-acylglycerol-3-phosphate *O*-acyltransferase 5 (AGPAT5) hosting pri-miR-4659a and zinc finger RAN-binding domain containing two (ZRANB2) hosting pri-miR-186 (Figure 8A). Furthermore, the splicing efficiency of introns containing pri-miRNAs downregulated by loss of NF90 was determined by RT-PCR for several targets (Figure 8B). Splicing efficiency was diminished for three pre-mRNAs containing NF90-associated pri-miRNAs: T-cell lymphoma invasion

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

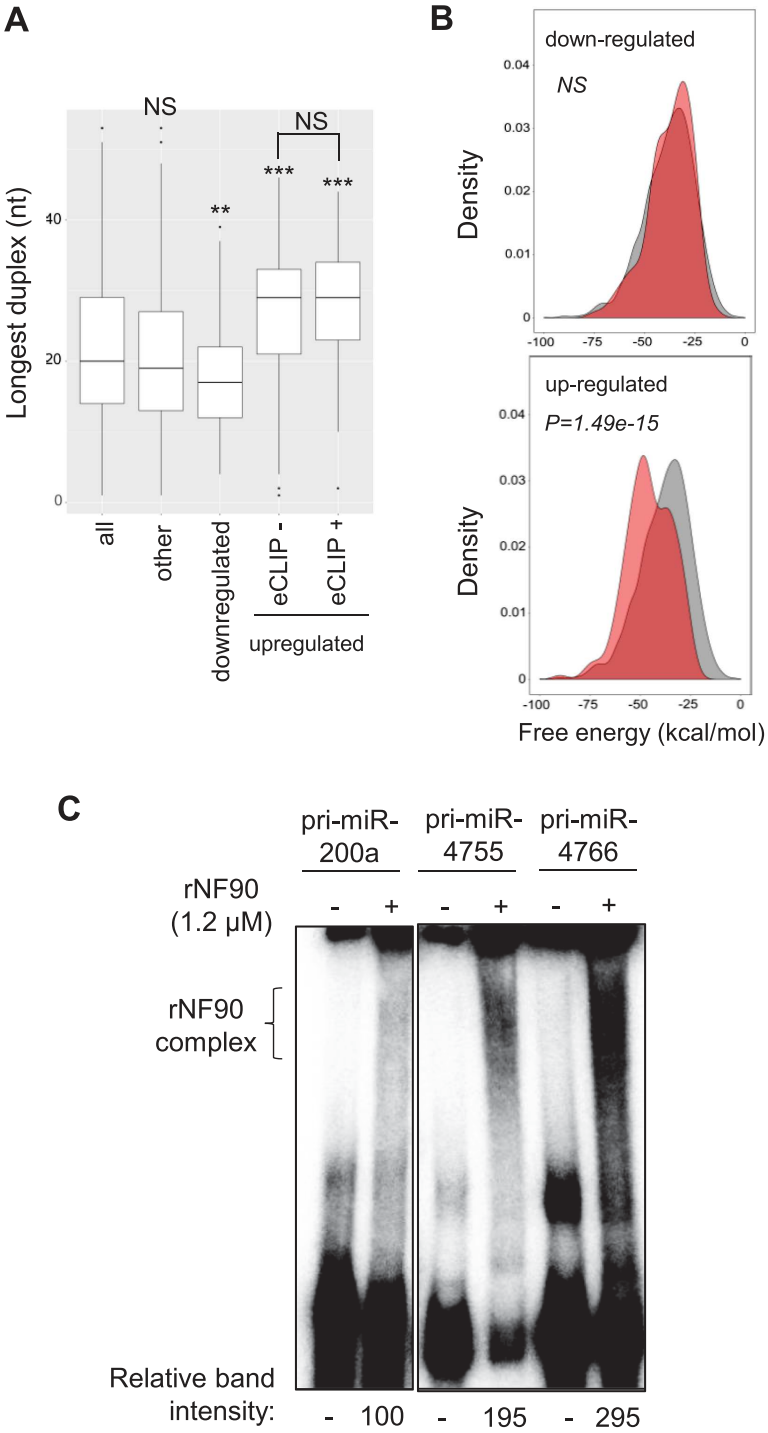


Figure 7. Pri-miRNAs whose mature products are upregulated following loss of NF90 share a similar structure. (**A**) Box plot representation of the longest duplex length of pri-miRNAs sorted into the indicated categories (*$P < 0.05$, ***$P < 0.001$, NS, not significant, Wilcoxon test). (**B**) Graphical representation of the free energy of pri-miRNAs whose mature products are downregulated or upregulated as indicated following loss of NF90 (red) compared to all pri-miRNAs (gray). (**C**) RNA EMSA performed using recombinant NF90 and probed with radiolabelled pri-miRNAs as indicated. rNF90-pri-miRNA complexes are indicated on the figure. Relative band intensities (normalized to pri-miR200a) are shown below.

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

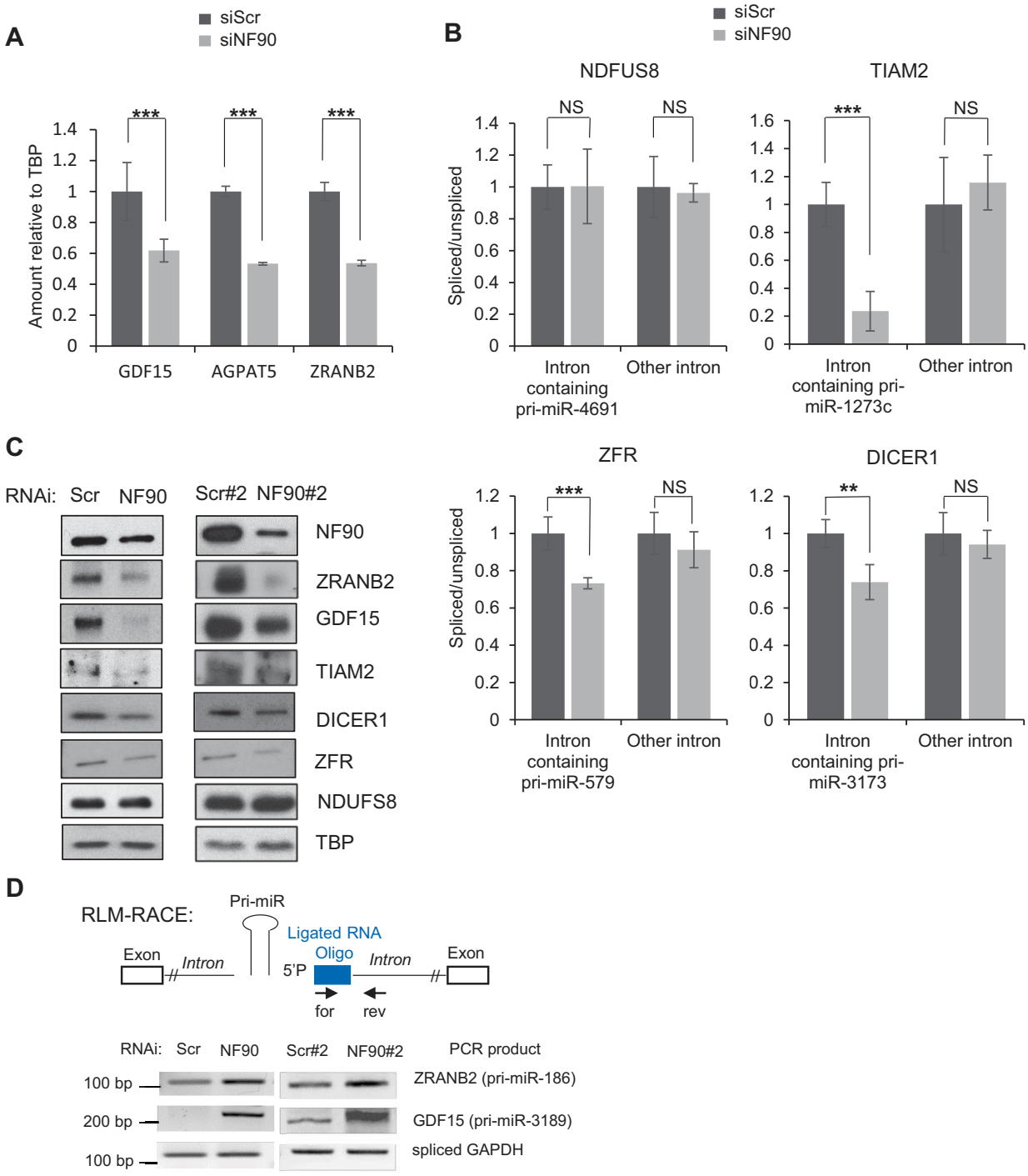


Figure 8. NF90 impacts expression of genes hosting pri-miRNAs. (**A**) Extracts of HepG2 cells transfected with siRNA targeting NF90 or a non-targeting control (Scr) as indicated were analyzed by RNA-seq and DESeq2. Data represent mean ± SEM obtained from three independent samples (***$P <$ 0.001, independent Student's *t* test). (**B**) The abundance of exon-intron junctions and exon-exon junctions in samples described in A was measured by RT-qPCR using PCR primers amplifying spliced or unspliced transcripts including introns containing pri-miRNAs or other introns. The splicing efficiency was calculated by the ratio of spliced to unspliced transcripts. Values obtained for the control sample (siScr) were attributed a value of 1. NS indicates 'Not Significant'. The graphs represent the mean ± SEM obtained from three or more independent experiments (**$P < 0.01$, ***$P < 0.001$, independent Student's *t* test). (**C**) Extracts of HepG2 cells transfected with siRNA targeting NF90 (NF90, NF90#2) or a non-targeting control (Scr, Scr#2) as indicated were analyzed by immunoblot using the antibodies indicated. (**D**) NF90 modulates transcript cleavage at the region containing miRNA. Extracts of HepG2 cells transfected with siRNAs targeting NF90 (NF90, NF90#2) or non-targeting controls (Scr, Scr#2) as indicated were analyzed by modified 5′ RLM-RACE. Forward and reverse primers used, and the predicted sizes of the PCR products are indicated.

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

and metastasis 2 (TIAM2), hosting pri-miR-1273c, Zinc Finger RNA binding protein (ZFR), hosting pri-miR-579, and DICER1, hosting pri-miR-3173 (Figure 8B). Interestingly, the splicing defect was detected for the intron containing the pri-miRNA but not for another intron within the same transcript (Figure 8B). In contrast, no significant effect was observed for NDUFS8, which hosts pri-miR-7113 and pri-miR-4691 that are not bound by NF90 and whose abundance are not affected by NF90 (Figure 8B).

The expression of these genes was analysed by western blot of extracts obtained from HepG2 cells transfected with control (Scr and Scr#2) and NF90-targeting (NF90 and NF90#2) siRNAs. All genes tested showed diminished expression upon loss of NF90, except NDUFS8 that showed no significant difference in expression (Figure 8C). Thus, NF90 modulates the expression of certain pri-miRNA host genes, including TIAM2, a known oncogene and metastasis factor in HCC (31,32).

Finally, to determine whether loss of gene expression correlated with increased pri-miRNA cropping following knock down of NF90, we performed modified RLM-5′ RACE as described previously (19), using extracts of cells transfected with control (Scr and Scr#2) and NF90-targeting (NF90 and NF90#2) siRNAs. Indeed, RLM RACE analysis showed enhanced cleavage of the intronic region of ZRANB2 hosting pri-miR-186 and GDF15 hosting pri-miR-3189 in extracts of NF90 knock down cells compared to controls (Figure 8D). This analysis indicates that loss of NF90 enhances transcript cleavage in the vicinity of the hosted pri-miRNA.

DISCUSSION

We and others have previously shown that NF90 can inhibit the processing of certain miRNA precursors (14,15,19). However, it was unclear how widespread the impact of NF90 might be on human miRNA biogenesis. Here, we have used genome-wide approaches to address the effect of NF90 on the miRNA pool in HepG2 HCC cells. Our data indicate that NF90 modulates the processing of a specific subset of miRNA precursors. NF90 is associated with at least 38 human pri-miRNAs, as indicated by analysis of eCLIP data obtained by Nussbacher and Yeo (24). Of these, 22 showed increased abundance of mature miRNA products following knock-down of NF90. Thus, association of NF90 with a pri-miRNA is likely to influence its fate. Most NF90-associated pri-miRNAs did not overlap with those bound by either DGCR8 or Drosha. Moreover, results obtained by RNA-EMSA support the idea that NF90 and Microprocessor may compete for the binding of the subset of pri-miRNAs, at least *in vitro*. Further analysis will be required to determine whether the competition also occurs *in vivo*. Of note, RIP analysis showed that loss of NF90/NF45 complex led to increased binding of Drosha at pri-miRNAs that were highly bound by NF90 in control conditions. Conversely, loss of Microprocessor increased binding by NF90 to pri-miRNAs that were not highly bound by NF90 in wild-type cells. Interestingly, for those pri-miRNAs that were bound by both NF90 and DGCR8, the binding profiles of the two factors were largely complementary. Furthermore, while the binding profile of DGCR8 was not noticeably different for this group compared to all pri-miRNAs bound by DGCR8, the binding profile of NF90 differed somewhat for this group compared to all pri-miRNAs bound by NF90. This could suggest that NF90 and DGCR8 might bind simultaneously to the pri-miRNA, and that the binding of DGCR8 may alter the binding mode of NF90 for such pri-miRNAs.

Since NF90 is a highly abundant and ubiquitously expressed protein, it might be expected that NF90-associated pri-miRNAs would be poorly processed in most cells. Indeed, the mature miRNA products of NF90 bound pri-miRNAs are very poorly expressed, or not expressed at all in control cells. They become readily detectable only upon loss of NF90. An exception is pri-miR-7-1, although interestingly, this miRNA shows tissue specific expression, being highly expressed only in brain and pancreas (33).

Our data suggests that pri-miRNAs upregulated after loss of NF90 share a common structure that might facilitate NF90 association with the stem region. This finding is consistent with a previous report showing structure-based recognition of adenovirus-expressed VA1 RNA by NF90 (34). Extensive mutational analysis of VA1 association with NF90 showed no specificity for nucleotide sequence but rather the requirement for a minihelix structure within the stem region. The pri-miRNAs identified in this study also exhibit a minihelix-like structure that appears to be necessary for NF90 binding. Indeed, RNA EMSA showed that NF90 association with pri-miR-3173 and pri-miR-186 could be diminished by introducing destabilizing mutations, while NF90 association could be acquired by increasing the stability of the stem region, as for pri-miR-200a.

Interestingly, our data predict that the subset of NF90-associated pri-miRNAs may extend beyond those detected by eCLIP analysis. Using the characteristics determined from the eCLIP-positive, upregulated pri-miRNA group, that is duplex length and free energy, we found that pri-miRNAs whose mature products were upregulated following loss of NF90 but were not positive by eCLIP analysis shared the same characteristics as the double positive group. The length of the duplex region and the free energy of the structure was comparable to that of double positive pri-miRNAs. RNA EMSA confirmed the predicted association with NF90 for two of these pri-miRNAs. Interestingly, both groups were significantly different to all pri-miRNAs or those that are unaffected by NF90 (other). Thus, it appears that the high specificity of eCLIP revealed a subset of pri-miRNAs that share a common structure. When this information was used to interrogate the group of pri-miRNAs who share the same biological response to loss of NF90, that is, upregulation of their mature products, we observed that both groups share the same characteristics. We predict that a certain number of the upregulated group likely do bind to NF90 but may escape detection by eCLIP. For example, as noted above, many of the pri-miRNAs are expressed at extremely low levels in control cells, which could make their association with NF90 difficult to detect.

Interestingly, pri-miR-7-1 processing has been shown to be influenced by another RBP, HuR, which recruits MSI2 to the terminal loop. Binding of HuR/MSI2 was found to

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

stabilize the stem region and led to diminished processing by microprocessor (35). It would be interesting to determine whether binding of HuR/MSI2 to pri-miR-7-1 might facilitate NF90 binding to the stem region, and compete with microprocessor. Similarly, it would be interesting to determine whether HuR/MSI2 can bind the terminal loop of other NF90-modulated pri-miRNAs in addition to pri-miR-7-1. NF90 may cooperate with other RBPs, such as HuR/MSI2 to control the processing of a subset of pri-miRNAs.

Another feature that the subset of NF90-modulated pri-miRNAs share is their restriction to human or primate lineages. Again, pri-miR-7-1 is an exception, being highly conserved throughout evolution. Thus, given that the subset of NF90-modulated pri-miRNAs are young and almost perfect hairpins, it is tempting to speculate that this group may have originated through recent insertion of repeat elements in the genome.

Interestingly, GO analysis of validated mRNA targets of the mature miRNAs showed significant enrichment for infection by viruses such as Epstein Barr Virus (EBV), hepatitis B virus (HBV) and human T lymphoma virus type 1 (HTLV1) and in viral carcinogenesis. Indeed, viral infection of cells induces translocation of NF90 from the nucleus to the cytoplasm (28). Thus, it is conceivable that pathological conditions such as viral infection could result in the coordinated processing of the NF90-modulated subset of pri-miRNAs, which target mRNAs important for viral replication.

Finally, transcriptomic analysis showed that association of NF90 with pri-miRNAs may diminish the expression of certain host genes, as described previously (19). Among the pri-miRNA-hosting transcripts that are downregulated after loss of NF90, two are noteworthy. The expression of TIAM2, hosting pri-miR-1273C, is down-regulated upon loss of NF90. TIAM2 is a known oncogene and metastasis factor in HCC (31,32). Levels of NF90 are elevated in HCC (14,20) and it would be interesting to determine whether NF90-dependent modulation of TIAM2 might contribute to pathogenesis. Loss of NF90 also diminished expression of growth differentiation factor 15 (GDF15), hosting pri-miR-3189. GDF15 is expressed and secreted by a limited number of tissues, including liver. When complexed with its receptor, GFRAL, in brain and CNS, GDF15 supresses appetite (see (36) for review). Cancer patients express high circulating levels of GDF15, which contributes to anorexia/cachexia. On the other hand, enhancement of GDF15 expression is a promising therapeutic strategy in the treatment of obesity. It would be interesting to determine whether high levels of NF90 in HCC may have a role in promoting expression of GDF15 from liver cells in cancer patients.

In summary, we have identified a subset of human pri-miRNAs that are bound by NF90. Analysis indicates that this subset shares a similar structure that appears to be favorable for NF90 binding. These data extend our knowledge of how processing of pri-miRNAs can be modulated by RBPs. This may be beneficial for understanding perturbations of miRNA levels in pathological conditions and could also open up novel treatment strategies using nanotherapeutics.

DATA AVAILABILITY

Small RNA-seq and RNA-seq data have been deposited at GEO (GSE132341).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We wish to thank Catherine Dargemont and Xavier Contreras for critical reading of the manuscript, and the Gene Regulation lab and Hervé Seitz for helpful discussions.

FUNDING

European Research Council [RNAmedTGS to R.K.]; MSD Avenir [HideInflame&Seq to R.K.]; Ministère de l'Enseignement Supérieur et de la Recherche et de l'Innovation scholarship (to G.G.); Japan Society for the Promotion of Science (Grant-in-aid for Young Scientists (B)) [17K15601, 19K16523 to T.H.]; Grant-in-aid for Scientific Research (C) [16K08590, 19K07370 to S.S.]. Funding for open access charge: [ERC, RNA MedTGS].
Conflict of interest statement. None declared.

REFERENCES

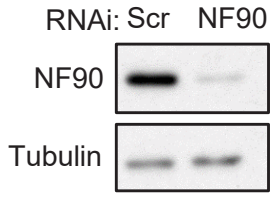
1. Ebert,M.S. and Sharp,P.A. (2012) Roles for microRNAs in conferring robustness to biological processes. *Cell*, **149**, 515–524.
2. Shenoy,A. and Blelloch,R.H. (2014) Regulation of microRNA function in somatic stem cell proliferation and differentiation. *Nat. Rev. Mol. Cell. Biol.*, **15**, 565–576.
3. Denli,A.M., Tops,B.B., Plasterk,R.H., Ketting,R.F. and Hannon,G.J. (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature*, **432**, 231–235.
4. Gregory,R.I., Yan,K.P., Amuthan,G., Chendrimada,T., Doratotaj,B., Cooch,N. and Shiekhattar,R. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature*, **432**, 235–240.
5. Ha,M. and Kim,V.N. (2014) Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.*, **15**, 509–524.
6. Finnegan,E.F. and Pasquinelli,A.E. (2013) MicroRNA biogenesis: regulating the regulators. *Crit Rev. Biochem. Mol. Biol.*, **48**, 51–68.
7. Krol,J., Loedige,I. and Filipowicz,W. (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.*, **11**, 597–610.
8. Mendell,J.T. and Olson,E.N. (2012) MicroRNAs in stress signaling and human disease. *Cell*, **148**, 1172–1187.
9. Winter,J., Jung,S., Keller,S., Gregory,R.I. and Diederichs,S. (2009) Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.*, **11**, 228–234.
10. Viswanathan,S.R., Daley,G.Q. and Gregory,R.I. (2008) Selective blockade of microRNA processing by Lin28. *Science*, **320**, 97–100.
11. Guil,S. and Caceres,J.F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.*, **14**, 591–596.
12. Michlewski,G., Guil,S., Semple,C.A. and Caceres,J.F. (2008) Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol. Cell*, **32**, 383–393.
13. Michlewski,G. and Caceres,J.F. (2019) Post-transcriptional control of miRNA biogenesis. *RNA*, **25**, 1–16.
14. Higuchi,T., Todaka,H., Sugiyama,Y., Ono,M., Tamaki,N., Hatano,E., Takezaki,Y., Hanazaki,K., Miwa,T., Lai,S. *et al.* (2016) Suppression of MicroRNA-7 (miR-7) biogenesis by nuclear factor 90-Nuclear factor 45 complex (NF90-NF45) controls cell proliferation in hepatocellular carcinoma. *J. Biol. Chem.*, **291**, 21074–21084.

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

15. Sakamoto,S., Aoki,K., Higuchi,T., Todaka,H., Morisawa,K., Tamaki,N., Hatano,E., Fukushima,A., Taniguchi,T. and Agata,Y. (2009) The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. *Mol. Cell. Biol.*, **29**, 3754–3769.
16. Heale,B.S., Keegan,L.P., McGurk,L., Michlewski,G., Brindle,J., Stanton,C.M., Caceres,J.F. and O'Connell,M.A. (2009) Editing independent effects of ADARs on the miRNA/siRNA pathways. *EMBO J.*, **28**, 3145–3156.
17. Yang,W., Chendrimada,T.P., Wang,Q., Higuchi,M., Seeburg,P.H., Shiekhattar,R. and Nishikura,K. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat. Struct. Mol. Biol.*, **13**, 13–21.
18. Jayachandran,U., Grey,H. and Cook,A.G. (2016) Nuclear factor 90 uses an ADAR2-like binding mode to recognize specific bases in dsRNA. *Nucleic Acids Res.*, **44**, 1924–1936.
19. Barbier,J., Chen,X., Sanchez,G., Cai,M., Helsmoortel,M., Higuchi,T., Giraud,P., Contreras,X., Yuan,G., Feng,Z. *et al.* (2018) An NF90/NF110-mediated feedback amplification loop regulates dicer expression and controls ovarian carcinoma progression. *Cell Res.*, **28**, 556–571.
20. Jiang,W., Huang,H., Ding,L., Zhu,P., Saiyin,H., Ji,G., Zuo,J., Han,D., Pan,Y., Ding,D *et al.*2015) Regulation of cell cycle of hepatocellular carcinoma by NF90 through modulation of cyclin E1 mRNA stability. *Oncogene*, **34**, 4460–4470.
21. Seco-Cervera,M., Gonzalez-Rodriguez,D., Ibanez-Cabellos,J.S., Peiro-Chova,L., Pallardo,F.V. and Garcia-Gimenez,J.L. (2018) Small RNA-seq analysis of circulating miRNAs to identify phenotypic variability in Friedreich's ataxia patients. *Sci Data*, **5**, 180021.
22. Kozomara,A., Birgaoanu,M. and Griffiths-Jones,S. (2019) miRBase: from microRNA sequences to function. *Nucleic Acids Res.*, **47**, D155–D162.
23. Bennasser,Y., Chable-Bessia,C., Triboulet,R., Gibbings,D., Gwizdek,C., Dargemont,C., Kremer,E.J., Voinnet,O. and Benkirane,M. (2011) Competition for XPO5 binding between Dicer mRNA, pre-miRNA and viral RNA regulates human Dicer levels. *Nat. Struct. Mol. Biol.*, **18**, 323–327.
24. Nussbacher,J.K. and Yeo,G.W. (2018) Systematic discovery of RNA binding proteins that regulate MicroRNA Levels. *Mol. Cell*, **69**, 1005–1016.
25. Chou,C.H., Shrestha,S., Yang,C.D., Chang,N.W., Lin,Y.L., Liao,K.W., Huang,W.C., Sun,T.H., Tu,S.J., Lee,W.H *et al.*2018) miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.*, **46**, D296–D302.
26. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
27. Nguyen,T.A., Jo,M.H., Choi,Y.G., Park,J., Kwon,S.C., Hohng,S., Kim,V.N. and Woo,J.S. (2015) Functional anatomy of the human Microprocessor. *Cell*, **161**, 1374–1387.
28. Li,X., Liu,C.X., Xue,W., Zhang,Y., Jiang,S., Yin,Q.F., Wei,J., Yao,R.W., Yang,L. and Chen,L.L. (2017) Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. *Mol. Cell*, **67**, 214–227.
29. Zheng,B., Zhou,J. and Wang,H. (2020) Host microRNAs and exosomes that modulate influenza virus infection. *Virus Res.*, **279**, 197885.
30. Khongnomnan,K., Makkoch,J., Poomipak,W., Poovorawan,Y. and Payungporn,S. (2015) Human miR-3145 inhibits influenza A viruses replication by targeting and silencing viral PB1 gene. *Exp. Biol. Med. (Maywood)*, **240**, 1630–1639.
31. Chen,J.S., Su,I.J., Leu,Y.W., Young,K.C. and Sun,H.S. (2012) Expression of T-cell lymphoma invasion and metastasis 2 (TIAM2) promotes proliferation and invasion of liver cancer. *Int. J. Cancer*, **130**, 1302–1313.
32. Yen,W.H., Ke,W.S., Hung,J.J., Chen,T.M., Chen,J.S. and Sun,H.S. (2016) Sp1-mediated ectopic expression of T-cell lymphoma invasion and metastasis 2 in hepatocellular carcinoma. *Cancer Med.*, **5**, 465–477.
33. Landgraf,P., Rusu,M., Sheridan,R., Sewer,A., Iovino,N., Aravin,A., Pfeffer,S., Rice,A., Kamphorst,A.O., Landthaler,M. *et al.* (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, **129**, 1401–1414.
34. Gwizdek,C., Ossareh-Nazari,B., Brownawell,A.M., Evers,S., Macara,I.G. and Dargemont,C. (2004) Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3. *J. Biol. Chem.*, **279**, 884–891.
35. Choudhury,N.R., de Lima Alves,F., de Andres-Aguayo,L., Graf,T., Caceres,J.F., Rappsilber,J. and Michlewski,G. (2013) Tissue-specific control of brain-enriched miR-7 biogenesis. *Genes Dev.*, **27**, 24–38.
36. Tsai,V.W.W., Husaini,Y., Sainsbury,A., Brown,D.A. and Breit,S.N. (2018) The MIC-1/GDF15-GFRAL pathway in energy homeostasis: implications for obesity, cachexia, and other associated diseases. *Cell Metab.*, **28**, 353–368.

Downloaded from https://academic.oup.com/nar/article-abstract/48/12/6874/5840576 by UPR 1142 - IGH/ Bibliothèque IFR3 user on 16 July 2020

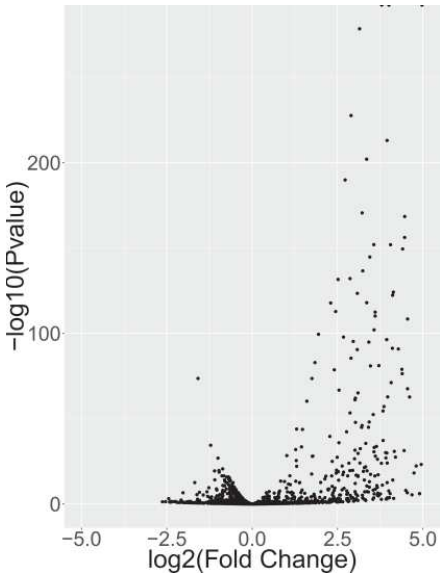
A



C

miRNAs	Upregulated	Downregulated
Mature	278	84
Precursors	217	77

B



D

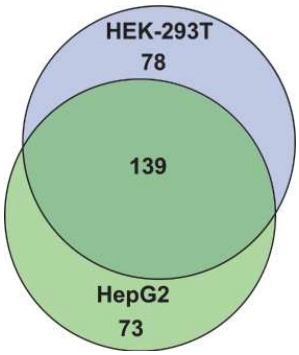


Figure S1. NF90 Modulates the Expression Level of miRNAs in HEK-293T cells. **(A)** Extracts of HEK-293T cells transfected with siRNA targeting NF90 or a non-targeting control (Scr) were analyzed by Western blot using the indicated antibodies. **(B)** Samples described in A were analyzed by small RNA-seq. Results are shown as log2 fold change versus –log10 p-value. **(C)** Table summarizing the number of mature miRNAs and pri-miRNAs modulated in HEK-293T cells upon loss of NF90, according to small-RNA seq. **(D)** Venn diagram representing the number of miRNAs upregulated following knock-down of NF90 in HepG2 versus HEK-293T cells.

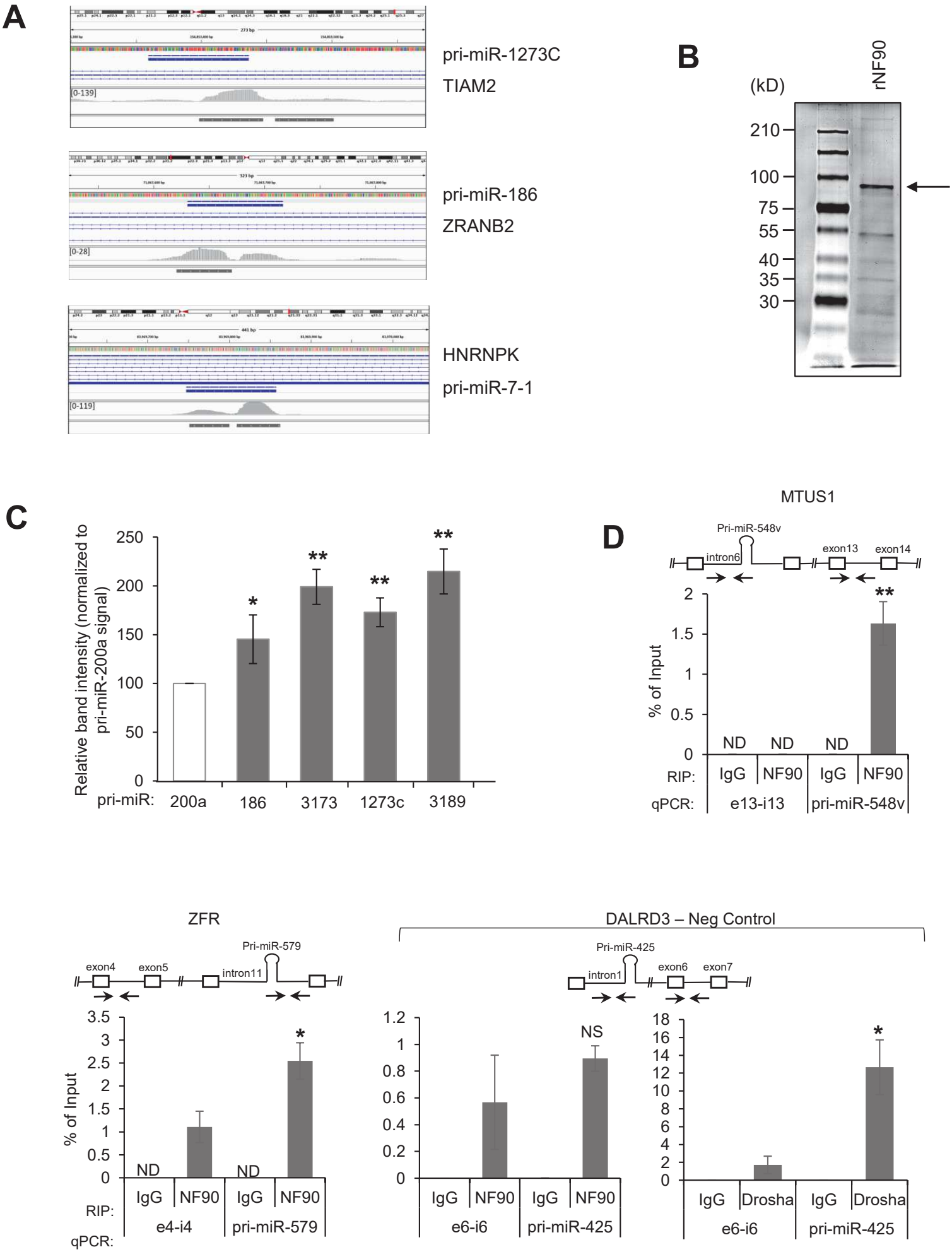

A
pri-miR-1273C
TIAM2
[0-139]
pri-miR-186
ZRANB2
[0-28]
HNRNPK
pri-miR-7-1
[0-119]
B
rNF90
(kD)
210
100
75
55
40
35
30
C
Relative band intensity (normalized to pri-miR-200a signal)
250
200
150
100
50
0
pri-miR:
200a
186
3173
1273c
3189
D
MTUS1
Pri-miR-548v
intron6
exon13
exon14
% of Input
2
1.5
1
0.5
0
ND
ND
ND
RIP:
IgG
NF90
IgG
NF90
qPCR:
e13-i13
pri-miR-548v
ZFR
Pri-miR-579
exon4
exon5
intron11
% of Input
3.5
3
2.5
2
1.5
1
0.5
0
ND
ND
RIP:
IgG
NF90
IgG
NF90
qPCR:
e4-i4
pri-miR-579
DALRD3 – Neg Control
Pri-miR-425
intron1
exon6
exon7
1.2
1
0.8
0.6
0.4
0.2
0
NS
IgG
NF90
IgG
NF90
e6-i6
pri-miR-425
18
16
14
12
10
8
6
4
2
0
IgG
Drosha
IgG
Drosha
e6-i6
pri-miR-425

Figure S2. (A) Browser shots of NF90 eCLIP read coverage over the pri-miRNAs indicated on the figure. Blue lines represent host gene showing localization of the pri-miRNA. eCLIP reads are shown in grey and locations of eCLIP peaks are shown as dark grey bars. The arrows indicate the strand from which the reads originated. **(B)** Recombinant NF90 used in RNA EMSA was analyzed by SDS-PAGE and Comassie brilliant blue staining. **(C)** RNA EMSA shown in Figure 2C performed using recombinant NF90 and probed with radiolabeled pri-miRNAs was carried out in three independent experiments. The graph shows the mean ± SD of the relative band intensity normalized to pri-miR-200a signal *(*P < 0.05 **P < 0.005, independent Student's t test).*

(D) HepG2 cells transfected with siRNA targeting NF90 or a non-targeting control (Scr) were subjected to RIP using anti-NF90, anti-Drosha or a control antibody. Immunoprecipitates were analyzed by RT-qPCR. ND indicates 'Not Detected'. Data represent mean ± SEM obtained from 3 independent experiments *(*P < 0.05, **P < 0.01, independent Student's t test)*.

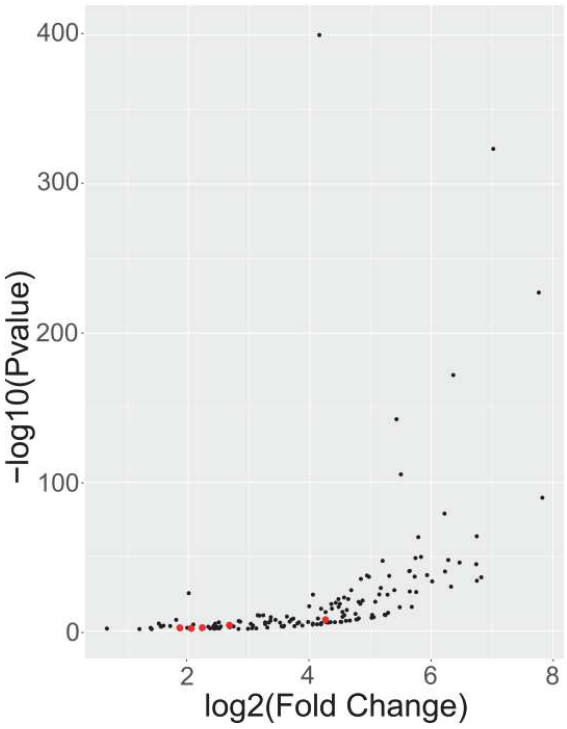


Figure S3. DotPlot of Drosha-associated pri-miRNAs, determined by eCLIP analysis. Red dots indicate the position of pri-miRNAs that are also positive for association with NF90.

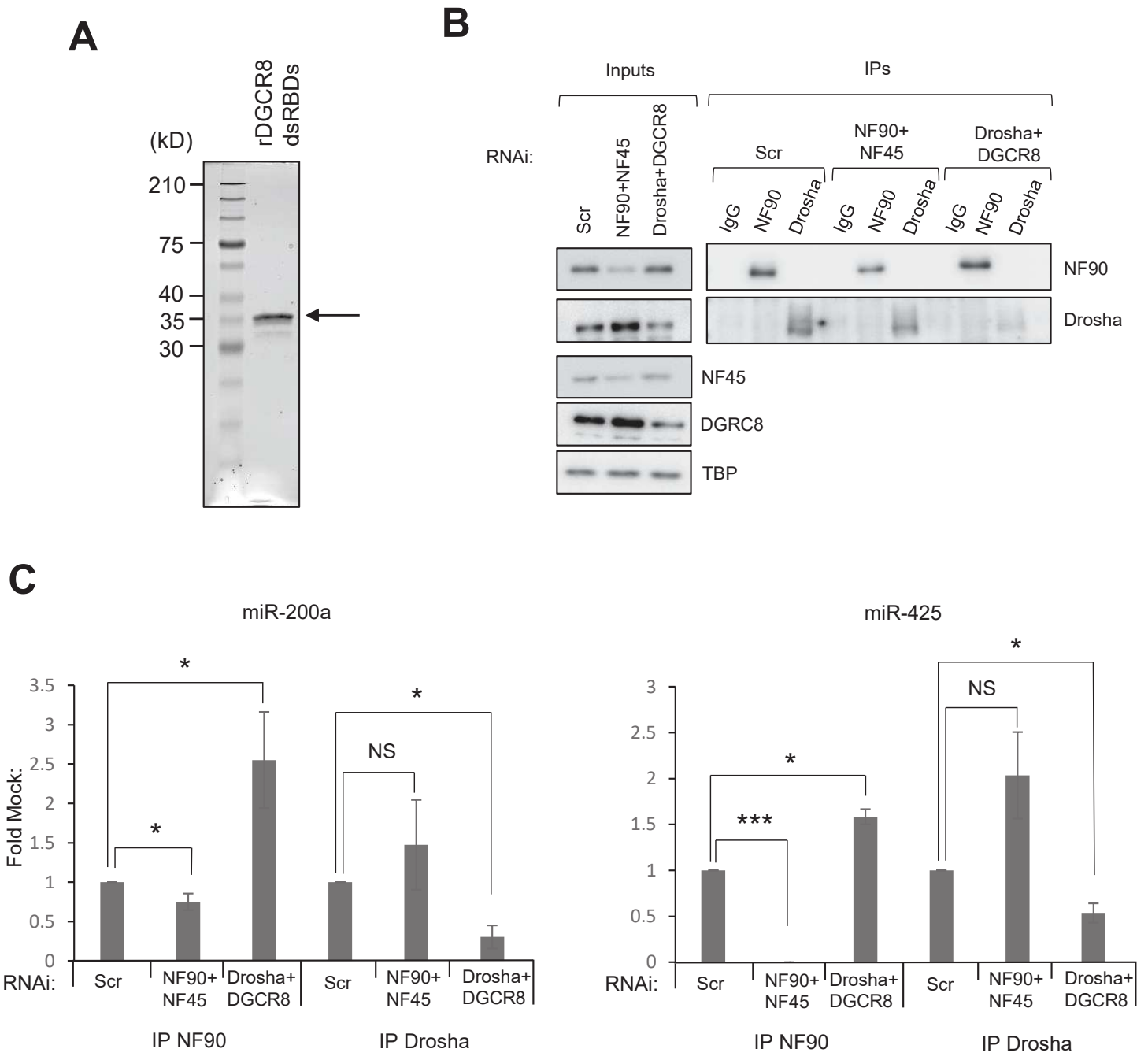


Figure S4. (A) Recombinant DGCR8 dsRBDs used in RNA EMSA was analyzed by SDS-PAGE and Comassie brilliant blue staining. **(B)** Extracts of HepG2 cells transfected with siRNAs targeting NF90 and NF45, Drosha and DGCR8 or a non-targeting control (Scr) were analyzed by Western blot using the indicated antibodies (left panel). The same extracts were used of RIP using antibodies anti-NF90, anti-Drosha or control IgG, as indicated (right panel). **(C)** HepG2 cells transfected with siRNA targeting NF90 and NF45 or Drosha and DGCR8 or a non-targeting control (Scr), as indicated, were subjected to RIP using anti-NF90, anti-Drosha or a control IgG antibody. Immunoprecipitates were analyzed by RT-qPCR. NS indicates 'Not Significant'. Data represent Fold mock (IgG) relative to the control sample (siScr), which was attributed a value of 1, obtained from 3 independent experiments *(*P < 0.05, ***P < 0.001, independent Student's t test)*.

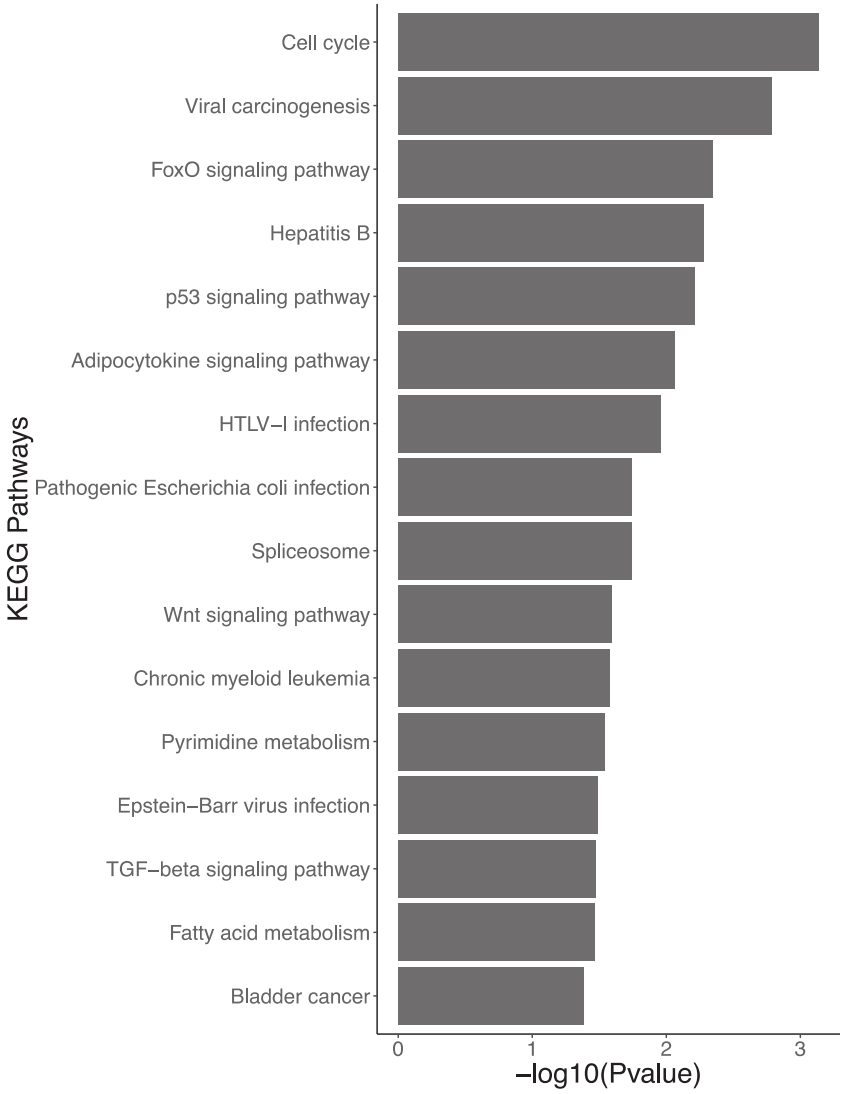


Figure S5. NF90 double positive pri-miRNAs target genes involved in viral infection and cancer. Gene ontology of validated targets of NF90-bound and upregulated miRNAs.

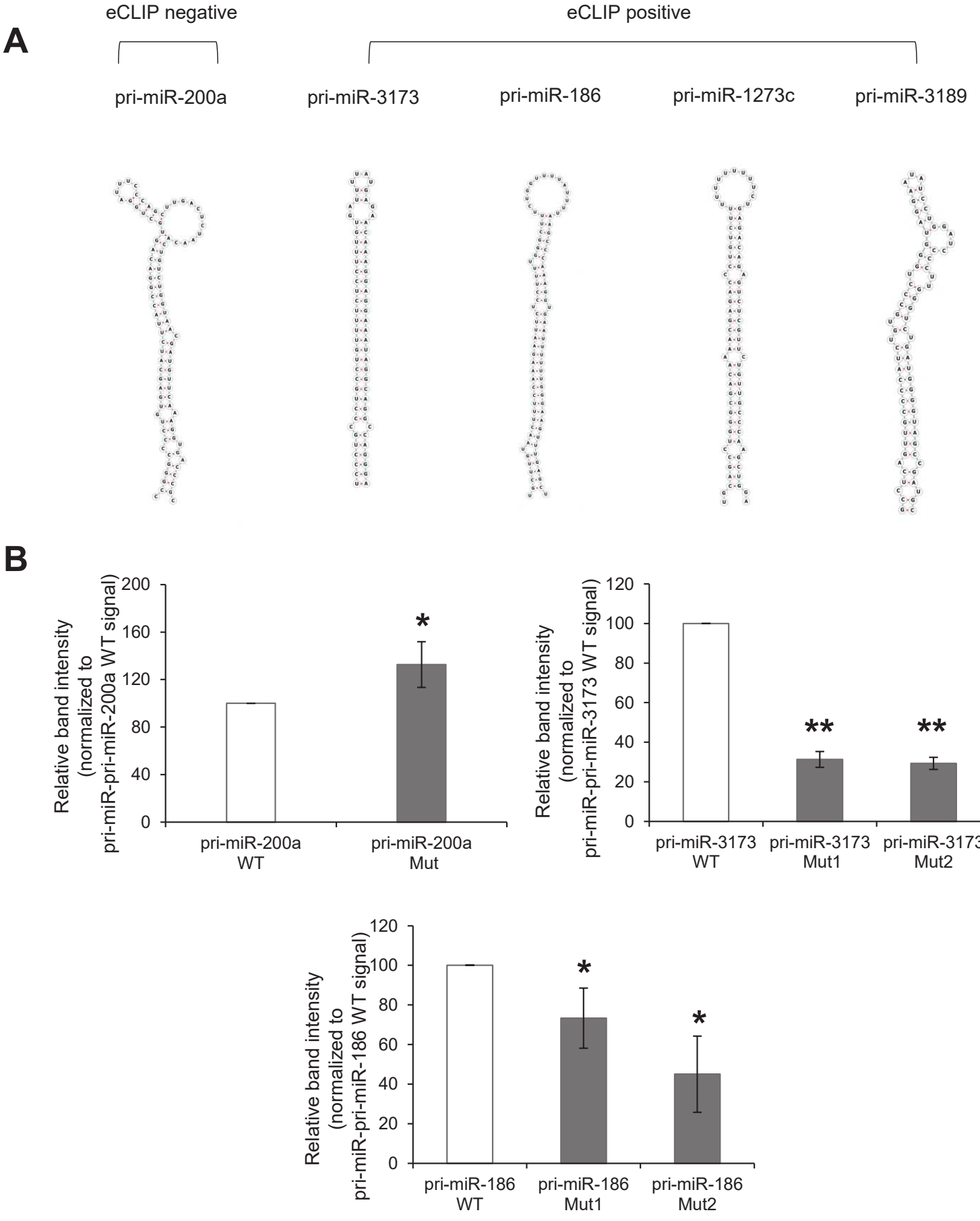


Figure S6. (A) NF90-associated pri-miRNAs are highly stable. Predicted folding of pri-miRNAs that are significantly associated or not with NF90 as indicated. RNA structures were predicted using FORNA. **(B)** RNA EMSA shown in Figure 5B performed using recombinant NF90 and probed with radiolabeled WT or mutant pri-miRNAs was carried out in three independent experiments. The graph shows the mean ± SD of the relative band intensity normalized to the WT signal *(*P < 0.05 **P < 0.005, independent Student's t test).*

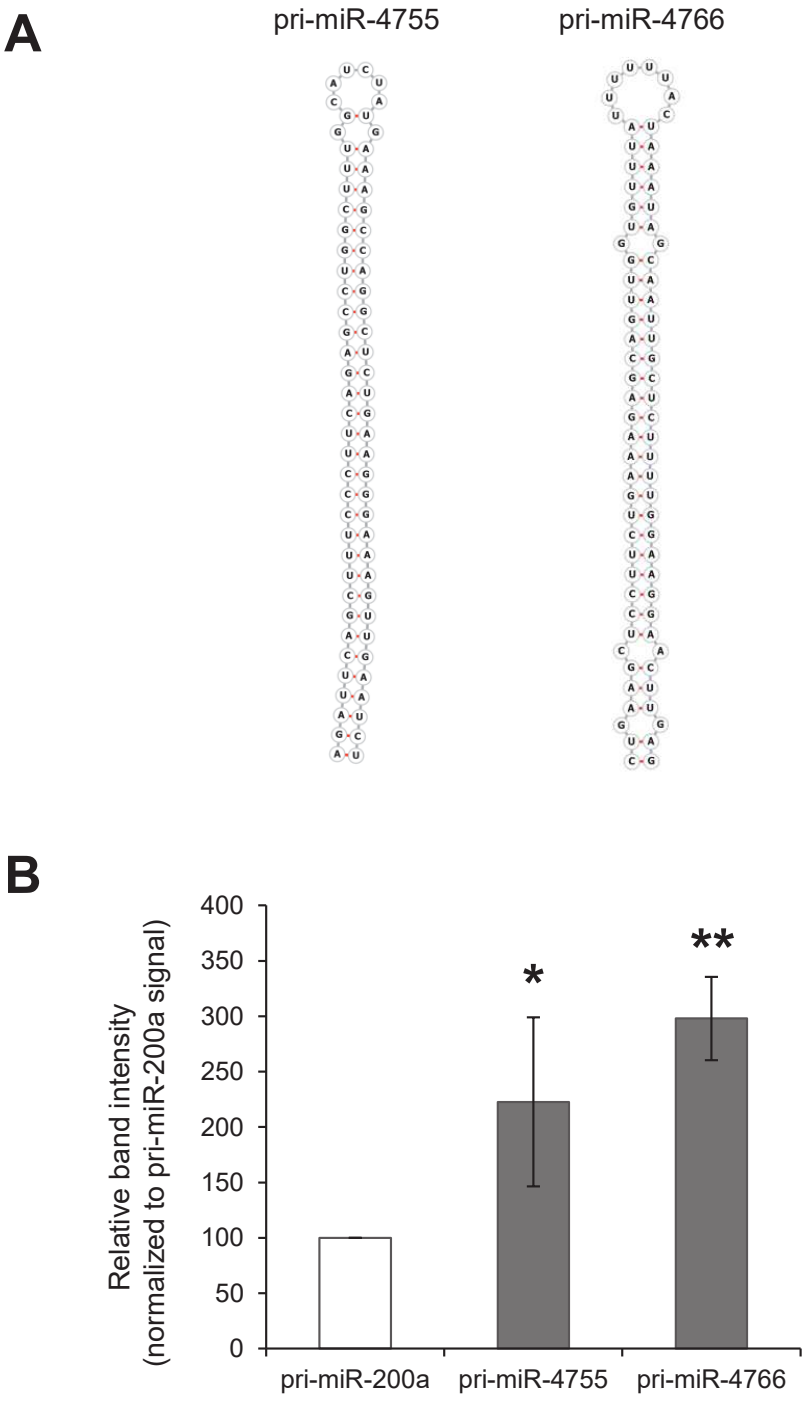


Figure S7. (A) NF90-modulated pri-miRNAs are highly stable. Predicted folding of examples of pri-miRNAs whose mature products are upregulated following loss of NF90. RNA structures were predicted using FORNA. **(B)** RNA EMSA shown in Figure 6C performed using recombinant NF90 and probed with radiolabeled pri-miR-200a, pri-miR-4755 and pri-miR-4766 was carried out in three independent experiments. The graph shows the mean ± SD of the relative band intensity normalized to pri-miR-200a *(*P < 0.05, **P < 0.05, independent Student's t test).*

Supplementary Table S1. Double stranded siRNAs used in this study.

siRNA	Sequence (5' to 3')
Scr	gcgcgcuuuguaggauucg(dTdT)
Scr#2	ucugcaagguuaggcgucu(dTdT)
NF90	ccaaggaacucuaucacaa(dTdT)
NF90#2	gaguugaaguauugauaac(dTdT)
Drosha	cgaguaggcuucgugacuu(dTdT)
DGCR8	caucggacaagagugugau(dTdT)
NF45	guggugauacucaagauucugccaa(dTdT)

Supplementary Table S2. Primary antibodies used in this study.

Antibody	Reference	Supplier
NF90	A303-651A	Bethyl Laboratories
GDF-15	sc-377195	SCBT
ZRANB2	sc-514200	SCBT
TIAM2	sc-514090	SCBT
NDUFS8	sc-515527	SCBT
TUBULIN	DM1A clone, T6199	Sigma-Aldrich
TBP	sc-421	SCBT
Drosha (IP)	Ab-12286	Abcam
Drosha (WB)	sc-33778	SCBT
DGCR8	Ab-82876	Abcam
Rabbit IgG	p120-101	Bethyl Laboratories
DICER1	sc-136981	SCBT
ZFR	A14281	ABClonal
NF45	A303-147A	Bethyl Laboratories

Supplementary Table S3. Primers used in this study.

Primer	Forward (5' to 3')	Reverse (5' to 3')
Spiced GAPDH	cac atc gct cag aca cca t	gag gtc aat gaa ggg gtc at
U6	ctc gct tcg gca gca cat ata c	gga acg ctt cac gaa ttt gcg tg
pri-miR-1273C (EMSA)	ctt ggg aag ctg agg tag gc	act tgg tac tga ggc gga gg
pri-miR-186 (EMSA)	aca gaa cac cca tca tat tc	gtt gac att cac atg ctt c
pri-miR-200a (EMSA)	ctg gct gct cac cgc tcc	gat gtg cct cgg tgg tgt cc
pri-miR-3173 (EMSA)	cat tgg agg tct agg gct ta	gtt ctt cct cgg cac aag
pri-miR-3189 (EMSA)	agc agc ccc cat atc taa tc	ctg gca tcc ctg tac ctc
pri-miR-4755 (EMSA)	aga gat gag gaa ggt tat ggc t	tgg ccc aaa cct cat aga c
pri-miR-4766 (EMSA)	ccc ttc tac ctt tct gaa gct c	cac aca ggt ggc act caa c
5'RLM-RACE pri-miR-186 (outer)	gct gat ggc gat gaa tga aca ctg (adapter)	aaa cca ggt ata tgg cac agc aac
5'RLM-RACE pri-miR-186 (inner)	cgc gga tcc gaa cac tgc gtt tgc tgg ctt tga tg (adapter)	tgt tga cat tca cat gct tca ggt
5'RLM-RACE pri-miR-3189 (outer)	gct gat ggc gat gaa tga aca ctg (adapter)	acc aca ccc cca ttg ttt ctct
5'RLM-RACE pri-miR-3189 (inner)	cgc gga tcc gaa cac tgc gtt tgc tgg ctt tga tg (adapter)	acc aca ccc cca ttg ttt ctct
ZRANB2 e3-i3 (RIP)	gag ccg agg cct att tag tg	aag gtt acc ctg gct tgt ca
ZRANB2 pri-miR-186 (RIP)	cct gaa gca tgt gaa tgt caa	cca ggt ata tgg cac agc aa
DICER i22-e23 (RIP)	ggc cat gat ttt aaa gtt gc	tcc tcc tcc tcg taa tcc tc
DICER pri-miR-3173 (RIP)	aac aga acc tgg aca ctg ag	aga cac caa cct cac tca ag
TIAM2 i12-e13 (RIP)	ggt ttg agt ttg cag cct tc	aga aaa cag ggc ctc cat ct
TIAM2 pri-miR-1273c (RIP)	ctg aaa tgc tgt ccc cat ct	tgc cca gtc tct tct cgt tt
GDF-15 i2-e3 (RIP)	ctc cca aag tgc tgg gat ta	aga gat acg cag gtg cag gt
GDF-15 pri-miR-3189 (RIP)	acg cta cga gga cct gct aa	tta gat atg ggg gct gct tg
MTUS1 e13-i13 (RIP)	aga aag cct gaa agc tgt gtt	aat gca ggg ctc aat ttc ac
MTUS1 pri-miR-548v (RIP)	tct cag cgt ggc tac tag gaa	tgt acg gct aca gca tct gg
ZFR e4-i4 (RIP)	tca gcc ttc tgt tgc tga aa	ggg ctg aaa gtc cag aaa tg
ZFR pri-miR579 (RIP)	ttt tgt gtc tgg cat cgt tc	gga aac aag ttg cat gtc ca
DALRD3 e6-i6 (RIP)	atc tgt ggc cct gtg aaa gt	gta tgc cgg aac ctg tgt tt
DALRD3 pri-miR-425 (RIP)	agg gct gca atg gta gtg ac	aag gtg cat gac ctg gag ac
NDUFS8 miRNA unspliced	ctt agc cgg agt cca gga g	aag ccg cag tag atg cac tt
NDUFS8 miRNA spliced	cca cca tca act acc cgt tc	aag ccg cag tag atg cac tt
NDUFS8 unspliced	caa tgg cag cgt cct aca gt	caa cgg gga cac cac act
NDUFS8 spliced	atg gca gcg tcc tac agt g	agc agc ata ggc gtg gtg
TIAM2 miRNA unspliced	aac caa ggt ttt gcg tga ag	acc cag gta gct gaa gac ga
TIAM2 miRNA spliced	gct cag cca cca cct ata cc	acc cag gta gct gaa gac ga
TIAM2 unspliced	gga tgc ttt gga tag ccg ta	cga atg tgt gga ttc act tc
TIAM2 spliced	atc agt gac tgg acg gga ag	tcg cat gtg tgg att cac tt
ZFR miRNA unspliced	gat gca agt tct tgg gct gt	cct gga gga cca tga gga ta
ZFR miRNA spliced	gga gta ctg gcg aag acg ag	cct gga gga cca tga gga ta
ZFR unspliced	tcc ttc cac ttg tct cat agc a	ttt cag caa cag aag gct ga
ZFR spliced	tta tgg agg cta ccc cac tg	cag cag ttg ctg ttg gtt gt
DICER1 miRNA unspliced	gga aga gtt tga atg gct ca	ggg ctt ttc att cat cca gtg
DICER1 miRNA spliced	gtc cga tgg ttc tcg aag	gtt cta gca cag ctt act g
DICER1 unspliced	aat tgc att ctc act act gca	gtc cac aat cca cca caa tc
DICER1 spliced	att gtc cat cat gtc ctc gc	gtc cac aat cca cca caa tc

Supplementary Table S4. Wild-type and mutant pri-miRNAs sequences used for RNA EMSA. The pri-miRNA sequence is shown in red and flanking sequence is shown in black.

	pri-miR-200a	pri-miR-3173	pri-miR-186
Wild-type	ctggctgctcaccgctccggttcttccctgggct tccacagcagcccctgcctgcctggcgggac cccacgtccctcccgggcccctgtgagcatct taccggacagtgctggatttcccagcttgactc taacactgtctggtaacgatgttcaaaggtga cccgccgctcgccggggacaccaccgagg cacatc	cattggaggtctagggcttattttccagat agaattgagtctttgttggtcttgggccag cttccctgccctgcctgttttctcctttgtgatt ttatgagaacaaaggaggaaataggca ggccagggaaacgatctctctccctctctt gtgccgaggaagaact	acagaacacccatcatattcttcccaaacatttttttcat tgcttgtaactttccaaagaattctccttttgggctttctg gttttattttaagcccaaaggtgaatttttttgggaagttt gagctaaattccttcaaccaaaatatacaagtgaag aaaaaaaatttgtatttaaacatttgcacatttacttct acctgaagcatgtgaatgtcaac
Mutant #1	ctggctgctcaccgctccggttcttccctgggct tccacagcagcccctgcctgcctggcgggac cccacgtccctcccgggcccctgtgagcatct taccggacagtgctggatttcccagcctgtctg gtaacgatgttcaaaggtgacccgccgctcg ccggggacaccaccgaggcacatc	cattggaggtctagggcttattttccagat agaattgagtctttgttggtcttgggccag cttccctgcgacgcctgcctgttttctccttt gtgattttatgagaacaaaggaggaaag cgctaggcaggccagggaaacgatctc tctccctctcttgtgccgaggaagaact	acagaacacccatcatattcttcccaaacatttttttcat tgcttgtaactttccaactaaagaattctccttttgggct ttctggttttattttaagcccacctctaaggtgaatttttttg ggaagtttgagctaaattccttcaaccaaaatataca agtgaagaaaaaaaaatttgtatttaaacatttgcaca tttacttctacctgaagcatgtgaatgtcaac
Mutant #2		cattggaggtctagggcttattttccagat agaattgagtctttgttggtcttgggccag cttccctgcaagtcctgttttctcctttgtgatt ttatgagaacaaaggagaccctaggca ggccagggaaacgatctctctccctctctt gtgccgaggaagaact	acagaacacccatcatattcttcccaaacatttttttcat tgcttgtccagttccaaagaattctccttttgggctttct ggttttattttaagcccaaaggtgaaccacgtgggaa gtttgagctaaattccttcaaccaaaatatacaagtg aagaaaaaaaatttgtatttaaacatttgcacatttac ttctacctgaagcatgtgaatgtcaac

Supplementary Table S5. NF90-associated pri-miRNAs, as determined by eCLIP analysis.

hsa-mir-1273c	hsa-mir-4485	hsa-mir-548d-1
hsa-mir-1290	hsa-mir-4635	hsa-mir-548u
hsa-mir-15b	hsa-mir-4659a	hsa-mir-548v
hsa-mir-186	hsa-mir-4687	hsa-mir-5581
hsa-mir-1914	hsa-mir-4712	hsa-mir-570
hsa-mir-3140	hsa-mir-4714	hsa-mir-578
hsa-mir-3145	hsa-mir-4730	hsa-mir-579
hsa-mir-3173	hsa-mir-4762	hsa-mir-606
hsa-mir-3189	hsa-mir-4775	hsa-mir-624
hsa-mir-3646	hsa-mir-4779	hsa-mir-6751
hsa-mir-3648-1	hsa-mir-4782	hsa-mir-6839
hsa-mir-3680-1	hsa-mir-548aq	hsa-mir-7-1
hsa-mir-3939	hsa-mir-548ar	

Supplementary Table S6. 'Double positive' miRNAs whose abundance increased following loss of NF90 and that were positive for NF90 association by eCLIP, and their host gene.

miRNA	Small RNA-seq		Host Gene
	Fold Change (log2)	p value (-Log10)	
hsa-mir-1273c	3.55	33.94	TIAM2
hsa-mir-186	1.01	4.55	ZRANB2
hsa-mir-3140	3.77	37.16	FBXW7
hsa-mir-3145	2.56	9.78	NHSL1
hsa-mir-3173	3.03	25.36	DICER1
hsa-mir-3189	2.16	19.36	GDF15
hsa-mir-3646	3.76	4.88	HNF4A
hsa-mir-3939	1.27	7.59	RP1-167A14.2
hsa-mir-4659a	2.21	10.93	AGPAT5
hsa-mir-4714	3.36	23.31	IGF1R
hsa-mir-4762	2.14	5.66	ATXN10
hsa-mir-4775	1.37	3.72	CCNYL1
hsa-mir-4779	4.03	10.9	IMMT
hsa-mir-4782	3.94	3.87	SLC35F5
hsa-mir-548ar	3.24	6.79	CDC16
hsa-mir-548u	2.24	2.38	PRIM2
hsa-mir-548v	2.01	7.81	MTUS1
hsa-mir-5581	2.69	19.23	MEAF6
hsa-mir-578	2.17	3.18	CPE
hsa-mir-579	4	73.29	ZFR
hsa-mir-624	2.68	25.49	STRN3
hsa-mir-7-1	0.99	5.76	HNRNPK

Supplementary Table S7. MiRNAs downregulated in abundance following loss of NF90 and that are associated with NF90 by eCLIP.

miRNA	**Fold Change (log2)**	**p value (-Log10)**	**Host Gene**
hsa-mir-1914	-0.86	2.88	UCKL1
hsa-mir-6751	-1.52	3.73	SYVN1

Appendix B

Chromatin-associated MRN complex protects highly transcribing genes from genomic instability

MOLECULAR BIOLOGY

Chromatin-associated MRN complex protects highly transcribing genes from genomic instability

Kader Salifou[1†], Callum Burnard[1,2†], Poornima Basavarajaiah[1†], Giuseppa Grasso[1], Marion Helsmoortel[1], Victor Mac[1], David Depierre[2], Céline Franckhauser[1], Emmanuelle Beyne[1], Xavier Contreras[1], Jérôme Dejardin[3], Sylvie Rouquier[1], Olivier Cuvier[2], Rosemary Kiernan[1]*

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

MRN-MDC1 plays a central role in the DNA damage response (DDR) and repair. Using proteomics of isolated chromatin fragments, we identified DDR factors, such as MDC1, among those highly associating with a genomic locus upon transcriptional activation. Purification of MDC1 in the absence of exogenous DNA damage revealed interactions with factors involved in gene expression and RNA processing, in addition to DDR factors. ChIP-seq showed that MRN subunits, MRE11 and NBS1, colocalized throughout the genome, notably at TSSs and bodies of actively transcribing genes, which was dependent on the RNAPII transcriptional complex rather than transcription per se. Depletion of MRN increased RNAPII abundance at MRE11/NBS1-bound genes. Prolonged MRE11 or NBS1 depletion induced single-nucleotide polymorphisms across actively transcribing MRN target genes. These data suggest that association of MRN with the transcriptional machinery constitutively scans active genes for transcription-induced DNA damage to preserve the integrity of the coding genome.

INTRODUCTION

Execution of the appropriate transcriptional program is fundamental during growth, development, and response to environmental stimuli. Although an essential DNA-dependent process, transcription comes at a cost. It can induce double-strand breaks (DSBs) that are a deleterious form of DNA damage and can compromise the genome if erroneously repaired. DSBs are predominantly repaired by two canonical pathways: nonhomologous end joining (NHEJ) or homologous recombination (HR). HR involves extensive DNA end resection and uses an intact copy of the damaged locus to produce error-free repair. NHEJ ligates broken DSB ends with no or limited processing and frequently results in genomic mutations. Both pathways compete to repair a DSB (*1*), and mechanisms underlying pathway choice are not entirely clear. The nuclease activities of MRE11, a subunit of the MRE11-Rad50-NBS1 complex (MRN), are a critical determinant of pathway choice (*2*). MRE11 endonuclease activity initiates resection, thereby licensing HR. The exonuclease activities of MRE11 and EXO1/BLM (exonuclease 1/BLM RecQ like helicase) bidirectionally resect toward and away from the DNA end, which commits to HR and disfavors NHEJ.

While the complexity of the DNA break is an important factor in DNA damage response (DDR) pathway choice, recent studies have also identified the local nuclear environment and epigenetic landscape as determinants of repair pathway usage, predisposing to either NHEJ or HR pathways (*3*, *4*). At actively transcribed regions, H3K36me3 is associated with a preference for DNA repair through the HR pathway (*5*–*7*). Nuclear organization has emerged as a key parameter in the formation of chromosomal translocations, implying that DSBs cluster into higher-order structures. Microscopy-based techniques and more recent chromosome conformation capture assays have shown that DSBs are mobile and form higher-order structures detected as clusters (*8*–*11*). Initially thought to be factories for DNA repair, recent data have suggested that clustering may protect DSBs from spurious pairing to suppress translocations. Thus, DSBs in actively transcribed chromatin are not repaired in G_1 but remain in clusters until HR can proceed in G_2 (*9*). DSB clustering depends on several DNA repair factors including the MRN complex (*8*, *9*, *11*). Thus, a model emerges in which DSBs in active chromatin are sequestered in higher-order structures that permit the HR repair pathway to proceed under conditions that minimize the formation of deleterious translocations. However, the molecular mechanism by which HR is chosen as a predominant pathway for DSB repair in transcriptionally active regions remains unclear. Furthermore, most studies have been performed by inducing DNA damage using irradiation or chemicals or artificially cutting the genome. Mechanisms involved in endogenous transcription-associated DNA repair remain poorly understood.

To address the recruitment of DNA damage repair (DDR) factors at transcriptionally active regions, we used an unbiased proteomic approach, proteomics of isolated chromatin fragments (PICh) (*12*), of the inducibly transcribed HIV-1 promoter. Upon transcriptional activation, DNA repair factors were among the most recruited interactants. In particular, mediator of DNA damage checkpoint (MDC1) that associates with the MRN complex was recruited specifically following activation of transcription. Identification of the interactome of endogenous MDC1 by tandem affinity purification followed by mass spectrometry similarly identified numerous factors implicated in transcription and cotranscriptional RNA processing. Coimmunoprecipitation analysis showed that MDC1 interacts with RNA polymerase II (RNAPII) and coactivators, such as P-TEFb. ChIP-seq (chromatin immunoprecipitation followed by high-throughput sequencing) analysis of MRN subunits, MRE11 and NBS1, revealed an association with genes, specifically those that were transcriptionally active. MRE11 and NBS1 demonstrated covariation with RNAPII upon heat shock–induced transcriptional activation. Blockade of transcriptional elongation altered the profile of NBS1 to mirror that of RNAPII, suggesting that MRN is recruited through its association with the transcriptional

[1]CNRS-UMR 9002, Institute of Human Genetics (IGH)/University of Montpellier, Gene Regulation Lab, 34396, France. [2]Center of Integrative Biology (CBI) CNRS/Université de Toulouse (UMR 5077), 31000, France. [3]CNRS-UMR 9002, IGH/University of Montpellier, Biology of Repeated Sequences Lab, 34396, France.
*Corresponding author. Email: rosemary.kiernan@igh.cnrs.fr
†These authors contributed equally to this work.

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

machinery rather than the level of transcription per se. Following MRE11 or NBS1 depletion, the occurrence of single-nucleotide polymorphisms (SNPs) was highly correlated to the binding of MRE11, NBS1, or RNAPII. These data suggest that the MRN complex, through its interaction with the transcription machinery, scans active genes for transcription-induced DNA damage.

RESULTS

DDR factors are associated with active transcription

We investigated the recruitment of DDR factors during transcription using an unbiased proteomic method, PICh (*12*) of the HIV-1 promoter region, which becomes highly transcribed upon addition of its cognate transactivator, Tat. PICh was performed using chromatin extracts of cells containing tandem integrated copies of the HIV-1 long terminal repeat (LTR) linked to MS2 binding sites (*13*), in the presence or absence of the transactivator protein, Tat (fig. S1, A and B). Mass spectrometry identified 438 proteins associated with the locus in the control condition and 622 proteins associated with the locus upon transactivation by Tat. Upon transcriptional activation, 58 proteins were highly gained [fold change (FC), >7], 294 were moderately gained (FC, >2 and <7), 29 were moderately lost (FC, <−2 and >−7), and 3 were highly lost (FC, <−7) (Fig. 1A and table S1). Gene ontology revealed a number of pathways that were enriched in cells expressing Tat, compared to control untreated cells (fig. S1C). As expected, factors involved in gene expression were among the most highly gained. A number of factors involved in DNA repair were associated with HIV-1 chromatin specifically upon transactivation of transcription with Tat. Factors implicated in both the NHEJ and HR pathways were identified (table S1). However, NHEJ factors, such as DNA-dependent protein kinase (DNA-PK) and Ku70/80 are known to be implicated in transcription (*14*, *15*), including Tat-mediated transcription of the HIV-1 LTR (*16*–*18*). Furthermore, DNA-PK has been shown to interact with HIV-1 Tat, while Ku70/80 has been shown to bind HIV-1 TAR RNA (*16*, *18*, *19*). Therefore, it is not clear whether recruitment of these factors following Tat transactivation was due to DNA damage or increased Tat recruitment and transcription. A key factor in the HR pathway, MDC1 was highly represented (Fig. 1A). To validate the interaction of DDR factors with actively transcribed HIV-1, we performed ChIP in cells containing an integrated HIV-1 LTR linked to a luciferase reporter, as described previously (*20*). Tat-induced activation of HIV-1 transcription increased recruitment of RNAPII as expected (fig. S1D). Similarly, association of MDC1, MRE11, and Ku80 were also significantly enhanced in the presence of Tat (fig. S1D), in agreement with the recruitment of these factors upon transcriptional activation identified by PICh. We furthermore determined that γH2AX, a marker of double-stranded DNA breaks, was also increased following Tat transactivation (fig. S1E), suggesting that DDR factors may be recruited following DNA damage induced by Tat-mediated transcription.

These data indicated that DDR factors such as MDC1 and MRN may be closely linked with active transcription. To further explore this, we identified the interactome of MDC1 in the absence of exogenously induced DNA damage or Tat transactivation. Using CRISPR-Cas9 genome editing technique, a Flag-hemagglutinin (HA) tag was appended to the N terminus of endogenous MDC1 protein in human embryonic kidney (HEK) 293T cells. Tandem affinity purification of Flag-HA MDC1 was performed, as described previously (Fig. 1B) (*20*–*22*). Gene ontology analysis of the interactants identified association of MDC1 with factors in the DNA damage repair pathway and the cell cycle, as expected (Fig. 1C). MDC1 interactants were also implicated in gene expression pathways (Fig. 1C and table S2). Association of endogenous MDC1 with RNA processing factors, such as SNRNP200, PRP8, HNRNPM, EFTUD2, and SF3B1, was validated by coimmunoprecipitation analysis (Fig. 1D, left) in cells that do not express HIV-1 Tat. Interaction between MDC1 and RNAPII (total and C-terminal repeat domain phosphorylated forms), as well as subunits of P-TEFb, cyclin-dependent kinase 9 (CDK9), and cyclin T1, was also detected by coimmunoprecipitation analysis (Fig. 1D, right), confirming a previous report (*23*). Moreover, immunoprecipitates of cyclin T1 contained MDC1 and MRN subunits, MRE11 and NBS1 (fig. S1F). Furthermore, MDC1 was specifically associated with the small, active P-TEFb complex, as shown by glycerol gradient analysis (fig. S1G). Together, these data show that the highly transcribing HIV-1 locus recruits DDR factors such as MDC1 and MRN and, conversely, that DDR factor MDC1 is physically associated with factors implicated in transcription and RNA processing, independently of the presence of HIV-1 Tat. These data reveal that the transcription and DDR pathways are intimately linked at the biochemical level through physical interactions.

MRN complex is associated with chromatin

Given its interaction with factors implicated in RNAPII transcription, we next sought to determine the profile of chromatin-associated MDC1/MRN complex across the coding genome. Anti-MDC1 antibodies proved unsuitable for ChIP-seq. We therefore performed ChIP-seq analysis of MRN subunits, NBS1 and MRE11, as well as RNAPII in cells in the absence of exogenous DNA damage. The average profile of signal for both MRE11 and NBS1 at active genes resembled that of RNAPII, although the amount of signal detected was lower (Fig. 2A). While the highest accumulation of MRE11 and NBS1 ChIP-seq reads occurred near the transcription start site (TSS), significant levels of binding were also detected in the gene body and downstream of the transcription end site (TES; Fig. 2A). The binding profile of MRE11 and NBS1 on representative genes, RBM17 and GNG12, is shown (Fig. 2B). Association of MDC1 with several genes, including RBM17 (RNA Binding Motif Protein 17) and GNG12 (G Protein Subunit Gamma 12), was also confirmed by ChIP–quantitative polymerase chain reaction (qPCR) (fig. S2A). More than 50,000 peaks were detected throughout the genome for either MRE11 or NBS1, and 110,000 peaks were detected for RNAPII. Although the highest signal for both MRE11 and NBS1 was found around TSSs (Fig. 2, A and B), the number of peaks overlapping TSSs represented 2.19 and 1.44%, respectively (Fig. 2C). MRE11 and NBS1 were also detected on TESs/transcription termination sites (<2%) and enhancers (3.16 and 1.74%, respectively) (Fig. 2C). The majority of MRE11 and NBS1 peaks were found on gene bodies (approximately 52%), albeit the average size of genes (>26 kb) (*24*) could account for this result because of the higher chance of overlap with the gene body compared to TSSs/TESs. However, the signal associated with peaks in gene bodies was significantly lower than that detected at TSSs (Fig. 2, B and C). Furthermore, the distribution of ChIP-seq signal for MRE11 and NBS1 along gene bodies was similar to that of RNAPII (fig. S2B). Note that peaks of MRE11 and NBS1 mapped approximately 20 base pairs (bp) upstream of RNAPII, on average (Fig. 2B and fig. S2C), which may be due to an artefact of cross-linking a large multisubunit complex during the ChIP experiment.

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

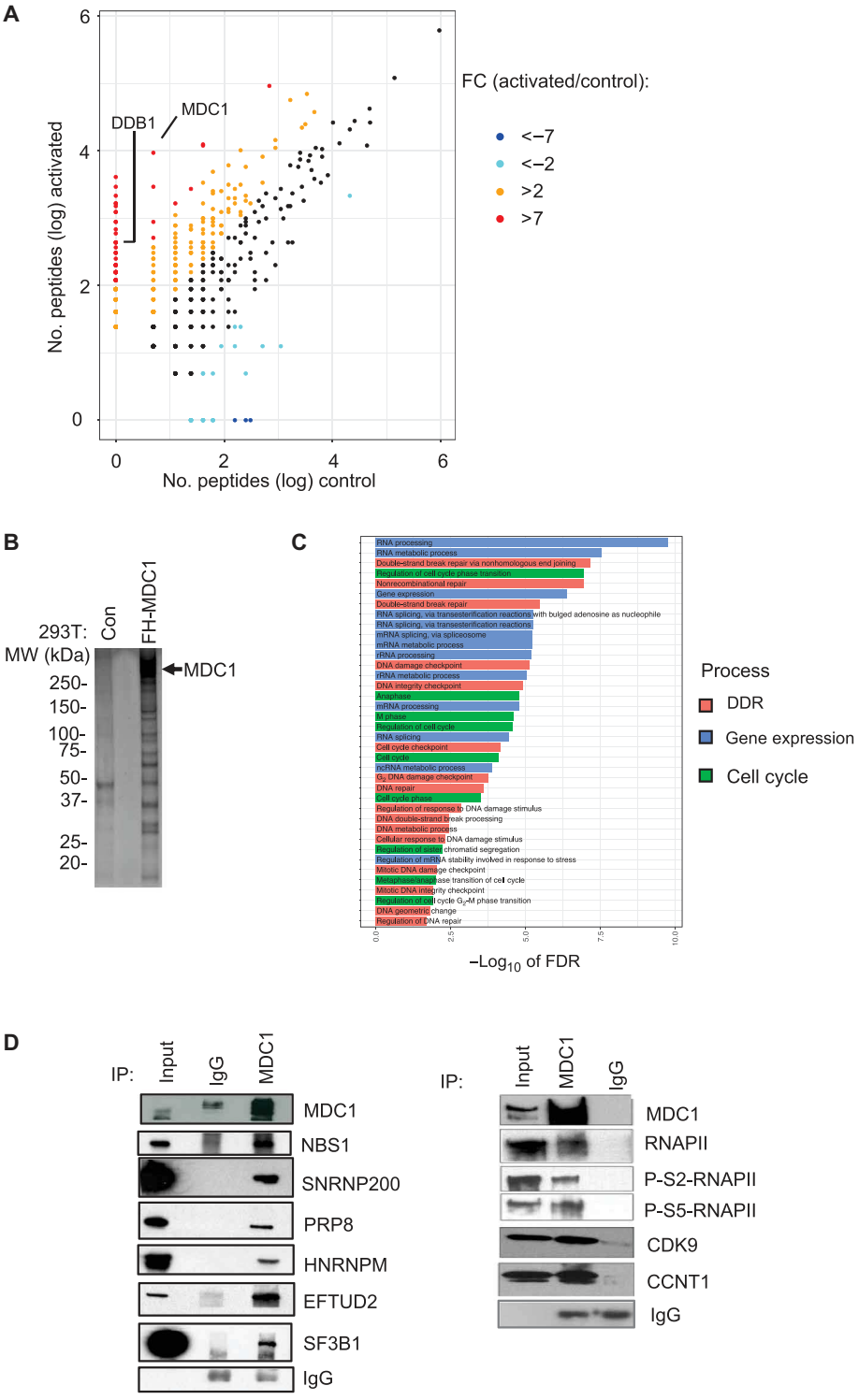


Fig. 1. DNA repair factors are associated with active transcription. (**A**) Scatter plot representation of proteins differentially associated with an integrated HIV-1 LTR-minigene upon HIV-1 promoter transactivation by Tat, from highly gained (FC, >7) to highly lost (FC, <7), as identified by PICh coupled to mass spectrometry. (**B**) SDS–polyacrylamide gel electrophoresis (SDS-PAGE) analysis followed by silver staining of eluates of tandem affinity purified nuclear extracts from HEK-293T expressing Flag-HA fused to endogenous MDC1 (FH-MDC1) or control cells (con), in the absence of HIV-1 Tat. (**C**) Gene Ontology analysis carried out submitting the interactants of FH-MDC1 identified by tandem affinity purification followed by mass spectrometry. GO terms having the highest significance are shown. FDR, false discovery rate. (**D**) Coimmunoprecipitation analysis of MDC1 performed using nuclear extracts of HEK-293T in the absence of HIV-1 Tat and the indicated antibodies. Normal immunoglobulin G (IgG) was used as a negative control. Nuclear extracts (5%) and immunoprecipitates were separated by SDS-PAGE and immunoblotted using the antibodies indicated in the figure.

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

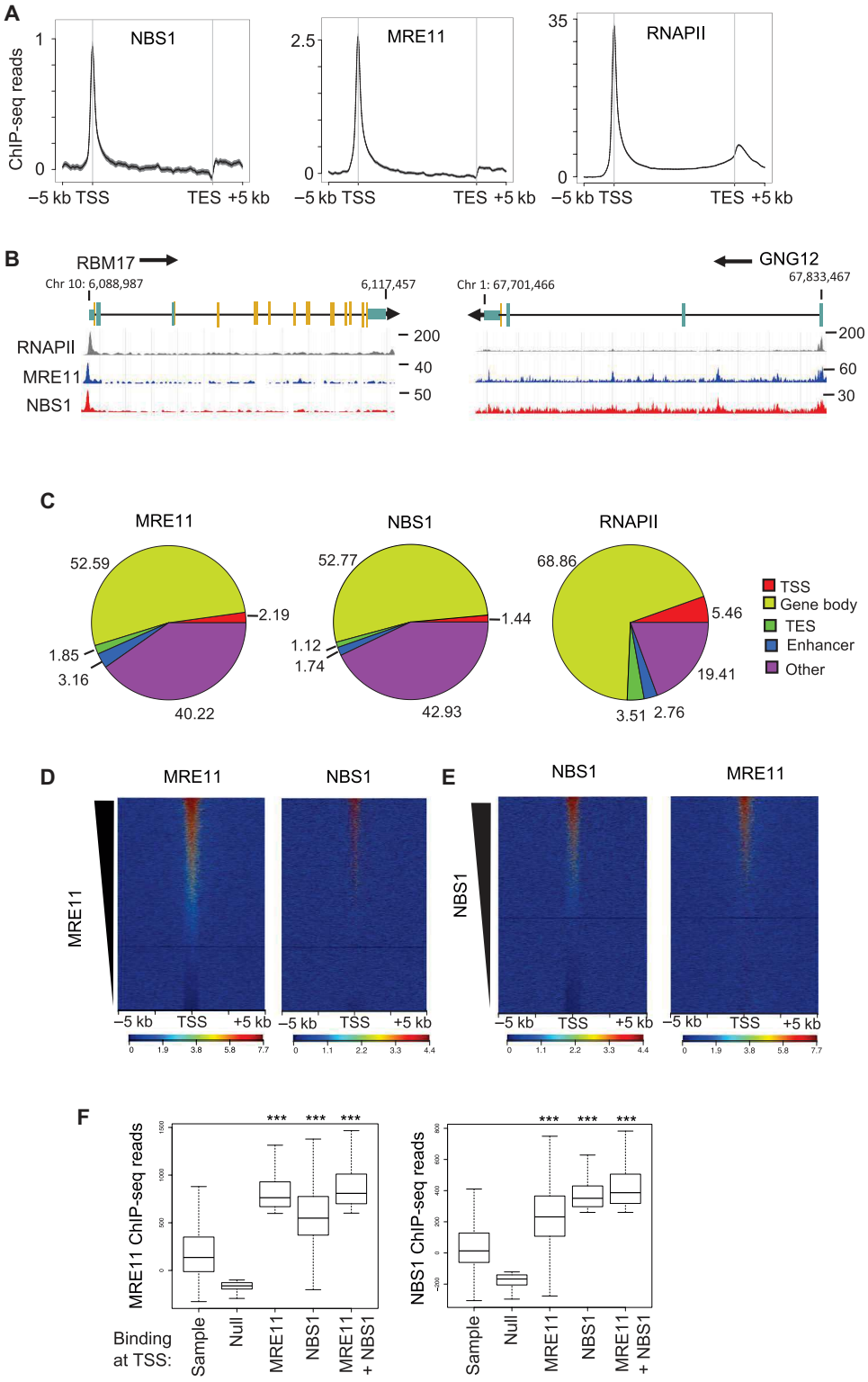


Fig. 2. MRN complex is associated with chromatin. (**A**) Scaled average density profiles of MRE11, NBS1, and RNAPII ChIP-seq reads across genes, ± 5 kb. (**B**) Browser shots of RNAPII, MRE11, or NBS1 ChIP-seq signal over a set of representative genes, as indicated. A schematic representation of the gene is shown above. (**C**) Pie charts showing the genomic distribution of ChIP-seq peaks of MRE11, NBS1, and RNAPII, as indicated. (**D** and **E**) ChIP-seq heatmaps centered and rank-ordered on MRE11 (D) or NBS1 (E). ChIP-seq reads of NBS1 (D) or MRE11 (E) were plotted respecting the same ranking. (**F**) Box plots of MRE11 or NBS1 ChIP-seq reads at genes for which the TSS was highly bound by MRE11, NBS1, both, or neither, as indicated, compared to a random sample of genes (***$P < 0.001$, Wilcoxon test; $n = 708$, 479, 1631, 1631, and 708, respectively).

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

Further genomic analysis of the localization of MRN on chromatin was performed by ranking genes according to the binding of MRE11 at TSSs (number of ChIP-seq reads per TSS, highest to lowest), as shown in Fig. 2D (left). Ranking genes by MRE11 reads also efficiently sorted the NBS1 signal (Fig. 2D). Similarly, gene ranking based on NBS1 read intensity at the TSS also efficiently ranked the MRE11 signal (Fig. 2E), highlighting a good correlation in their loading onto genes. Next, genes were grouped on the basis of high TSS binding of either MRE11 or NBS1, both factors together, or neither factor. As expected, genes highly bound by MRE11 had significantly more ChIP-seq reads for this factor compared to the random sample group (Fig. 2F, compare box 3 with 1; $P < 1 \times 10^{-20}$). The group of genes that were highly bound by NBS1 also had a significantly high number of MRE11 reads at the TSS compared to the sample group ($P < 1 \times 10^{-20}$; left box plot, compare box 4 with 1). Moreover, the highest number of MRE11 reads were found at cobound genes when compared to the sample group (left box plot, compare box 5 with 1; $P < 1 \times 10^{-20}$) or to TSSs highly bound by MRE11 only ($P < 1 \times 10^{-7}$; left box plot, compare box 5 with 3). The same trend was observed for NBS1 binding (Fig. 2F, right box plot). TSSs highly bound by NBS1 showed a high number of NBS1 reads as expected (compare box 4 with 1; $P < 1 \times 10^{-20}$), while TSSs highly bound by MRE11 also showed significantly high binding of NBS1 compared to the random sample group of genes (compare box 3 with 1; $P < 1 \times 10^{-20}$). As observed for MRE11, the highest number of NBS1 reads were found at genes that were cobound by MRE11 (compare box 5 to 4; $P < 1 \times 10^{-13}$). Correlation of MRE11 and NBS1 ChIP-seq reads furthermore confirmed that the two factors co-associate with the same regions (fig. S2D). These data indicate that MRE11 and NBS1 co-associate with a subset of human genes, likely in the context of the MRN DNA repair complex. They furthermore support proteomic data (Fig. 1) that uncovered previously unidentified biochemical interactions between MDC1/MRN complex and factors implicated in transcription and cotranscriptional mRNA processing.

MRN complex is associated with actively transcribed regions

Since MRE11 and NBS1 were found to be highly associated with TSSs and gene bodies, we wished to further investigate their association with transcription. First, genes were ranked according to the number of RNAPII ChIP-seq reads surrounding the TSS (±1 kb), from highest to lowest, which likely reflects the transcriptional activity of each gene. Next, respecting the same ranking, ChIP-seq reads for MRE11 and NBS1 were plotted. The resulting heatmaps (Fig. 3A) show that ranking genes by RNAPII also efficiently ranked both MRE11 and NBS1. Similarly, the averaged density profiles showed that MRE11 and NBS1 were more associated with genes having high RNAPII at the TSS than those having low RNAPII (Fig. 3A, top).

We next analyzed the amount of RNAPII at the TSS of genes highly bound by MRE11, NBS1, both MRE11 and NBS1, or neither factor (null) compared to a random sample group of genes (sample). Genes highly bound by either NBS1 or MRE11, alone or together, had significantly higher levels of RNAPII at the TSS compared to the random sample group of genes (Fig. 3B; boxes 3, 4, and 5, compared to box 1; $P > 1 \times 10^{-20}$). Pairwise correlation of MRE11, NBS1, and RNAPII ChIP-seq reads (figs. S2D and S3A) suggested that these factors co-associate with the same regions. To further demonstrate co-association of MRE11, NBS1, and RNAPII at cellular genes and HIV-1 LTR, Re-ChIP was performed (fig. S3B). This analysis demonstrated that chromatin fragments immunoprecipitated using NBS1 or MRE11 antibody could be reprecipitated by antibody against RNAPII, indicating that both proteins were colocalized on chromatin.

The association of MRE11 and NBS1 at the TSS of genes harboring varying amounts of RNAPII at the TSS was validated by ChIP-qPCR. The abundance of MRE11 or NBS1 largely mirrored that of RNAPII (fig. S3C). Furthermore, intersection analyses of ChIP-seq data showed a preferential colocalization of MRE11/NBS1 ChIP-seq peaks with active genes, defined as the top 50% of expressed genes when ranked by normalized RNA sequencing (RNA-seq) reads (*25*). Considering the peaks of each factor found at TSSs, 84% of MRE11 peaks ($n = 1597$), 79% of NBS1 peaks ($n = 687$), and 88% of MRE11 + NBS1 peaks ($n = 672$) were found at TSSs of active genes. Since RNAPII is also found at active enhancers, we asked whether MRE11 and NBS1 might associate with these regions. Both MRE11 and NBS1 were localized to enhancers, as compared to TSSs and random intergenic regions (Fig. 3C). Similar to TSSs, MRE11 and NBS1 were detected at enhancers with high levels of RNAPII, which are likely active enhancers, and were poorly associated with enhancers having low levels of RNAPII (Fig. 3D). Together, these data indicate the MRE11 and NBS1 subunits of MRN complex are associated with actively transcribing regions in a manner that appears to be tightly correlated to that of RNAPII.

We next tested whether the MRN complex co-associates with RNAPII following induction of a specific transcriptional program. To this end, cells were heat-shocked to induce the expression of a subset of genes. ChIP-seq of RNAPII, MRE11, and NBS1 was performed in heat-shocked and untreated cells. MRE11 and NBS1 binding was analyzed at genes induced upon heat shock, that is, those showing an increase in RNAPII association. As shown in Fig. 3E, both MRE11 and NBS1 became significantly associated with the TSS (left) and gene body (right) of up-regulated genes compared to a sample group of genes. Increased association of RNAPII and NBS1 was confirmed by ChIP-qPCR at several heat shock genes (Fig. 3F). The DSB marker, γH2AX was also increased following heat shock (Fig. 3G), suggesting that higher levels of transcription were associated with increased DNA damage at these regions. On the other hand, α-amanitin treatment was used to diminish RNAPII association with genes. As shown in fig. S3D, α-amanitin treatment induced loss of RNAPII as well as both MRE11 and NBS1 at target genes. Together, these data confirm that the binding of MRN tightly correlates with that of actively transcribing RNAPII.

MRN association with chromatin depends on RNAPII rather than transcription levels

Association of the MRN complex with actively transcribing regions is highly correlated to that of RNAPII (Fig. 3). However, it is not clear whether MRN is recruited through its interaction with the transcriptional machinery (Fig. 1) (*23*) or during the process of transcription. To distinguish between these possibilities, we sought to analyze MRN association under conditions where the abundance of RNAPII is not strictly correlated with the amount of transcription. To do so, RNAPII transcriptional elongation was blocked using the adenosine analog 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB). While DRB abolishes transcription in the coding region, it does not greatly affect the amount of transcriptional initiation. Consequently, DRB treatment causes RNAPII to accumulate at the 5′ end of genes disproportionately to the amount of transcription at the site. After 1 hour of DRB treatment, RNAPII accumulated near the TSS and diminished in the coding region and particularly at the

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

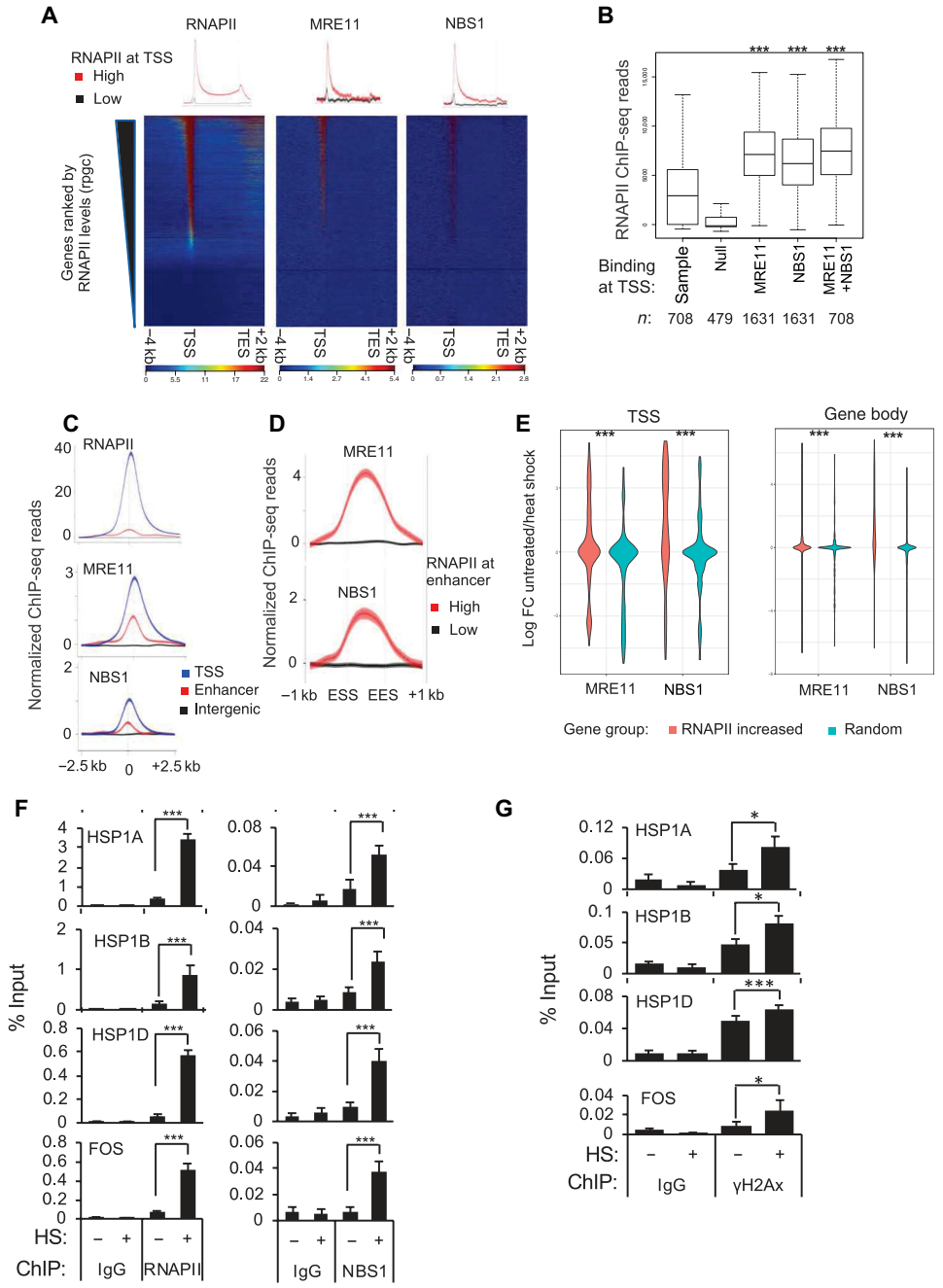


Fig. 3. MRN complex is associated with actively transcribed regions. (**A**) Heatmaps rank-ordered on normalized RNAPII ChIP-seq reads at the TSS and centered on genes. ChIP-seq reads of MRE11 or NBS1 were plotted respecting the same ranking. Averaged profiles of RNAPII, MRE11, or NBS1 ChIP-seq reads at the highest- and lowest-ranked RNAPII genes are shown above. (**B**) Box plots of normalized RNAPII ChIP-seq reads at genes highly bound at the TSS by MRE11, NBS1, both MRE11 and NBS1, or neither (null), compared to a random sample of genes (***$P < 0.001$, Wilcoxon test). (**C**) Averaged profiles of RNAPII, MRE11, or NBS1 ChIP-seq reads centered on TSS, enhancers, or intergenic regions. (**D**) Averaged profiles of MRE11 or NBS1 ChIP-seq reads at enhancers having the highest- or lowest-ranked RNAPII signal (ESS, enhancer start site; EES, enhancer end site). (**E**) Violin plots showing changes in MRE11 or NBS1 ChIP-seq reads at the TSS and body of genes at which RNAPII signal was increased in heat shocked samples relative to untreated controls, compared to a random group of genes in the same conditions (***$P < 0.001$, Wilcoxon test). (**F**) Binding of NBS1 or RNAPII was analyzed by ChIP-qPCR at the indicated genes from cells exposed or not to heat shock (HS). Data represent means ± SEM (***$P < 0.001$, Student's t test; $n = 3$). (**G**) Association of γH2AX was analyzed by ChIP-qPCR at the indicated genes from cells exposed or not to heat shock. Data represent means ± SEM (***$P < 0.001$ and *$P < 0.05$, Student's t test; $n = 3$).

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

3′ end of genes in the termination window, indicating that RNAPII elongation was efficiently inhibited (Fig. 4A). A longer blockade with DRB (3-hour treatment) caused RNAPII to accumulate at TSS boundaries, causing apparent accumulation near the start of the gene body as well as at upstream regions. By 1 hour after release, the profile of RNAPII had returned to that in untreated cells at the TSS and in the gene body, with a small reduction still evident at the TES.

Analysis of the profile of NBS1 under conditions of DRB treatment and release showed that it largely mirrored that of RNAPII, although the effects were more modest. The profiles showed the same qualitative effects, such as loss of signal at the TES, and leaking into TSS adjacent regions observed after 3 hours of treatment. The change in profiles of NBS1 and RNAPII were also quantitatively significant at both the TSS and TES regions (Fig. 4B). RNAPII and NBS1 association with chromatin during transcriptional blockade and release were confirmed by ChIP-qPCR at representative genes (Fig. 4C). Last, we analyzed the correlation between RNAPII and NBS1 association at individual genes for each condition. Association of the two factors with genes was positively correlated at each condition, with the highest correlation at 1 and 3-hour DRB treatment (Fig. 4D). Thus, overall, although the effects measured for NBS1 are more modest than those for RNAPII, the similarities in their binding profiles and dynamics suggest that MRN is likely recruited through its association with the RNAPII transcription complex rather than the amount of transcription per se.

Binding of MRN affects the transcriptional output of target genes

We next wondered whether the presence of MRN affects the level of transcription at its target genes. To test this, MRE11 and NBS1 were

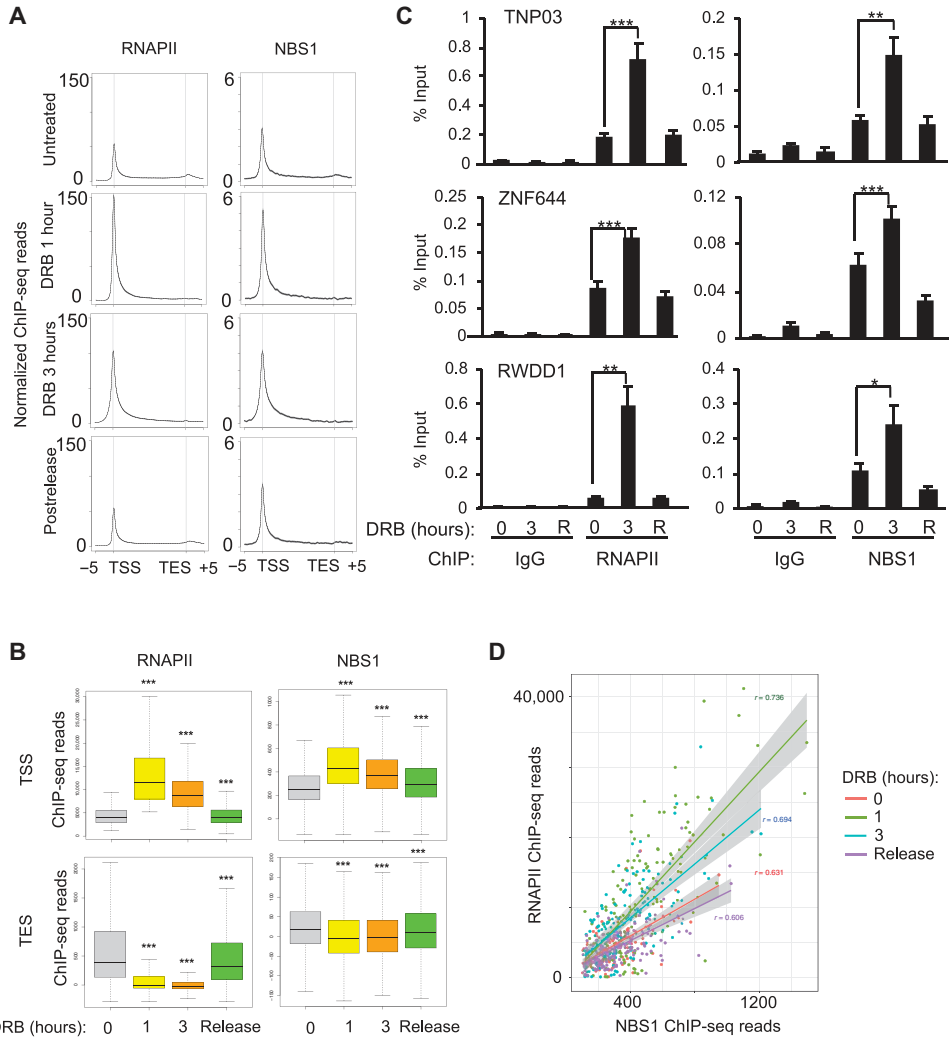


Fig. 4. MRN association with chromatin is correlated with the amount of RNAPII rather than transcription. (**A**) Average density profiles of ChIP-seq reads of RNAPII and NBS1 across genes showing the highest increase in RNAPII reads at 1-hour DRB treatment compared to a mock-treated control. ChIP-seq reads for RNAPII and NBS1 are shown at the same genes ±5 kb following treatment with DRB for 1, 3, or 1 hour after release, as indicated. (**B**) Box plots of RNAPII and NBS1 ChIP-seq reads at the TSS and TES of regions shown in (A) in DRB-treated samples compared to nontreated controls (0) (***$P < 0.001$, Wilcoxon test; $n = 4924$). (**C**) Association of RNAPII and NBS1 with selected target genes was measured by ChIP-qPCR following mock treatment or DRB treatment for 3 or 1 hour after release, as indicated. Data represent means ± SEM obtained from three independent experiments (***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$, independent Student's t test). (**D**) Scatter plot showing ChIP-seq reads of RNAPII and NBS1 at genes shown in (A) in untreated samples or samples treated with DRB, as indicated. The coefficient of correlation (r) was calculated for each condition as indicated on the graph.

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

depleted by RNA interference (RNAi). Consistent with previous findings showing that MRE11 is required for stability of the MRN complex (*26*), knockdown of MRE11 induced loss of MRE11 and destabilized NBS1 in cell extracts (fig. S4A). In contrast, RNAi-mediated depletion of NBS1 did not have a significant impact on MRE11 (fig. S4A), as reported previously (*26*). Next, RNAi-depleted cells were analyzed by ChIP-seq for the recruitment of each factor and RNAPII. expected, depletion of MRE11 led to a significant loss of MRE11 signal at genes highly bound at the TSS by MRE11, NBS1, or MRE11 + NBS1 (fig. S4B, left; compare blue boxes to gray boxes). Consistent with the reduction of NBS1 in extracts following knockdown of MRE11, NBS1 binding to chromatin was also significantly reduced under the same conditions (fig. S4B, right; compare blue boxes to gray boxes). Similarly, depletion of NBS1 reduced NBS1 binding at TSSs highly bound by NBS1, alone or together with MRE11 (right; compare red boxes to gray boxes). Depletion of NBS1 also diminished MRE11 signal at genes highly bound by MRE11, alone or together with NBS1 (left; red boxes 9 and 15), supporting the idea of cooperative binding to chromatin.

We next analyzed RNAPII association with genes in MRN-depleted and control cells. RNAPII accumulated over genes highly bound at the TSS by MRE11, NBS1, or both, following depletion of either MRE11 or NBS1 compared to a control knockdown (Fig. 5A). RNAPII binding was significantly increased at the TSS of genes highly bound by MRE11, NBS1, or both following loss of MRE11 or NBS1 [Fig. 5B (left and right, respectively) and fig. S5A] and across gene bodies (fig. S5B), which was shown by ChIP-qPCR at several representative genes (fig. S5C). Depletion of either NBS1 or MRE11 altered RNAPII association at a common set of genes, not only at genes where RNAPII levels were increased (fig. S5D, right matrix; compare deciles 1 on *X* and *Y* axes) but also at genes where RNAPII association was decreased (fig. S5D, right matrix; compare deciles 10 on *X* and *Y* axes; $P < 1 \times 10^{-20}$ by hypergeometric test). We next intersected the changes in RNAPII association with occupancy of MRE11 or NBS1. As shown in Fig. 5C, genes showing the highest increase in RNAPII ChIP-seq reads (deciles 1 on *X* and *Y* axes) were also highly associated with MRE11 (left) or NBS1 (right). Although depletion of NBS1 or MRE11 led to loss of RNAPII at a common set of genes (deciles 10 on *X* and *Y* axes), these, in contrast, were not significantly bound by either factor.

To determine whether the increase in RNAPII association with genes that occurred upon depletion of MRN reflected an increase in processive transcription, we determined the association of the elongating, serine-2 phosphorylated form of RNAPII (Ser^2) at genes highly bound by MRE11 and/or NBS1. Ser^2 RNAPII accumulated over genes bound by MRE11, NBS1, or both, following depletion of either MRE11 or NBS1 compared to controls (Fig. 5D) in a significant manner (Fig. 5E). MRN was also associated with active enhancers (Fig. 3, C and D). Analysis of RNAPII ChIP-seq showed that, like at genes, loss of MRE11 or NBS1 led to an increase in RNAPII association at enhancer regions (fig. S5E). However, in contrast to genes, Ser^2 RNAPII ChIP-seq reads were not significantly increased upon depletion of the factors (P value < 0.05). Therefore, MRE11 and NBS1 have a global impact on RNAPII levels genome-wide, frequently at the same regions. Up-regulation of RNAPII occurred at MRN-bound regions, whereas down-regulation of RNAPII may be independent of MRN association. These data suggest that MRN modulates the transcriptional output of target genes.

MRN protects actively transcribed regions from genomic instability

Given the key role of MRN in genomic stability, we next assessed SNPs by whole-genome sequencing in cells exposed to depletion of MRE11 or NBS1 or a control depletion. We discarded common SNPs found in all samples since these are likely to be false positives because of the background genomic environment. Approximately 30,000 SNPs unique to shCon cells were identified, while nearly 100,000 SNPs were detected in MRE11- or NBS1-depleted cells (Fig. 6A). Within genes, the majority of SNPs localized to gene bodies (Fig. 6A), which may reflect the genomic distributions of MRE11 and NBS1 (Fig. 2C). A ranking test by gene set enrichment analysis (GSEA) showed that SNPs found in MRE11- or NBS1-depleted cells could be readily predicted by the binding of MRE11 or NBS1 at TSSs (Fig. 6B). Furthermore, in keeping with the strict association between RNAPII and MRN occupancy at genes, SNP frequency was also well predicted by binding of RNAPII at the TSS.

Association of the DSB marker, γH2AX, with MRN-bound genes following loss of either MRE11 or NBS1 compared to control samples was also analyzed (Fig. 6C). γH2AX was detected at higher levels following loss of either MRE11 or NBS1 compared to a control knockdown at genes bound by MRE11, NBS1, or both factors (Fig. 6C and fig. S6). To determine whether the SNPs detected corresponded to regions showing increased levels of γH2AX, a GSEA was performed. As shown in Fig. 6D, SNPs found in MRE11- or NBS1-depleted cells could be predicted by increased detection of γH2AX. Together, these data suggest that binding of MRN over gene bodies preserves genes from DNA damage–induced mutations detected as SNPs. This highlights the key function of the MRN complex in maintaining genomic stability at actively transcribed regions.

DISCUSSION

Recent data strongly point to a central role for the MRN complex in the resolution of transcription-associated DNA repair. However, mechanistic details about the recruitment of MRN to sites of transcription-associated DNA damage are not clear. Using the unbiased proteomic technique, PICh, we found that DNA repair factors were among the most recruited proteins that specifically interacted with the chromatin region during transcription of the inducible HIV-1 LTR. In particular, MDC1, which interacts with MRN complex, became highly associated with the actively transcribed locus and was largely absent before transcriptional stimulation. Furthermore, we identified the interactome of MDC1 in cells in the absence of exogenous DNA damage. Consistent with results obtained by PICh at the highly active HIV-1 minigene, MDC1 interacted with many factors implicated in transcription and cotranscriptional RNA processing, indicating an association between MDC1 and active transcription of cellular genes. A robust interaction between MDC1 and RNAPII, as well as P-TEFb, was detected in the absence of exogenous DNA damage, suggesting a constitutive association with transcription. ChIP-seq revealed widespread localization of endogenous MRE11 and NBS1 across genes. Association of both MRE11 and NBS1 with genes was strongly correlated with transcriptional activity. Our data further suggest that MRN association with active genes is dependent on the presence of the RNAPII transcriptional complex rather than the level of transcription per se.

While it is unclear precisely how MRN associates with the RNAPII complex, note that MDC1 interacts with topoisomerase II

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

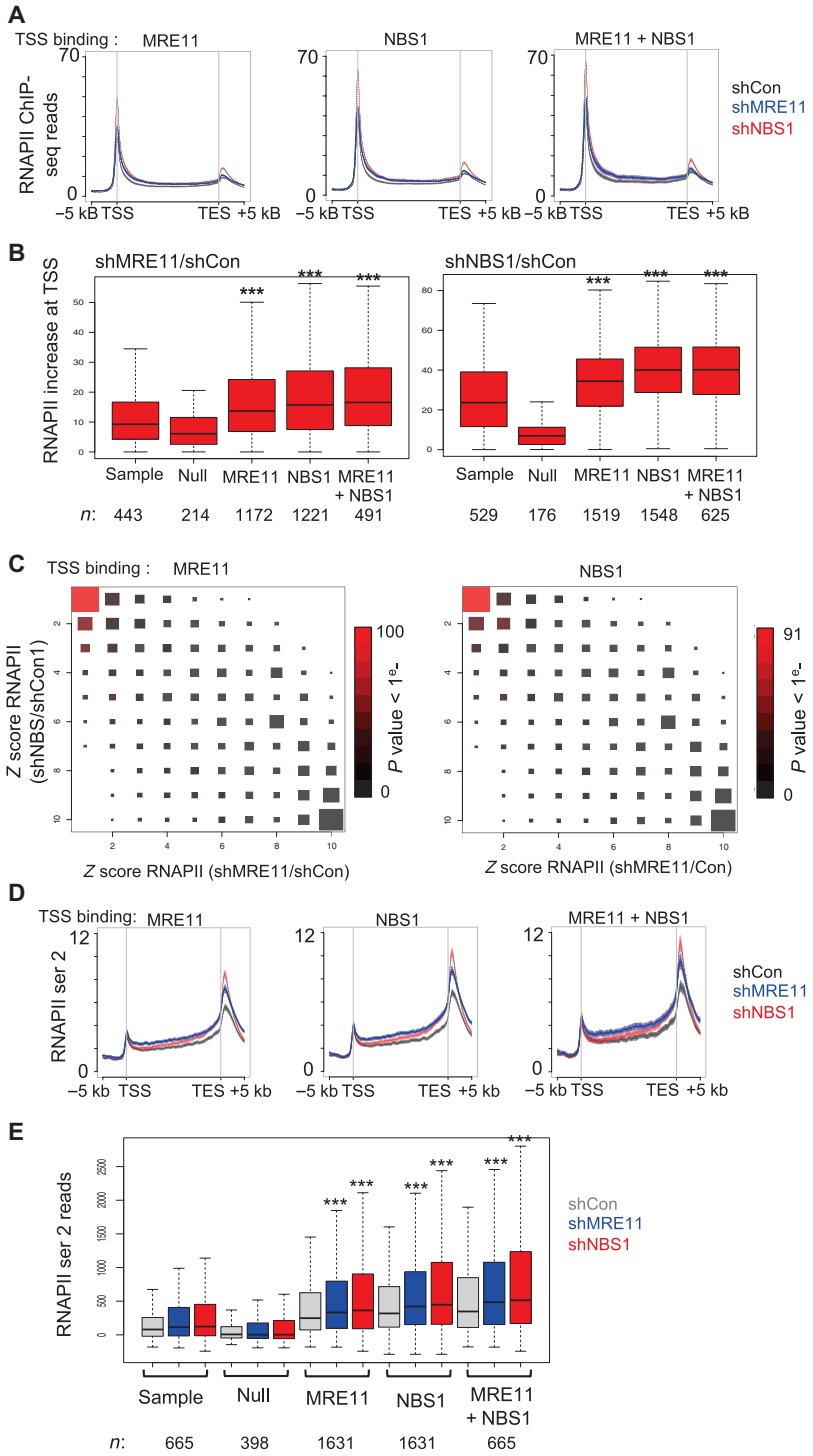


Fig. 5. Binding of MRN affects the transcriptional output of target genes. (**A**) Averaged profiles of RNAPII ChIP-seq reads across genes (±5 kb) highly bound at the TSS by MRE11, NBS1, or both in samples treated with short hairpin RNA (shRNA) targeting MRE11, NBS1, or a nontargeting control. (**B**) Box plots of RNAPII increase at the TSS of genes highly bound at the TSS by MRE11, NBS1, both MRE11 and NBS1, or neither (null), compared to a sample group of genes, following knockdown of MRE11 or NBS1, relative to a control knockdown. (***$P < 0.001$, Wilcoxon test). (**C**) Hinton diagram showing changes in RNAPII ChIP-seq reads at genes following knockdown of MRE11 or NBS1, relative to a control knockdown, ranked by deciles from the most up-regulated (1) to the most down-regulated (10), at genes highly bound at the TSS by either MRE11 (left) or NBS1 (right). The scale represents *P* value for the intersection of RNAPII increase with TSS binding of MRE11 or NBS1 (10^{-0} to 10^{-100} and 10^{-0} to 10^{-91}, respectively, hypergeometric test). The size of the box represents the number of genes, as shown in fig. S5D. (**D**) Averaged profiles of phospho-Ser^2 RNAPII ChIP-seq reads across genes highly bound at the TSS by MRE11, NBS1, or both following knockdown of MRE11 or NBS1 or a control. (**E**) Box plots representing phospho-Ser^2 RNAPII ChIP-seq reads at the TES of genes for which the TSS was highly bound by MRE11, NBS1, both MRE11 and NBS1, or neither (null), compared to a sample group of genes, following knockdown of MRE11 or NBS1 or a control (***$P < 0.001$, Wilcoxon test).

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

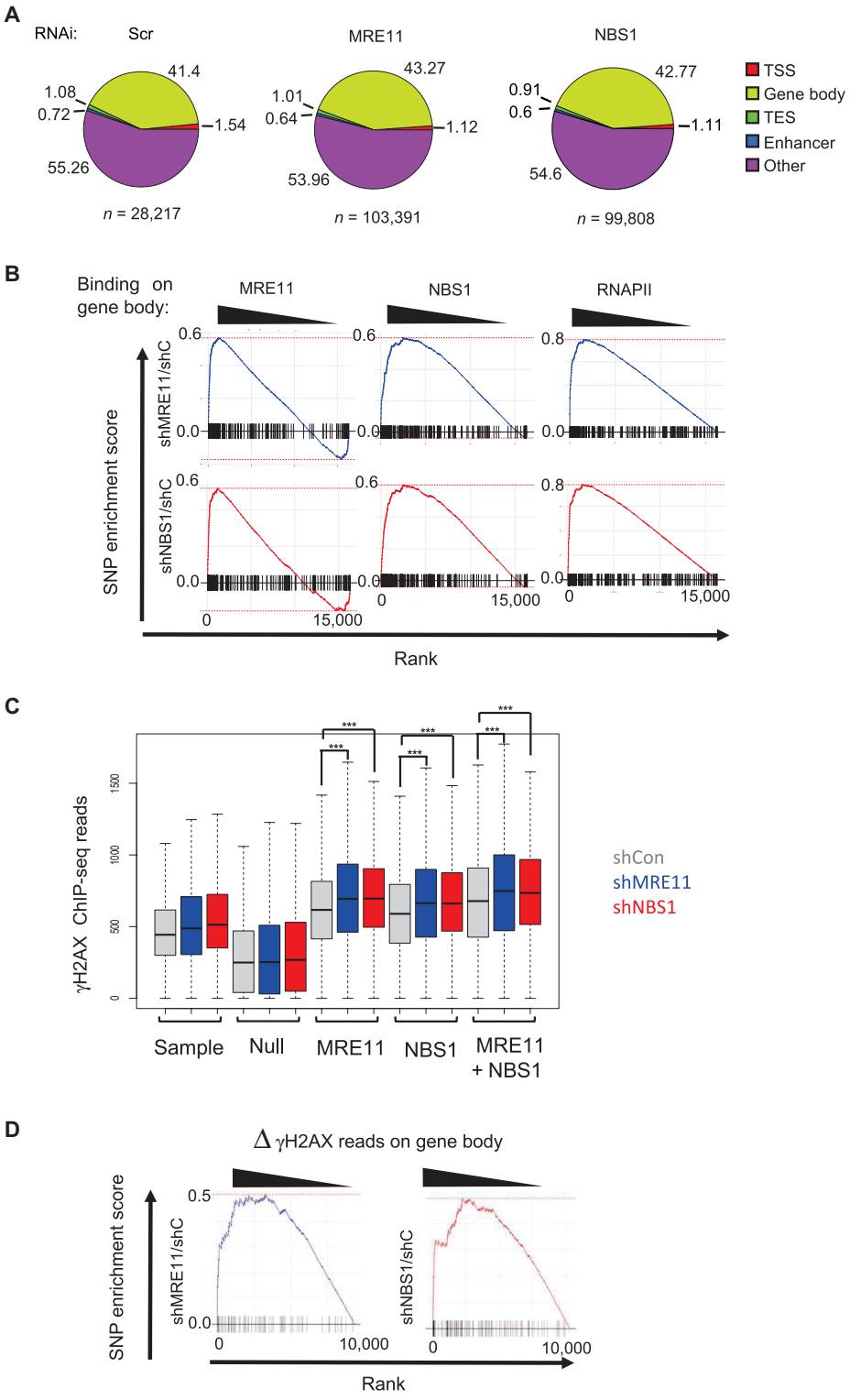


Fig. 6. MRN protects actively transcribed regions from genomic instability. (**A**) Pie charts showing the genomic distribution of SNPs detected by analyzing genomic DNA extracted from cells following depletion of MRE11, NBS1, or a control, as indicated. (**B**) GSEA of SNPs detected by analyzing genomic DNA extracted from cells following depletion of MRE11 and NBS1 compared to shCon sample, as indicated, ranked by ChIP-seq reads of MRE11, NBS1, or RNAPII in gene bodies, as indicated (all tests were significant; $P < 10^{-3}$, Wilcoxon test). (**C**) Box plots representing γH2AX ChIP-seq read depth across the body of genes for which the TSS was highly bound by MRE11 or NBS1, both, or neither (null), compared to a sample group of genes, following knockdown of MRE11 or NBS1 or a control knockdown, as indicated (***$P < 0.001$, Wilcoxon test). (**D**) GSEA of SNPs detected by analyzing genomic DNA extracted from cells following depletion of MRE11 or NBS1 compared to a control depletion (shCon), ranked by the increase in γH2AX ChIP-seq reads across the gene body under the same conditions (all tests were significant; $P < 10^{-3}$, Wilcoxon test).

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

(TOP2) through its C-terminal BRCT domains (*26*). TOP2 is a component of the RNAPII transcriptional complex and is required for transcription through chromatin (*27*). Whether MDC1 associates with the RNAPII complex through TOP2 or another transcription-associated factor, MRN, is found at the most highly transcribed genes, and its association with genes is highly correlated to that of RNAPII. Notably, upon depletion of either MRE11 or NBS1, MRN target genes accumulated DNA damage in an MRN- and RNAPII-dependent manner. Together, these data suggest that MRN is associated with the RNAPII transcription complex probably to scan for transcription-associated DNA damage and initiate repair.

NHEJ and HR pathways compete for the repair of DSBs (*1*). The degree of chromatin compaction is thought to influence repair pathway choice. DSBs occurring in open chromatin undergo end resection and are predominantly repaired by HR, which results in faithful repair and suppresses dangerous mutations from arising in coding regions of the genome (*7*, *9*, *28*). The preference for using HR to repair DSBs in transcribed regions is not due to cell cycle–dependent availability of factors. Increasing evidence indicates that DNA damage occurring on active genes is repaired through a specialized mechanism involving HR. The current model suggests that upon DNA damage, the damaged site is targeted by MRN complex and undergoes the initial steps in HR, which disfavors repair of the site by NHEJ. These MRN-marked sites subsequently pair and cluster into higher-order structures to be repaired flawlessly during G_2 phase of the cell cycle by the HR pathway. This mechanism preserves the fidelity of the genome in coding regions. While the preference for HR at actively transcribed is likely due to a combination of factors including nuclear position and the epigenetic landscape, we speculate that the physical link between MRN, which is required for commitment to HR, and RNAPII transcription complex might contribute to the preference for HR at transcribed regions. Being a component of the transcription complex, MRN may have immediate access to transcription-induced DSBs, facilitating end resection and commitment to HR.

Several lines of evidence indicate that the presence of DSBs in coding regions leads to transcriptional arrest. For example, transcriptional silencing occurs on chromatin in the vicinity of DSBs in an ATM (ataxia telangiectasia mutated)–dependent manner (*29*). Paradoxically, we observed that transcription was modestly increased in cells depleted of MRE11 or NBS1. There are several possible explanations for this apparently conflicting result. It has recently been shown that in response to ultraviolet irradiation, RNAPII is released from promoter-proximal regions into gene bodies to promote detection of DNA damage within genes (*30*). Thus, the increased RNAPII occupancy observed upon loss of MRN could also be a scanning mechanism to rapidly detect DNA damage within transcribed genes. Alternatively, the measured increase in transcription could be due to failure to establish HR at damaged sites in the absence of MRN. The use of HR to repair damage at active genes privileges genome fidelity at a slight cost to gene expression. Transcription of genes that have sustained DNA damage will be arrested until repair is completed in G_2. Therefore, at the global level, expression of the gene will be somewhat diminished, but DNA repair will be flawless. In cells depleted of MRN, which is required to commit to HR, DNA repair most likely occurs by NHEJ. In this case, the cost to transcription will be minimal, as NHEJ operates throughout the cell cycle. However, repair will likely occur at the expense of genome fidelity, as NHEJ repair is characterized by the appearance of SNPs. This is the outcome in cells depleted of MRN. RNAPII occupancy at both the TSS and gene body was higher than in control cells. However, whole-genome sequencing revealed that DNA repair was highly error prone, with the appearance of thousands of SNPs in the gene bodies of target genes. Furthermore, the number of SNPs detected was highly correlated to the abundance of MRN at the gene under control conditions and the presence of γH2AX in MRN-depleted conditions. This scenario implies that MRN does not directly influence transcription but that the absence of MRN predisposes to an alternative DNA repair pathway that is more favorable for gene expression but at the cost of genome fidelity.

MATERIALS AND METHODS

Cell culture, treatment, and lentiviral infection

HeLa cells were grown in Dulbecco's modified Eagle's minimal essential medium (DMEM) (Sigma-Aldrich, D6429), supplemented with 10% fetal calf serum (FCS; Eurobio Scientific, CVFSVF00-01) and containing 1% penicillin-streptomycin (Sigma-Aldrich, P4333). HEK-293T were grown in Hepes-modified DMEM (Sigma-Aldrich, D6171), supplemented with 10% FCS (Eurobio Scientific, CVFSVF00-01) and containing 1% penicillin-streptomycin (Sigma-Aldrich, P4333). All cells were grown in a humidified incubator at 37°C with 5% CO_2. Where indicated, cells were subjected to heat shock at 42°C for 1 hour, and the medium was replaced before fixation. Cells were treated with α-amanitin (Sigma-Aldrich, A2263) for 16 hours at 20 μg/ml or with 150 μM DRB (Sigma-Aldrich, D1916) or mock-treated with dimethyl sulfoxide where indicated for 0, 1, or 3 hours. After 3 hours of DRB treatment, cells were washed and incubated for 1 hour in fresh cell culture medium.

Production of short hairpin RNA (shRNA)–expressing lentiviral particles was performed as described previously (*20*) using plasmids expressing shRNAs targeting MRE11 (Sigma-Aldrich MISSION shRNA, TRCN000338391), NBS1 (Sigma-Aldrich MISSION shRNA, TRCN0000288622), or a nontargeting control (obtained through Addgene, plasmid 1864), as shown in table S5. For knockdown experiments, HeLa cells were transduced with lentiviral particles and harvested 4 days later, as described previously (*20*).

Transactivation of the luciferase reporter gene was achieved by transfecting HeLa cells in 15-cm dishes with 8 μg of pcDNA3 vector or pcDNA3-Tat.Flag using TransIT-X2 transfection reagent (Mirus, MIR 6000), according to the manufacturer's instructions. After 24 hours, cells were used for ChIP or Re-ChIP assays.

Antibodies

Antibodies used in this study are shown in table S3.

CRISPR-Cas9–mediated editing of endogenous MDC1 gene

A single guide RNA (sgRNA) targeting the MDC1 gene around the ATG translation start site was cloned in pSpCas9 (BB)-2A–green fluorescent protein plasmid (Addgene, no. 48138). The plasmid was then transfected into HEK-293T cells along with a single-stranded oligodeoxynucleotide (table S4) harboring the Flag-HA sequence flanked by homology sequences to MDC1 around the cleavage site. Single cells were isolated and amplified. HEK-293T clones expressing Flag-HA MDC1 were identified by PCR and confirmed by Western blot using anti-HA and anti-Flag antibodies.

Coimmunoprecipitation analysis

Coimmunoprecipitation was performed using nuclear extracts of HEK-293T cells. Cells were lysed in ice-cold hypotonic buffer [20 mM

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

tris (pH 7.6), 10 mM KCl, and 1.5 mM $MgCl_2$] supplemented with EDTA-free complete protease inhibitor mixture (Roche) for 15 min on ice. NP-40 was added at 0.5% final, and extracts were centrifuged 1 min at 14,000*g*/4°C. The pellet (nuclei) was resuspended in nuclease buffer [20 mM tris (pH 7.6), 150 mM NaCl, 1.5 mM $MgCl_2$, 2.5 mM $CaCl_2$, and 0.5 μl of 100 mM phenylmethylsulfonyl fluoride] and incubated with micrococcal nuclease (2×10^3 U/ml; New England Biolabs) for 2 hours at 4°C. Lysates were cleared by centrifugation at 14,000*g*/4°C for 10 min and diluted in immunoprecipitation buffer [50 mM tris (pH 7.6), 150 mM NaCl, and 1% NP-40] supplemented with protease inhibitors. Protein concentration was determined using the Bradford reagent (Bio-Rad). Immunoprecipitations were performed on 400 μg of protein extracts with the indicated antibodies and rotated overnight at 4°C. Protein G Dynabeads were washed three times in immunoprecipitation buffer, added to protein extracts/antibody solution, and incubated for 2 hours at 4°C. Immunoprecipitates were washed extensively with the immunoprecipitation buffer, resuspended in protein sample loading buffer, boiled for 5 min, and analyzed by Western blotting using the antibodies shown in table S3.

Proteomics of isolated chromatin fragments

Proteomics of isolated chromatin segments was performed as described previously (*12*) using U20S cells expressing HIV-LTR-MS2 (*13*). Sequences of probes used are shown in table S4. Silver staining was performed according to the manufacturer's instruction (SilverQuest, Invitrogen). Mass spectrometry was performed at the Taplin Facility, Harvard University, Boston, MA.

MDC1 protein complex purification

MDC1 complexes were purified from nuclear extracts of HEK-293T cells stably expressing Flag-HA-MDC1 by two-step affinity chromatography, as described previously (*20*). Sequential Flag and HA immunoprecipitations were performed on equal amounts of proteins. Silver staining was performed according to the manufacturer's instructions (SilverQuest, Invitrogen). Mass spectrometry was performed at the Taplin Facility, Harvard University, Boston, MA.

Glycerol gradient sedimentation analysis

Separation of active and inactive P-TEFb complexes was performed as described previously (*31*). Briefly, glycerol gradients (10 to 30%) were formed by pipetting 2 ml of each of the glycerol fraction (10, 15, 20, 25, and 30% v/v) in buffer A [20 mM Hepes (pH 7.9), 0.3 M KCl, 0.2 mM EDTA, and 0.1% NP-40] into centrifugation tubes (Beckman, 331372). Gradients were formed by standing for 6 hours at 4°C. Cells were lysed in 0.5 ml of buffer A [10 mM tris-HCl (pH 7.4), 150 mM NaCl, 2 mM EDTA, 1% NP-40, and 0.1% protease inhibitor] for 30 min at 4°C. The lysates were centrifuged at 10,000*g* for 10 min, and the supernatants were loaded into tubes with the preformed glycerol gradients. Protein complexes were then fractionated by centrifugation in an SW 41 Ti rotor (Beckman) at 38,000 rpm for 21 hours. Fractions (0.5 ml) were collected, precipitated with trichloroacetic acid, and lastly analyzed by immunoblotting with the appropriate antibodies.

ChIP, library preparation, and sequencing

ChIP-seq (*32*) was performed from HeLa cells using the ChIP-IT High Sensitivity Kit from Active Motif (reference 53040) according to the manufacturer's instructions. Sonication was performed using the Qsonica Q700 Sonicator with microtip of 1/8 inches (reference 4418) at 11% amplitude and 13 min of processing time (30-s "ON" and 30-s "OFF"). Each ChIP used 30 μg of chromatin along with 4 μg of antibody detecting MRE11, NBS1, RNAPII, or phospho-Ser^2 RNAPII (table S3). ChIP-seq libraries were constructed using the Next Gen DNA Library Kit (Active Motif, 53216 and 53264). Library quality was assessed using Agilent 2100 Bioanalyzer and Agilent High Sensitivity DNA assay. High-throughput sequencing was performed by sequence-by-synthesis technique using a NextSeq 500 (Illumina) at the Genom'ic Facility, Institut Cochin, Paris.

ChIP and Re-ChIP–qPCR

RNAPII, NBS1, MRE11, and MDC1 ChIP were performed using the iDeal ChIP-qPCR Kit (Diagenode, catalog no. C01010180) following the manufacturer's instructions. HeLa cells were sonicated using the Bioruptor Pico (Diagenode, catalog no. B01060001) for 8 cycles of 30-s ON and 30-s OFF at high-power setting. For RNAPII, MRE11, and MDC1 ChIP, 3 μg of antibody was used, while 1.5 μg of antibody was used for NBS1 ChIP. γH2AX ChIP was performed using the ChIP-IT Express Enzymatic Kit (Active Motif, catalog no. 53009), following the manufacturer's instructions. Chromatin was digested for 7 min, and 50 μg of chromatin and 3 μg of antibody were used for each γH2AX ChIP.

Re-ChIP was performed using the Re-ChIP-IT Kit (Active Motif, catalog no. 53016) according to the manufacturer's instruction. Chromatin for the Re-ChIP was prepared using the ChIP-IT High Sensitivity Kit (Active Motif, catalog no. 53040), as described above. For each Re-ChIP, 50 μg of chromatin and 3 μg of antibody were used. Antibody used and sequences of primers used for real-time qPCR analysis are shown in table S4.

Whole-genome sequencing

Genomic DNA was extracted from cells treated with lentiviral particles expressing shCon, shMRE11, or shNBS1 using the QIAGEN DNeasy Blood and Tissue Kit (reference 69504), according to the manufacturer's instructions. To avoid RNA contamination, extracts were treated with ribonuclease A (QIAGEN, 19101) according to the manufacturer's instructions. Equal amounts of DNA were used for library preparation. Whole-genome sequencing was performed by Novogene, Cambridge, UK.

Bioinformatic analysis

For PICh and mass spectrometry proteomics data analysis, uniquely mapping peptides were counted for each protein in each condition. Proteins whose abundance was greater than sevenfold than that in the control condition were analyzed using Gene Ontology enrichment analysis (*33*, *34*), searching for enriched biological processes.

For analysis of ChIP-seq data, sequencing reads were first filtered, using fastq_illumina_filter, and quality control of filtered reads was performed using FastQC. Filtered reads were then aligned onto the HG38 genome (*24*) using the Burrows-Wheeler Aligner (*35*) with default parameters. The sorted BAM files generated by SAMtools (*36*) keeping only reads with a mapping quality at least 30 were then normalized by DeepTools' (*37*) bamCoverage function, with a bin size of 10 bp. RPGC normalization was applied, with an effective genome size of 2,913,022,398 bp according to DeepTools' user manual instructions. Files were then further normalized by subtracting an RPGC (reads per genomic content) normalized input

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

data file, using bigwigCompare. The input data used for normalization were generated by averaging four inputs from separate assays, so as to minimize variability and biases that can be introduced during input normalization.

From these normalized data files, peak calling was performed using NormR's enrichR function, searching for enrichment of each BAM file of ChIP-seq reads against the input BAM file using a false discovery rate correction. Genomic Ranges (*38*) was then used to determine overlap between the peak range and genomic features of interest, such as genes with a TSS and TES from GRCh38 and enhancers in HeLa S3 from ENCODE. Profile matrices were extracted from the normalized data files using DeepTools' computeMatrix using a bin size of 10 bp. For each gene, matrices for both TSS and TES were used, with 5-kb flanking on each side of the feature, as well as gene bodies, which were scaled to 4 kb in length and with 4-kb flanking before the TSS and 2-kb flanking after the TES. For enhancers, the enhancer body was scaled to 500 bp, with 1-kb flanking on either side. Using these profile matrices, quantification of normalized reads was calculated by summing the score of each appropriate bin for the feature. Unless indicated otherwise, this was from start to end for gene body and for enhancers and 500 bp before and after TSS and TES.

RNAPII binding variation between conditions was calculated using *z* scores, which were calculated as follows $Z\ \text{score} = \frac{(KD - WT)}{(KD - WT)/2}$ (i.e., difference weighted by mean signal), which transforms the distribution of variations into a normal distribution, allowing for better statistical interpretation of the variations.

Scatter plots, box plots, violin plots, pie charts, and bar plots were created using either basic R plotting functions or ggplot2 functions (*39*). For box plots, the gene groups analyzed were "sample" (randomly chosen genes within the sample group), "null" (genes within the lowest decile of ChIP-seq reads for a given protein), MRE11 or NBS1 (genes within the highest decile of ChIP-seq reads for individual factor binding), and "MRE11 + NBS1" (genes within the highest decile of ChIP-seq reads for both factors). Average binding profiles of proteins across genomic features of interest were generated using the plotAverage function "seqplots" (*40*). Heatmaps on genomic features were created using genomation's (*41*) gridHeat function, using the profile matrices generated by DeepTools. GSEA was performed using fgsea as described previously (*42*), and all enrichment plots were created using the plotEnrichment function of the package. Decile matrices were created using color2D matplot from the plotrix package. Genes ranked by *z* score of RNAPII variation were split into deciles, for both shMRE11 versus shCon and shNBS1 versus shCon. Then, for each pair of deciles, the number of genes in the intersection of those two deciles was saved into a matrix, creating a 10 by 10 matrix of integers. The number of genes in each intersection was tested for significance by hypergeometric test, creating a 10 by 10 matrix of *P* values. The genes at each intersection were also tested by hypergeometric test for significant enrichment in either top MRE11 or NBS1 binding levels (decile 1 of genes ranked by MRE11 or NBS1 binding). The group of active genes was determined by first ranking all genes from highest to lowest expression, using HTSeq count (Bioconductor) run on RNA-seq data in HeLa cells described in (*25*). The top 50% of ranked genes were selected as "active genes." For analysis of whole-genome sequencing data, following processing by ANNOVAR (*43*) on the sequencing platform, SNP location was then overlapped with gene bodies using Genomic Ranges functions, which yields a quantification per gene.

Statistical analysis

Data presented as histograms are shown as means ± SD. Comparison between two groups was analyzed by two-tailed Student's *t* test, and asterisks represented significance defined as $*P < 0.05$, $**P < 0.01$, or $***P < 0.001$. Other statistical methods are described above in the "Bioinformatic analysis" section.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/21/eabb2947/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. D. Setiaputra, D. Durocher, Shieldin—The protector of DNA ends. *EMBO Rep.* **20**, e47560 (2019).
2. A. Shibata, D. Moiani, A. S. Arvai, J. Perry, S. M. Harding, M. M. Genois, R. Maity, S. van Rossum-Fikkert, A. Kertokalio, F. Romoli, A. Ismail, E. Ismalaj, E. Petricci, M. J. Neale, R. G. Bristow, J. Y. Masson, C. Wyman, P. A. Jeggo, J. A. Tainer, DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. *Mol. Cell* **53**, 7–18 (2014).
3. A. Janssen, G. A. Breuer, E. K. Brinkman, A. I. van der Meulen, S. V. Borden, B. van Steensel, R. S. Bindra, J. R. LaRocque, G. H. Karpen, A single double-strand break system reveals repair dynamics and mechanisms in heterochromatin and euchromatin. *Genes Dev.* **30**, 1645–1657 (2016).
4. C. Lemaître, A. Grabarz, K. Tsouroula, L. Andronov, A. Furst, T. Pankotai, V. Heyer, M. Rogier, K. M. Attwood, P. Kessler, G. Dellaire, B. Klaholz, B. Reina-San-Martin, E. Soutoglou, Nuclear position dictates DNA repair pathway choice. *Genes Dev.* **28**, 2450–2463 (2014).
5. S. X. Pfister, S. Ahrabi, L. P. Zalmas, S. Sarkar, F. Aymard, C. Z. Bachrati, T. Helleday, G. Legube, N. B. la Thangue, A. C. G. Porter, T. C. Humphrey, SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep.* **7**, 2006–2018 (2014).
6. F. Aymard, B. Bugler, C. K. Schmidt, E. Guillou, P. Caron, S. Briois, J. S. Iacovoni, V. Daburon, K. M. Miller, S. P. Jackson, G. Legube, Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat. Struct. Mol. Biol.* **21**, 366–374 (2014).
7. T. Yasuhara, R. Kato, Y. Hagiwara, B. Shiotani, M. Yamauchi, S. Nakada, A. Shibata, K. Miyagawa, Human Rad52 promotes XPG-mediated R-loop processing to initiate transcription-associated homologous recombination repair. *Cell* **175**, 558–570.e11 (2018).
8. J. A. Aten, J. Stap, P. M. Krawczyk, C. van Oven, R. A. Hoebe, J. Essers, R. Kanaar, Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* **303**, 92–95 (2004).
9. F. Aymard, M. Aguirrebengoa, E. Guillou, B. M. Javierre, B. Bugler, C. Arnould, V. Rocher, J. S. Iacovoni, A. Biernacka, M. Skrzypczak, K. Ginalski, M. Rowicka, P. Fraser, G. Legube, Genome-wide mapping of long-range contacts unveils clustering of DNA double-strand breaks at damaged active genes. *Nat. Struct. Mol. Biol.* **24**, 353–361 (2017).
10. P. M. Krawczyk, T. Borovski, J. Stap, T. Cijsouw, R. ten Cate, J. P. Medema, R. Kanaar, N. A. P. Franken, J. A. Aten, Chromatin mobility is increased at sites of DNA double-strand breaks. *J. Cell Sci.* **125**, 2127–2133 (2012).
11. V. Roukos, T. C. Voss, C. K. Schmidt, S. Lee, D. Wangsa, T. Misteli, Spatial dynamics of chromosome translocations in living cells. *Science* **341**, 660–664 (2013).
12. J. Dejardin, R. E. Kingston, Purification of proteins associated with specific genomic Loci. *Cell* **136**, 175–186 (2009).
13. S. Boireau, P. Maiuri, E. Basyuk, M. de la Mata, A. Knezevich, B. Pradet-Balade, V. Bäcker, A. Kornblihtt, A. Marcello, E. Bertrand, The transcriptional cycle of HIV-1 in real-time and live cells. *J. Cell Biol.* **179**, 291–304 (2007).
14. J. F. Goodwin, K. E. Knudsen, Beyond DNA repair: DNA-PK function in cancer. *Cancer Discov.* **4**, 1126–1139 (2014).
15. C. A. Sartorius, G. S. Takimoto, J. K. Richer, L. Tung, K. B. Horwitz, Association of the Kuautoantigen/DNA-dependent protein kinase holoenzyme and poly(ADP-ribose) polymerase with the DNA binding domain of progesterone receptors. *J. Mol. Endocrinol.* **24**, 165–182 (2000).
16. R. F. Chun, O. J. Semmes, C. Neuveut, K. T. Jeang, Modulation of Sp1 phosphorylation by human immunodeficiency virus type 1 Tat. *J. Virol.* **72**, 2615–2629 (1998).
17. S. Tyagi, A. Ochem, M. Tyagi, DNA-dependent protein kinase interacts functionally with the RNA polymerase II complex recruited at the human immunodeficiency virus (HIV) long terminal repeat and plays an important role in HIV gene expression. *J. Gen. Virol.* **92**, 1710–1720 (2011).
18. S. M. Zhang, H. Zhang, T.-Y. Yang, T.-Y. Ying, P.-X. Yang, X.-D. Liu, S.-J. Tang, P.-K. Zhou, Interaction between HIV-1 Tat and DNA-PKcs modulates HIV transcription and class switch recombination. *Int. J. Biol. Sci.* **10**, 1138–1149 (2014).

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

19. W. Kaczmarski, S. A. Khan, Lupus autoantigen Ku protein binds HIV-1 TAR RNA in vitro. *Biochem. Biophys. Res. Commun.* **196**, 935–942 (1993).
20. X. Contreras, K. Salifou, G. Sanchez, M. Helsmoortel, E. Beyne, L. Bluy, S. Pelletier, E. Rousset, S. Rouquier, R. Kiernan, Nuclear RNA surveillance complexes silence HIV-1 transcription. *PLOS Pathog.* **14**, e1006950 (2018).
21. D. Latreille, L. Bluy, M. Benkirane, R. E. Kiernan, Identification of histone 3 variant 2 interacting factors. *Nucleic Acids Res.* **42**, 3542–3550 (2014).
22. B. Sobhian, N. Laguette, A. Yatim, M. Nakamura, Y. Levy, R. Kiernan, M. Benkirane, HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol. Cell* **38**, 439–451 (2010).
23. F. Michelini, S. Pitchiaya, V. Vitelli, S. Sharma, U. Gioia, F. Pessina, M. Cabrini, Y. Wang, I. Capozzo, F. Iannelli, V. Matti, S. Francia, G. V. Shivashankar, N. G. Walter, F. d'Adda di Fagagna, Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. *Nat. Cell Biol.* **19**, 1400–1411 (2017).
24. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, Y. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.-F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.-M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H.-C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, J. Szustakowki; International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
25. B. Stadelmayer, G. Micas, A. Gamot, P. Martin, N. Malirat, S. Koval, R. Raffel, B. Sobhian, D. Severac, S. Rialle, H. Parrinello, O. Cuvier, M. Benkirane, Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat. Commun.* **5**, 5531 (2014).
26. Z. H. Zhong, W. Q. Jiang, A. J. Cesare, A. A. Neumann, R. Wadhwa, R. R. Reddel, Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. *J. Biol. Chem.* **282**, 29314–29322 (2007).
27. N. Mondal, J. D. Parvin, DNA topoisomerase IIalpha is required for RNA polymerase II transcription on chromatin templates. *Nature* **413**, 435–438 (2001).
28. W. T. Lu, B. R. Hawley, G. L. Skalka, R. A. Baldock, E. M. Smith, A. S. Bader, M. Malewicz, F. Z. Watts, A. Wilczynska, M. Bushell, Drosha drives the formation of DNA:RNA hybrids around DNA break sites to facilitate DNA repair. *Nat. Commun.* **9**, 532 (2018).
29. N. M. Shanbhag, I. U. Rafalska-Metcalf, C. Balane-Bolivar, S. M. Janicki, R. A. Greenberg, ATM-dependent chromatin changes silence transcription in cis to DNA double-strand breaks. *Cell* **141**, 970–981 (2010).
30. M. D. Lavigne, D. Konstantopoulos, K. Z. Ntakou-Zamplara, A. Liakos, M. Fousteri, Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate. *Nat. Commun.* **8**, 2076 (2017).
31. X. Contreras, M. Barboric, T. Lenasi, B. M. Peterlin, HMBA releases P-TEFb from HEXIM1 and 7SK snRNA via PI3K/Akt and activates HIV transcription. *PLOS Pathog.* **3**, 1459–1469 (2007).
32. A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, K. Zhao, High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
33. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
34. The Gene Ontology Consortium, The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
35. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
36. J. K. Bonfield, J. Marshall, P. Danecek, H. Li, V. Ohan, A. Whitwham, T. Keane, R. M. Davies, HTSlib: C library for reading/writing high-throughput sequencing data. *Gigascience* **10**, giab007 (2021).
37. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
38. M. Lawrence, W. Huber, H. Pagès, P. Aboyoun, M. Carlson, R. Gentleman, M. T. Morgan, V. J. Carey, Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).
39. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
40. P. Stempor, J. Ahringer, SeqPlots—Interactive software for exploratory data analyses, pattern discovery and visualization in genomics. *Wellcome Open Res.* **1**, 14 (2016).
41. A. Akalin, V. Franke, K. Vlahovicek, C. E. Mason, D. Schubeler, Genomation: A toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics* **31**, 1127–1129 (2015).
42. A. Heurteau, C. Perrois, D. Depierre, O. Fosseprez, J. Humbert, S. Schaak, O. Cuvier, Insulator-based loops mediate the spreading of H3K27me3 over distant micro-domains repressing euchromatin genes. *Genome Biol.* **21**, 193 (2020).
43. K. Wang, M. Li, H. Hakonarson, ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

Acknowledgments: We thank C. Dargemont and members of the Cuvier and Kiernan laboratories for helpful criticism, J. Hamroune and the Genom'ic platform for sequencing, A. Heurteau and Gen2i for help with bioinformatics, and L. Bluy for technical assistance. **Funding:** This work was supported by ARC, ANRS, MSD Avenir (Hide, Inflame&Seq), and European Research Council (CoG RNAmedTGS) to R.K.; ARC (4027) to S.R.; and Fondation pour la Recherche Medicale (FRM DEQ20160334940) to O.C. **Author contributions:** K.S., P.B., G.G., M.H., V.M., C.F., X.C., and S.R performed experiments and analyzed data. C.B. and D.D. analyzed data. R.K. and S.R. designed the experiments. O.C. designed data analysis. R.K. wrote the manuscript with input from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. ChIP-seq data have been deposited at GEO (GSE 143591). Additional data related to this paper may be requested from the authors.

Submitted 14 February 2020
Accepted 31 March 2021
Published 21 May 2021
10.1126/sciadv.abb2947

Citation: K. Salifou, C. Burnard, P. Basavarajaiah, G. Grasso, M. Helsmoortel, V. Mac, D. Depierre, C. Franckhauser, E. Beyne, X. Contreras, J. Dejardin, S. Rouquier, O. Cuvier, R. Kiernan, Chromatin-associated MRN complex protects highly transcribing genes from genomic instability. *Sci. Adv.* **7**, eabb2947 (2021).

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

Science Advances

Chromatin-associated MRN complex protects highly transcribing genes from genomic instability

Kader Salifou, Callum Burnard, Poornima Basavarajaiah, Giuseppa Grasso, Marion Helsmoortel, Victor Mac, David Depierre, Céline Franckhauser, Emmanuelle Beyne, Xavier Contreras, Jérôme Dejardin, Sylvie Rouquier, Olivier Cuvier and Rosemary Kiernan

Sci Adv **7** (21), eabb2947.
DOI: 10.1126/sciadv.abb2947

ARTICLE TOOLS http://advances.sciencemag.org/content/7/21/eabb2947

SUPPLEMENTARY MATERIALS http://advances.sciencemag.org/content/suppl/2021/05/17/7.21.eabb2947.DC1

REFERENCES This article cites 42 articles, 11 of which you can access for free
http://advances.sciencemag.org/content/7/21/eabb2947#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Downloaded from http://advances.sciencemag.org/ on June 9, 2021

Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/7/21/eabb2947/DC1

Supplementary Materials for

Chromatin-associated MRN complex protects highly transcribing genes from genomic instability

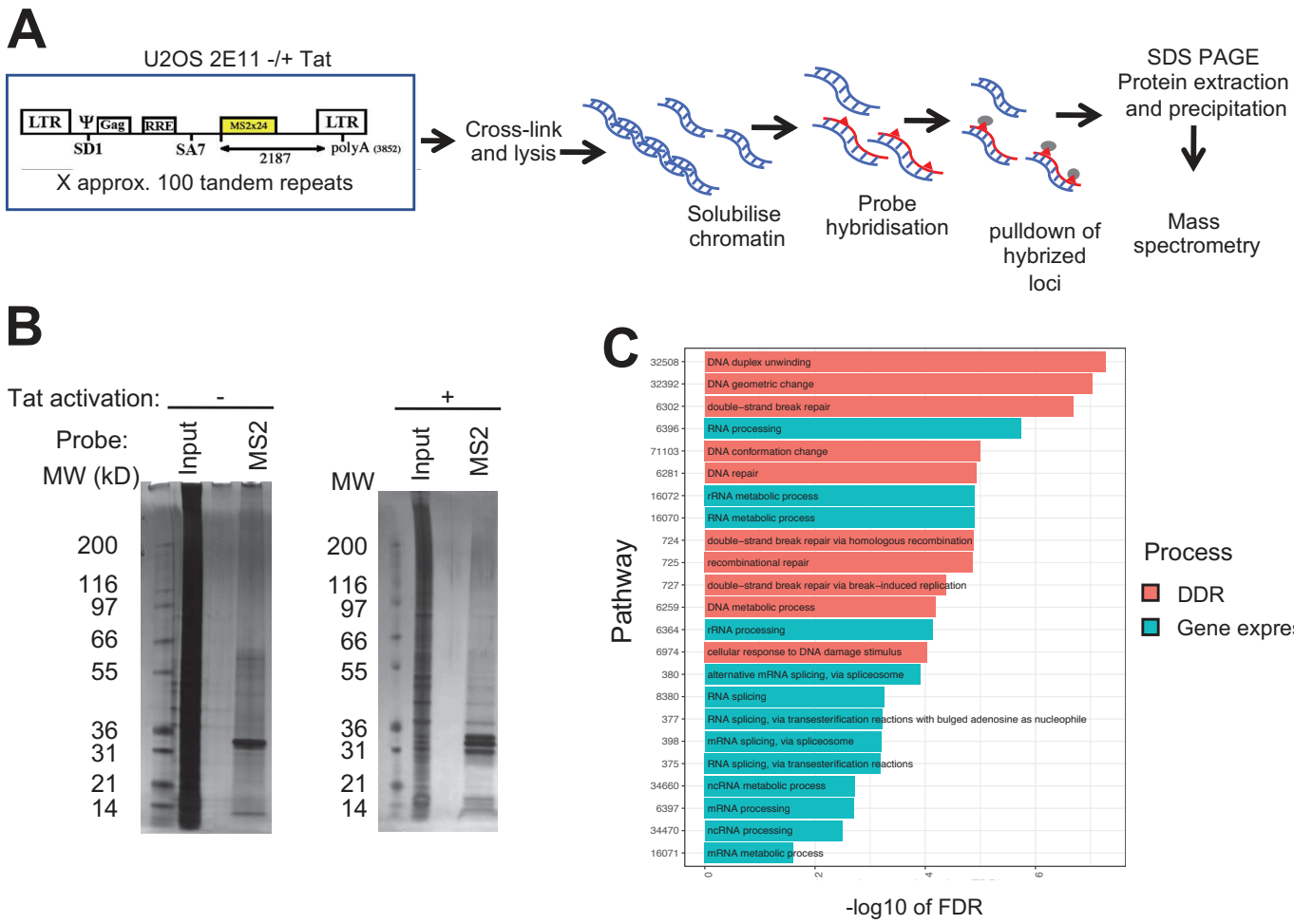
Kader Salifou, Callum Burnard, Poornima Basavarajaiah, Giuseppa Grasso, Marion Helsmoortel, Victor Mac, David Depierre, Céline Franckhauser, Emmanuelle Beyne, Xavier Contreras, Jérôme Dejardin, Sylvie Rouquier, Olivier Cuvier, Rosemary Kiernan*

*Corresponding author. Email: rosemary.kiernan@igh.cnrs.fr

Published 21 May 2021, *Sci. Adv.* **7**, eabb2947 (2021)
DOI: 10.1126/sciadv.abb2947

This PDF file includes:

Figs. S1 to S6
Tables S1 to S5



Supplementary Figure S1.

Fig. S1. DDR factors are associated with transcriptional activation.

A. Schematic representation of PICh experiment.

B. Silver stained SDS-PAGE gel of eluates obtained by PICh from cells treated with Tat or nontreated, as indicated.

C. GO analysis of proteins detected by PICh that were highly gained upon transactivation (+ Tat) of the HIV-1 LTR.

D. DDR factors associate with the HIV-1 LTR-reporter minigene following transcriptional activation. HeLa-LTR-luc cells, transfected with Tat or an empty vector as indicated, were analyzed by ChIP-qPCR for association of DDR factors with the promoter-proximal region of LTR-luc. Results are shown relative to the non-activated condition. Data represent mean ± SEM obtained from 3 independent experiments (***$P < 0.001$, *$P < 0.05$, NS indicates not significant, independent Student's *t* test).

E. γH2AX enrichment at HIV-1 LTR following transcriptional activation. HeLa-LTR-luc cells, transfected with Tat or an empty vector, as indicated, were analyzed by ChIP-qPCR for enrichment of γH2AX at the luciferase region. Results are shown in % of Input. Data represent mean ± SEM obtained from 3 independent experiments (***P < 0.001, *P < 0.05, NS indicates not significant, independent Student's t test).

F. P-TEFb subunit, cyclin T1 (CCNT1) interacts with DDR factors. Nuclear extracts of HEK293T cells, in the absence of HIV-1 Tat expression, were immunoprecipitated using anti-CCNT1 or control IgG. Immunoprecipitates and an aliquot of extract (input) were analyzed by Western blotting using the antibodies indicated.

G. Glycerol gradient analysis (10%-30%) of cell extracts described in (F) to separate the small active and large, inactive complexes of P-TEFb. TCA precipitates of the indicated fractions were analyzed by Western blotting using the indicated antibodies.

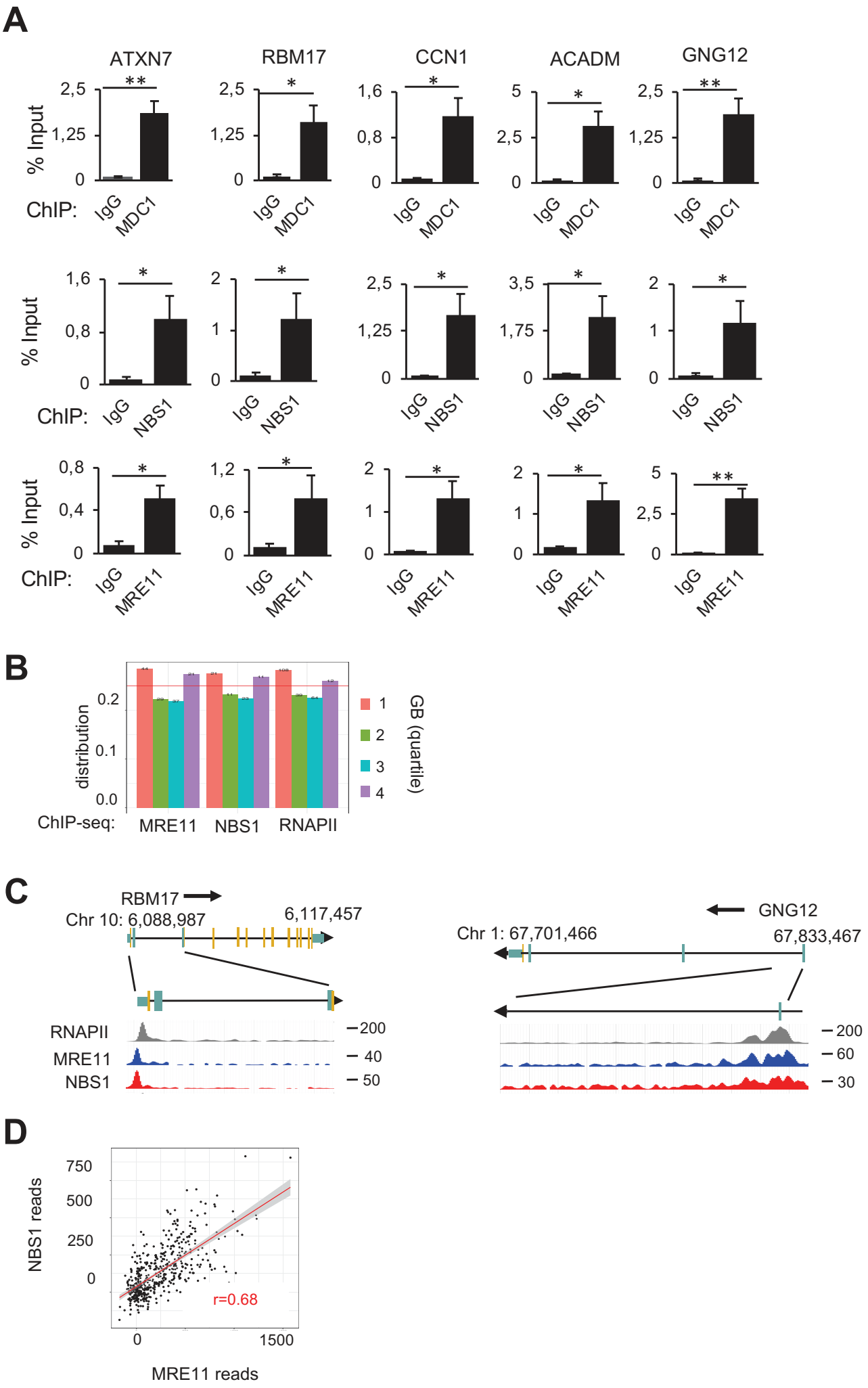

A
ATXN7
RBM17
CCN1
ACADM
GNG12
% Input
ChIP:
IgG
MDC1
NBS1
MRE11
B
distribution
GB (quartile)
ChIP-seq:
MRE11
NBS1
RNAPII
C
RBM17
Chr 10: 6,088,987
6,117,457
GNG12
Chr 1: 67,701,466
67,833,467
RNAPII
MRE11
NBS1
D
NBS1 reads
MRE11 reads
r=0.68

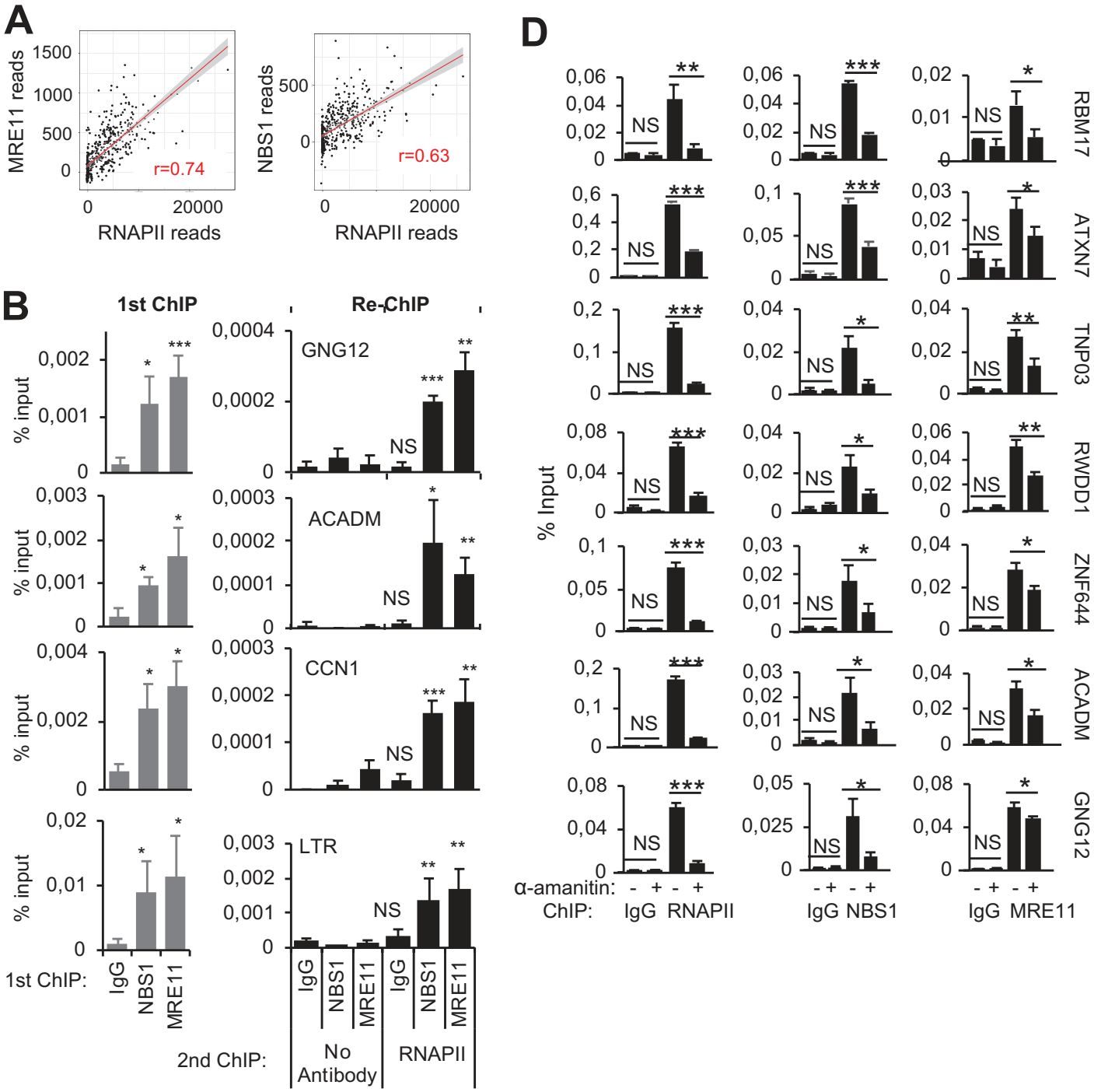
Fig. S2. DDR factors associate on chromatin.

A. DDR factors associate with TSS of actively transcribed genes. ChIP-qPCR of MDC1, NBS1 and MRE11 at the TSS of selected target genes. Results are shown in % of Input. Data represent mean ± SEM obtained from 3 independent experiments (***P < 0.001, *P < 0.05, NS indicates not significant, independent Student's t test).

B. ChIP-seq reads for MRE11, NBS1 or RNAPII were counted across gene body (GB) quartiles, as indicated. The red line represents 25%. The numbers above each column show the p value (as 1^{e}-) for the difference from the expected value (25%).

C. Browser shots of RNAPII, MRE11 and NBS1 ChIP-seq reads showing the TSS region of selected genes. Schematic representations of the gene and the zoomed region are shown above.

D. Scatter plot showing ChIP-seq reads of MRE11 and NBS1 at target genes. The co-efficient of correlation (r) is indicated on the graph.



Supplementary Figure S3.

Fig. S3. Association of MRN with genes correlates with that of RNAPII.

A. Scatter plot showing ChIP-seq reads of MRE11 and RNAPII (left) or NBS1 and RNAPII (right) at target genes. The co-efficient of correlation (r) calculated is indicated on the graph.

B. Re-ChIP analysis showing co-localization of RNAPII and MRN subunits on chromatin. Eluates from an initial ChIP using NBS1, MRE11 or a control IgG antibody (Ist ChIP) were used for RNAPII ChIP or no antibody as a control (re-ChIP). Results of qPCR after 1st and 2nd ChIPs are shown in % of input for 1st ChIP. Data represent mean ± SEM obtained from 3 independent experiments (***P < 0.001, *P < 0.05, NS indicates not significant, independent Student's t test).

C. Genes having variable amounts of RNAPII at the TSS were analyzed by ChIP-qPCR for binding of MRN subunits, MRE11 and NBS1. Binding, measured as % input, was compared to an IgG negative control. Data represent mean ± SEM obtained from 3 independent experiments (****P* < 0.001, **P* < 0.05, independent Student's *t* test).

D. Binding of MRE11, NBS1 or RNAPII was analyzed by ChIP-qPCR at the TSS region of target genes, indicated at right of the figure, from cells treated or not with a-amanitin, as indicated. Data represent mean ± SEM obtained from 3 independent experiments (***P < 0.001, **P < 0.01, *P < 0.05, NS indicates not significant, independent Student's t test).

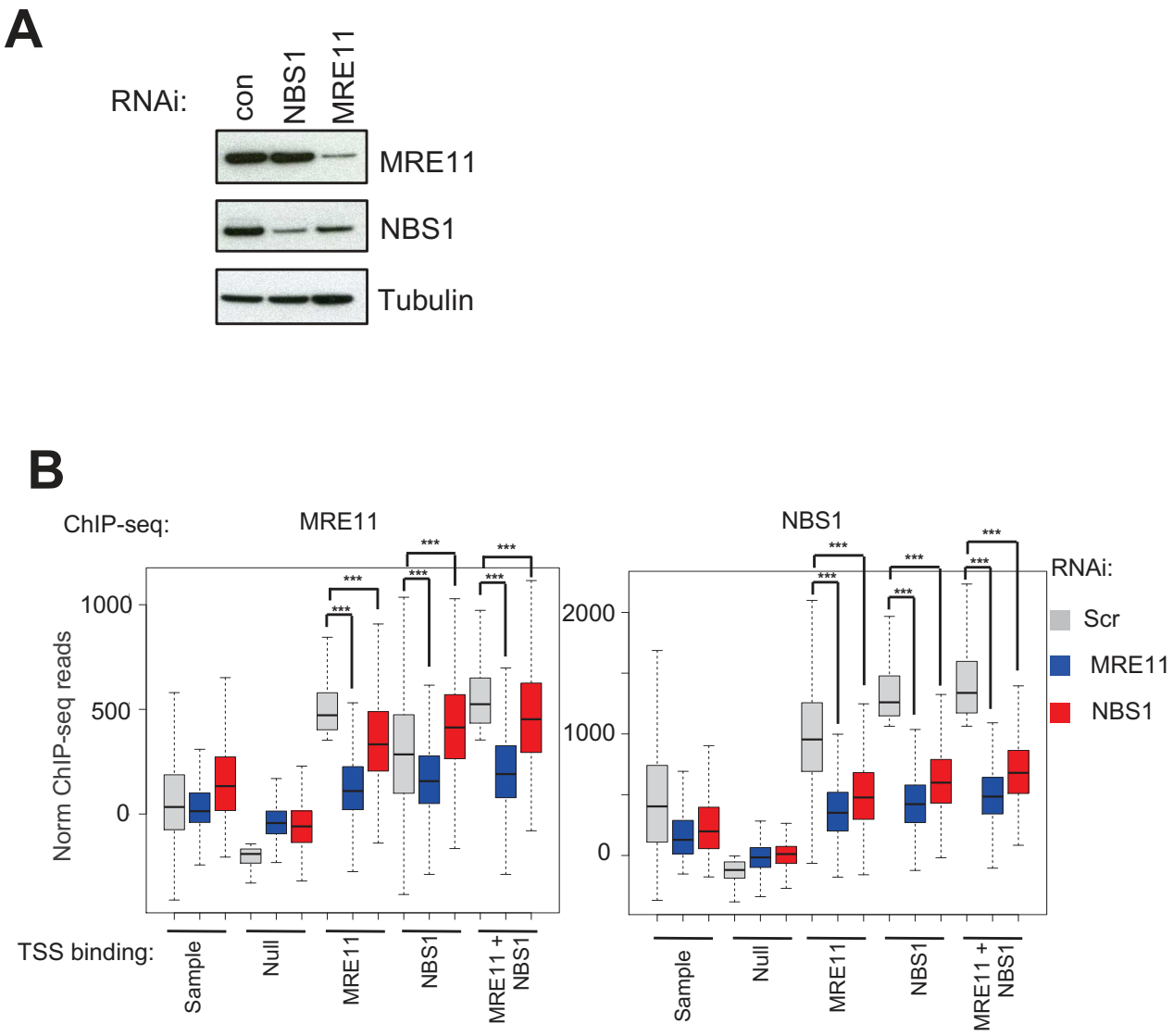
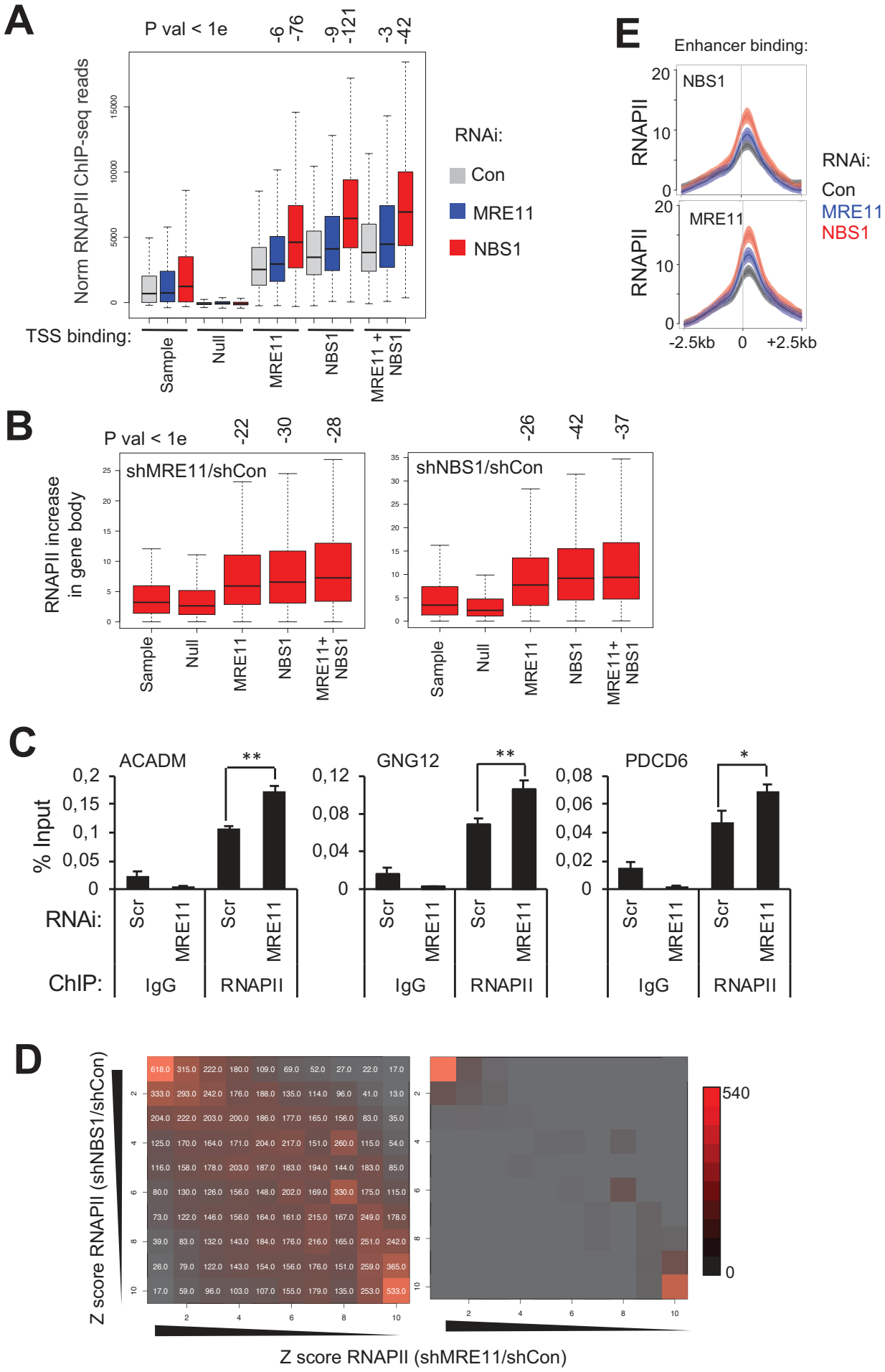


Fig. S4. RNAi against MRE11 or NBS1 induces depletion of MRN from target genes.

A. Whole cell extracts from cells treated with shRNAs targeting MRE11, NBS1 or a nontargeting control, as indicated, were analysed by Western blot using the indicated antibodies.

B. Box plots of ChIP-seq reads of MRE11 (left box plot) or NBS1 (right box plot) at genes highly bound at the TSS by MRE11, NBS1, both or neither (null), compared to a sample group of genes, following knock-down of MRE11 or NBS1, or a control knock-down, as indicated (***, $p < 0.001$, Wilcoxon test).



Supplementary Figure S5.

Fig. S5. Binding of MRN Impacts the Transcriptional Output of Target Genes

A. Box plots of RNAPII ChIP-seq reads at the TSS of genes highly bound by MRE11, NBS1, both or neither (null), compared to a sample group of genes, following knock-down of MRE11 or NBS1 or a control knock-down, as indicated. P-values (Wilcoxon test) are shown above the boxes.

B. Box plots of RNAPII increase as z-scores over genes highly bound at the TSS by MRE11, NBS1, both or neither (null), compared to a sample group of genes, following knock-down of MRE11 or NBS1, relative to a control knock-down. P-values (Wilcoxon test) are shown above the boxes.

C. Binding of RNAPII was analyzed by ChIP-qPCR at the TSS region of target genes, following downregulation of MRE11, as indicated. Data represent mean ± SEM obtained from 3 independent experiments (***P < 0.001, **P < 0.01, *P < 0.05, NS indicates not significant, independent Student's t test).

D. Matrix showing RNAPII increase as Z-scores at the TSS of genes following knock-down of MRE11 or NBS1, relative to a control knock-down, as indicated. Boxes in the left matrix indicate the number of genes in each decile. The right matrix indicates the P-value (Hypergeometric test).

E. Average density profile of RNAPII ChIP-seq reads at enhancers bound by MRE11 or NBS1, as indicated, following depletion of MRE11, NBS1 or a control knock-down, as indicated.

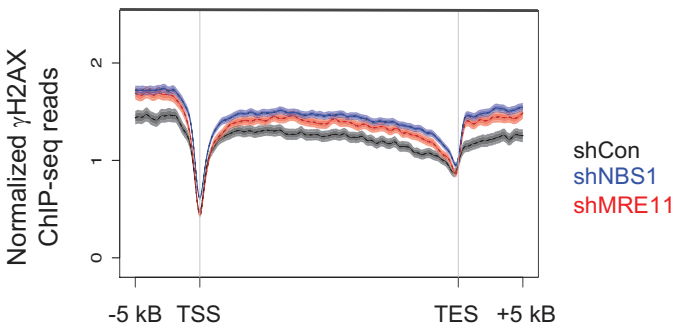


Fig. S6. Loss of MRN is associated with increased γH2AX association at MRN-bound genes

Average density profile of γH2AX ChIP-seq reads at MRN-bound active genes following depletion of MRE11, NBS1 or a control knock-down, as indicated.

Table S1. Identification of proteins whose association with the HIV-1 minigene was significantly changed in the presence Tat transcriptional activator compared to the nontreated sample.

UniProt ID	logFC	UniProt ID	logFC	UniProt ID	logFC	UniProt ID	logFC
RL1D1_HUMAN	3,61	PUF60_HUMAN	2,08	IMA2_HUMAN	1,79	PDIA4_HUMAN	1,61
HP1B3_HUMAN	3,47	G6PI_HUMAN	2,08	MECP2_HUMAN	1,79	SPAS2_HUMAN	1,61
RPA34_HUMAN	3,33	MOGS_HUMAN	2,08	FUBP1_HUMAN	1,79	SYK_HUMAN	1,61
MDC1_HUMAN	3,28	CALX_HUMAN	2,08	RFC4_HUMAN	1,79	SSRP1_HUMAN	1,61
XRCC5_HUMAN	3,22	HSP72_HUMAN	2,08	CCD47_HUMAN	1,79	FUS_HUMAN	1,61
HNRPM_HUMAN	3,22	MCM7_HUMAN	2,08	SYDC_HUMAN	1,79	CX067_HUMAN	1,61
TKT_HUMAN	3,18	PA2G4_HUMAN	2,08	COX2_HUMAN	1,79	HAT1_HUMAN	1,61
AHNK_HUMAN	3,09	GRP75_HUMAN	2,05	HSP7C_HUMAN	1,70	AATM_HUMAN	1,61
NOP56_HUMAN	3,09	LDHB_HUMAN	2,01	DHX9_HUMAN	1,70	VRK1_HUMAN	1,61
MCM6_HUMAN	2,94	ENPL_HUMAN	1,95	ATPA_HUMAN	1,70	NCLN_HUMAN	1,61
H90B3_HUMAN	2,94	SYEP_HUMAN	1,95	NPM_HUMAN	1,70	RBBP4_HUMAN	1,61
HS90A_HUMAN	2,83	BOP1_HUMAN	1,95	HSP71_HUMAN	1,67	QCR2_HUMAN	1,61
XRCC6_HUMAN	2,77	MCM4_HUMAN	1,95	SYIC_HUMAN	1,61	PRS6A_HUMAN	1,61
MOES_HUMAN	2,77	UBP5_HUMAN	1,95	ACTN1_HUMAN	1,61	HNRDL_HUMAN	1,61
MCM5_HUMAN	2,64	EIF3C_HUMAN	1,95	PSIP1_HUMAN	1,61	ROA1_HUMAN	1,61
HNRPR_HUMAN	2,64	SRSF3_HUMAN	1,95	HMGB1_HUMAN	1,61	HMGA2_HUMAN	1,61
DDX17_HUMAN	2,64	IMB1_HUMAN	1,95	RPA1_HUMAN	1,61	RS15_HUMAN	1,61
PHB2_HUMAN	2,64	NAT10_HUMAN	1,95	RRP5_HUMAN	1,61	DEK_HUMAN	1,58
TFR1_HUMAN	2,56	TSR1_HUMAN	1,95	HEAT1_HUMAN	1,61	HNRPD_HUMAN	1,54
DDX21_HUMAN	2,48	ILF2_HUMAN	1,95	TRIPC_HUMAN	1,61	NUCL_HUMAN	1,53
TIF1B_HUMAN	2,48	DDX18_HUMAN	1,95	E7ETY2_HUMAN	1,61	P5CS_HUMAN	1,50
DDX27_HUMAN	2,48	LKHA4_HUMAN	1,95	FND3B_HUMAN	1,61	HSP74_HUMAN	1,50
RECQ1_HUMAN	2,48	ZN512_HUMAN	1,95	STT3B_HUMAN	1,61	HDGF_HUMAN	1,50
LMNB2_HUMAN	2,47	DDX1_HUMAN	1,95	LRRF1_HUMAN	1,61	RS2_HUMAN	1,47
SFPQ_HUMAN	2,40	NOG1_HUMAN	1,95	SMHD1_HUMAN	1,61	BAF_HUMAN	1,47
KI67_HUMAN	2,30	FBRL_HUMAN	1,95	NOL6_HUMAN	1,61	GRP78_HUMAN	1,46
RRMJ3_HUMAN	2,30	TGM2_HUMAN	1,95	NOMO1_HUMAN	1,61	MCM2_HUMAN	1,45
RBM39_HUMAN	2,30	ABCE1_HUMAN	1,95	ACLY_HUMAN	1,61	LDHA_HUMAN	1,45
DDX3X_HUMAN	2,30	EIF3L_HUMAN	1,95	SMCA1_HUMAN	1,61	LAP2A_HUMAN	1,42
KV309_HUMAN	2,30	DNJC9_HUMAN	1,95	WDR12_HUMAN	1,61	SP16H_HUMAN	1,39
TCPB_HUMAN	2,30	LA_HUMAN	1,95	MYO1C_HUMAN	1,61	HS90B_HUMAN	1,39
DDX5_HUMAN	2,30	LAP2B_HUMAN	1,95	SUN2_HUMAN	1,61	ALDOC_HUMAN	1,39
MCM3_HUMAN	2,25	THOC4_HUMAN	1,95	DCTN2_HUMAN	1,61	TOP1M_HUMAN	1,39
ESYT1_HUMAN	2,20	NOP2_HUMAN	1,92	CTNA1_HUMAN	1,61	H2AW_HUMAN	1,39
DDB1_HUMAN	2,20	LYRIC_HUMAN	1,90	STAT1_HUMAN	1,61	RCC1_HUMAN	1,39
IMMT_HUMAN	2,20	NUMA1_HUMAN	1,87	HDGR2_HUMAN	1,61	APEX1_HUMAN	1,39
STIP1_HUMAN	2,20	UBA1_HUMAN	1,87	UHRF1_HUMAN	1,61	CHD4_HUMAN	1,39
ASNS_HUMAN	2,20	TOP2A_HUMAN	1,79	UBP2L_HUMAN	1,61	SPT6H_HUMAN	1,39
MK67I_HUMAN	2,20	FLNB_HUMAN	1,79	HNRPF_HUMAN	1,61	KIF4A_HUMAN	1,39
TM214_HUMAN	2,20	GCN1L_HUMAN	1,79	CMC2_HUMAN	1,61	MSH6_HUMAN	1,39
YBOX1_HUMAN	2,20	SYLC_HUMAN	1,79	PUR9_HUMAN	1,61	MBB1A_HUMAN	1,39
GUAA_HUMAN	2,20	MAGD2_HUMAN	1,79	NSUN2_HUMAN	1,61	BAZ2A_HUMAN	1,39
HNRPQ_HUMAN	2,20	ACTN4_HUMAN	1,79	EWS_HUMAN	1,61	NOLC1_HUMAN	1,39
HNRCL_HUMAN	2,20	PABP1_HUMAN	1,79	VPS35_HUMAN	1,61	ASPH_HUMAN	1,39
LMNB1_HUMAN	2,13	ITB1_HUMAN	1,79	FUBP2_HUMAN	1,61	TTF1_HUMAN	1,39
CKAP4_HUMAN	2,12	DLDH_HUMAN	1,79	RCC2_HUMAN	1,61	FXR1_HUMAN	1,39
BAZ1B_HUMAN	2,08	ILF3_HUMAN	1,79	D6RBZ0_HUMAN	1,61	PWP2_HUMAN	1,39
LPPRC_HUMAN	2,08	DHX15_HUMAN	1,79	DPYL3_HUMAN	1,61	XPC_HUMAN	1,39
TBL3_HUMAN	2,08	SYQ_HUMAN	1,79	GNL3_HUMAN	1,61	WDR36_HUMAN	1,39

UniProt ID	logFC	UniProt ID	logFC	UniProt ID	logFC	UniProt ID	logFC
MAN1_HUMAN	1,39	UBC9_HUMAN	1,39	EIF3B_HUMAN	0,98	CPNE3_HUMAN	-0,92
COPG1_HUMAN	1,39	FLNC_HUMAN	1,36	TCP4_HUMAN	0,98	RL15_HUMAN	-0,92
ACSL3_HUMAN	1,39	COF1_HUMAN	1,34	ARF1_HUMAN	0,98	ACTB_HUMAN	-0,99
XPO1_HUMAN	1,39	MAP1B_HUMAN	1,34	TBB1_HUMAN	0,96	CATA_HUMAN	-1,10
MVP_HUMAN	1,39	LMNA_HUMAN	1,32	HNRH1_HUMAN	0,96	NCRP1_HUMAN	-1,10
ABCF1_HUMAN	1,39	CLIC1_HUMAN	1,30	FEN1_HUMAN	0,96	ECHB_HUMAN	-1,39
XRN2_HUMAN	1,39	SRRT_HUMAN	1,25	HMGB3_HUMAN	0,96	TGM5_HUMAN	-1,39
U5S1_HUMAN	1,39	SMCA5_HUMAN	1,25	HMGB2_HUMAN	0,96	ADRO_HUMAN	-1,39
NCKP1_HUMAN	1,39	RS6_HUMAN	1,25	EZRI_HUMAN	0,94	DHCR7_HUMAN	-1,39
MATR3_HUMAN	1,39	TBA3C_HUMAN	1,25	RSSA_HUMAN	0,94	DPM1_HUMAN	-1,39
ELYS_HUMAN	1,39	DBPA_HUMAN	1,25	H2AY_HUMAN	0,93	BAP31_HUMAN	-1,39
DDX56_HUMAN	1,39	SERPH_HUMAN	1,25	PTN1_HUMAN	0,92	RU2B_HUMAN	-1,39
TBR1_HUMAN	1,39	HS105_HUMAN	1,25	TBB4B_HUMAN	0,92	RL35A_HUMAN	-1,39
PSMD2_HUMAN	1,39	LASP1_HUMAN	1,25	BUB3_HUMAN	0,92	RL34_HUMAN	-1,39
NOL11_HUMAN	1,39	RL6_HUMAN	1,25	VDAC2_HUMAN	0,92	SMD1_HUMAN	-1,39
ECHA_HUMAN	1,39	HNRPU_HUMAN	1,24	AL7A1_HUMAN	0,92	PYC_HUMAN	-1,61
PESC_HUMAN	1,39	TBB5_HUMAN	1,22	PGAM1_HUMAN	0,92	KV201_HUMAN	-1,61
SRPR_HUMAN	1,39	NOP58_HUMAN	1,22	RS26_HUMAN	0,92	2AAA_HUMAN	-1,61
SYRC_HUMAN	1,39	FLNA_HUMAN	1,21	NACA_HUMAN	0,92	GP179_HUMAN	-1,61
E2AK2_HUMAN	1,39	MYH9_HUMAN	1,20	RL8_HUMAN	0,92	RUVB2_HUMAN	-1,61
NPL4_HUMAN	1,39	TLN1_HUMAN	1,20	DDX6_HUMAN	0,92	H14_HUMAN	-1,61
TCPE_HUMAN	1,39	RL13_HUMAN	1,20	CNDP2_HUMAN	0,92	FABP5_HUMAN	-1,70
GFPT1_HUMAN	1,39	DDX47_HUMAN	1,20	GDIR1_HUMAN	0,92	ACTBL_HUMAN	-1,79
RFA1_HUMAN	1,39	SNUT1_HUMAN	1,20	RS3A_HUMAN	0,92	POF1B_HUMAN	-1,79
STRAP_HUMAN	1,39	RL19_HUMAN	1,20	RS29_HUMAN	0,92	DHRS2_HUMAN	-1,79
ROA3_HUMAN	1,39	RL7_HUMAN	1,20	CALM_HUMAN	0,92	ACTBM_HUMAN	-1,95
ELAV1_HUMAN	1,39	CG050_HUMAN	1,20	AK1A1_HUMAN	0,92	TRFL_HUMAN	-2,20
PAL4A_HUMAN	1,39	PARP1_HUMAN	1,19	ANXA5_HUMAN	0,92	HUTH_HUMAN	-2,40
G3PT_HUMAN	1,39	RBMX_HUMAN	1,18	RHOA_HUMAN	0,92	FRPD4_HUMAN	-2,48
CAP1_HUMAN	1,39	H90B2_HUMAN	1,16	RS23_HUMAN	0,92		
NP1L1_HUMAN	1,39	HGB1A_HUMAN	1,15	PLEC_HUMAN	0,91		
IMP3_HUMAN	1,39	ROA2_HUMAN	1,15	VIME_HUMAN	0,91		
ACTZ_HUMAN	1,39	CLH1_HUMAN	1,10	ENOA_HUMAN	0,90		
BCCIP_HUMAN	1,39	TOP1_HUMAN	1,10	RS18_HUMAN	0,88		
PLIN3_HUMAN	1,39	LBR_HUMAN	1,10	PRKDC_HUMAN	0,86		
CDC37_HUMAN	1,39	PEBP1_HUMAN	1,10	RL10_HUMAN	0,85		
METK2_HUMAN	1,39	RPN2_HUMAN	1,10	TBB3_HUMAN	0,85		
EIF3E_HUMAN	1,39	G3BP1_HUMAN	1,10	SERA_HUMAN	0,85		
FCL_HUMAN	1,39	TPIS_HUMAN	1,10	RBM8A_HUMAN	0,85		
EBP2_HUMAN	1,39	RL18_HUMAN	1,10	TBB4A_HUMAN	0,83		
DHSO_HUMAN	1,39	DNJB1_HUMAN	1,10	RL3_HUMAN	0,83		
BIEA_HUMAN	1,39	IF2A_HUMAN	1,10	H2B1A_HUMAN	0,82		
RLA0L_HUMAN	1,39	HNRH3_HUMAN	1,10	PUR6_HUMAN	0,81		
H2B2C_HUMAN	1,39	H2AV_HUMAN	1,10	CBX1_HUMAN	0,79		
RANG_HUMAN	1,39	HMGN2_HUMAN	1,10	ATPB_HUMAN	0,79		
NTPCR_HUMAN	1,39	GBLP_HUMAN	1,06	PHB_HUMAN	0,79		
HTR5A_HUMAN	1,39	DKC1_HUMAN	1,03	LRC59_HUMAN	0,77		
DUT_HUMAN	1,39	PDIA6_HUMAN	1,03	CDSN_HUMAN	0,77		
SKP1_HUMAN	1,39	CBX3_HUMAN	1,03	RS3_HUMAN	0,74		
PPAC_HUMAN	1,39	HNRPK_HUMAN	1,02	RPN1_HUMAN	0,73		
DUS3_HUMAN	1,39	RS8_HUMAN	1,01	RS9_HUMAN	-0,81		
ARPC4_HUMAN	1,39	STT3A_HUMAN	0,98	ZCH23_HUMAN	-0,85		
TMEDA_HUMAN	1,39	PPM1G_HUMAN	0,98	PRDX2_HUMAN	-0,92		

Table S2. Identification of proteins interacting with endogenous MDC1 by tandem affinity purification and mass spectrometry. Results are shown as LogFE, (log fold enrichment) of the number of peptides found in MDC1 sample compared to the control sample.

ID	logFE	ID	logFE	ID	logFE	ID	logFE	ID	logFE
MDC1	4,68	RPL6	1,79	SRRM1	1,39	ZNF326	1,10	RPL19	0,69
RAD50	4,26	SRSF1	1,79	RBM8A	1,39	SNRPD3	1,10	GRWD1	0,69
CD2AP	3,69	FAM175A	1,79	DDX1	1,39	PSMA5	1,10	ALDOC	0,69
NBN	3,56	USP28	1,79	ADH5	1,39	TUBAL3	1,10	SRSF5	0,69
MRE11A	3,47	SLTM	1,79	NUMA1	1,39	HSP90AA5P	1,10	NEFM	0,69
TRIP12	3,26	RPS16	1,79	NCOA5	1,39	HMGB2	1,10	RPL34	0,69
ANAPC1	3,22	KPNA2	1,79	SNRNP40	1,39	TRA2A	1,10	NUCKS1	0,69
SH3KBP1	3,14	PPP1CA	1,79	HNRNPUL2	1,39	RBM17	1,10	YBX3	0,69
PRPF8	3,09	BRE	1,79	THRAP3	1,39	ACTN1	1,10	PRPS1	0,69
TEX10	3,04	PNP	1,79	SNRPA1	1,39	RBBP4	1,10	PREP	0,69
TP53BP1	3,00	ZFR	1,79	RPS13	1,39	GAR1	1,10	CCT7	0,69
PELP1	2,89	HNRNPH3	1,79	NME1	1,39	AMOTL1	1,10	HIST1H2AB	0,69
SF3B1	2,89	XRN2	1,79	ILF3	1,34	GLO1	1,10	RBM39	0,69
SNRNP200	2,89	PABPC1	1,79	ILF2	1,32	EIF4A1	1,10	PDHB	0,69
YLPM1	2,77	PARP1	1,79	NPM1	1,30	HSPH1	1,10	HIST1H2BB	0,69
CDC27	2,77	RBM6	1,79	HNRNPL	1,25	TOP1	1,10	DNAJA1	0,69
CDC16	2,64	RPL7	1,79	ELAVL1	1,25	PRMT5	1,10	PRSS1	0,69
LAS1L	2,64	RPL4	1,79	XRCC6	1,25	MAGOHB	1,10	ACLY	0,69
DDX21	2,64	HNRNPR	1,61	CELSR3	1,10	SLC25A5	1,10	MFAP1	0,69
ANAPC5	2,64	SYNCRIP	1,61	ANP32E	1,10	ZNF638	1,10	RPS26	0,69
ANAPC7	2,64	RCN2	1,61	STAU1	1,10	RPL24	1,10	TUBA1B	0,69
ADAR	2,64	HNRNPD	1,61	HP1BP3	1,10	RPL27	1,10	KHDRBS1	0,69
AMOT	2,64	LTA4H	1,61	SRSF7	1,10	HSP90AB2P	1,10	RPL10	0,69
ANAPC4	2,56	HNRNPH2	1,61	PSMA2	1,10	RPL8	1,10	ACAT1	0,69
SRRM2	2,56	RBM14	1,61	YWHAZ	1,10	ATIC	1,10	DDB1	0,69
HNRNPM	2,56	SNRPD2	1,61	RPLP0	1,10	PCNA	1,10	PRPF4B	0,69
HNRNPC	2,48	PRRC2A	1,61	SAP18	1,10	HSPA5	1,05	CRNKL1	0,69
CAPZA1	2,48	DKC1	1,61	PGD	1,10	DHX15	1,03	SRSF3	0,69
IGF2BP1	2,48	IGF2BP3	1,61	CDC5L	1,10	CAPRIN1	0,98	NACA	0,69
UIMC1	2,40	RAN	1,61	SRSF10	1,10	PSAT1	0,98	APOA1	0,69
SENP3	2,40	SRSF4	1,61	RNPS1	1,10	SF3B2	0,98	PSMB4	0,69
RALY	2,30	HNRNPCL3	1,50	PTBP1	1,10	HNRNPUL1	0,96	RPL31	0,69
EFTUD2	2,30	PDIA3	1,39	RPLP0P6	1,10	RPL7A	0,92	RPS17L	0,69
BRCC3	2,30	MTHFD1	1,39	EEF1B2	1,10	ALYREF	0,92	PRPF6	0,69
MATR3	2,30	HNRNPK	1,39	PSMB3	1,10	HDGF	0,92	CWC22	0,69
MDN1	2,30	WDR1	1,39	H1FX	1,10	SF3A1	0,92	YWHAQ	0,69
ANAPC2	2,30	HNRNPA1	1,39	EEF1D	1,10	GDI2	0,92	CCT6A	0,69
SF3B3	2,25	GDI1	1,39	NAP1L1	1,10	RPL9	0,92	HSPE1	0,69
WDR18	2,20	ETFA	1,39	TARDBP	1,10	EZR	0,92	SHMT2	0,69
TP53	2,08	YWHAB	1,39	SNW1	1,10	U2SURP	0,85	IMPDH2	0,69
CAPZB	2,08	RPSA	1,39	PYGL	1,10	GAPDH	0,85	GARS	0,69
RPL3	2,08	CIRBP	1,39	PRKRA	1,10	RPS4X	0,85	ACTN4	0,69
CAPZA2	1,95	TRA2B	1,39	CLTC	1,10	TPI1	0,81	DLD	0,69
BABAM1	1,95	PRPF19	1,39	SNRPE	1,10	ANXA5	0,81	SSRP1	0,69
NOL9	1,95	APEX1	1,39	RBM10	1,10	HSPA9	0,79	CCT2	0,69
TJP1	1,95	XRCC5	1,39	RPL27A	1,10	DIAPH1	0,69	WDR61	0,69
CHERP	1,95	RPL15	1,39	HPRT1	1,10	KAT2A	0,69	RRP1B	0,69
USP7	1,95	ANAPC10	1,39	PSMB5	1,10	HSBP1	0,69	SF3B4	0,69
RSL1D1	1,95	NOP56	1,39	HIST1H4A	1,10	DCDC2C	0,69	OAT	0,69
CDC23	1,95	RPS15A	1,39	SUB1	1,10	TMC7	0,69	SHMT1	0,69

ID	logFE	ID	logFE	ID	logFE	ID	logFE	ID	logFE
RPN2	0,69	CWC22	0,69	NHP2	0,69	HDAC1	0,69	CKB	0,64
CCT3	0,69	YWHAQ	0,69	RPS10	0,69	CPNE3	0,69	AHCY	0,59
C21orf33	0,69	CCT6A	0,69	SKIV2L2	0,69	PSMB6	0,69	UBA1	0,59
ZCCHC8	0,69	HSPE1	0,69	LUC7L2	0,69	RPS11	0,69	DDX3X	0,56
RRP9	0,69	SHMT2	0,69	RDX	0,69	SUPT16H	0,69	SSB	0,56
RPL37A	0,69	IMPDH2	0,69	RNPEP	0,69	PCYOX1	0,69	HNRNPAB	0,56
FUBP3	0,69	GARS	0,69	PFAS	0,69	CS	0,69	CA2	0,56
GTPBP4	0,69	ACTN4	0,69	STRBP	0,69	SNRPB2	0,69	HSPA1L	0,51
TJP2	0,69	DLD	0,69	APEH	0,69	UHRF1	0,69	GTF2I	0,51
RPL23	0,69	SSRP1	0,69	BCAS2	0,69	CDH7	0,69	DHX9	0,51
RPS28	0,69	CCT2	0,69	OXCT1	0,69	ZNF346	0,69	TALDO1	0,51
GSTP1	0,69	WDR61	0,69	GRHPR	0,69	FADD	0,69	PA2G4	0,51
RPLP1	0,69	RRP1B	0,69	RPL5	0,69	PARP10	0,69	NASP	0,51
EIF6	0,69	SF3B4	0,69	SMNDC1	0,69	MAD2L1	0,69	SAFB2	0,51
GANAB	0,69	OAT	0,69	HMGB3	0,69	CLIC1	0,69	MDH1	0,51
SERBP1	0,69	SHMT1	0,69	NKRF	0,69	PLOD2	0,69	DDX5	0,49
STMN1	0,69	RPN2	0,69	SNRPB	0,69	FAM171A1	0,69	RBMX	0,47
RBM39	0,69	CCT3	0,69	PPAT	0,69	RPL36	0,69	HSPA4	0,47
PDHB	0,69	C21orf33	0,69	NUDT5	0,69	TRIO	0,69	TUBB2A	0,45
HIST1H2BB	0,69	ZCCHC8	0,69	PABPN1	0,69	E4F1	0,69	TKT	0,41
DNAJA1	0,69	RRP9	0,69	NCBP1	0,69	RSRC1	0,69	ACTB	0,41
PRSS1	0,69	RPL37A	0,69	PRPH	0,69	BRF1	0,69	HNRNPDL	0,41
ACLY	0,69	FUBP3	0,69	ARFIP1	0,69	SF1	0,69	RPS2	0,41
MFAP1	0,69	GTPBP4	0,69	CPSF1	0,69	ERH	0,69	ANP32A	0,41
RPS26	0,69	TJP2	0,69	FHL1	0,69	GAPDHS	0,69	HIST1H1A	0,41
TUBA1B	0,69	RPL23	0,69	GOT1	0,69	YBX1	0,69	YWHAE	0,41
KHDRBS1	0,69	RPS28	0,69	SYNE1	0,69	EWSR1	0,69	RPL18	0,41
RPL10	0,69	GSTP1	0,69	PTMS	0,69	GOT2	0,69	RPS14	0,41
ACAT1	0,69	RPLP1	0,69	HIST1H2AA	0,69	HNRNPF	0,69	RANBP1	0,41
DDB1	0,69	EIF6	0,69	FARSA	0,69	RPL12	0,69	TUBB3	0,41
PRPF4B	0,69	GANAB	0,69	PRMT1	0,69	PGAM1	0,69	NPEPPS	0,41
CRNKL1	0,69	STMN1	0,69	ESRRA	0,69	TUBB4A	0,69	CANX	0,41
SRSF3	0,69	PCBP1	0,69	DDX50	0,69	ACO2	0,69	PYGB	0,41
NACA	0,69	NANS	0,69	KHSRP	0,69	PPIB	0,69	PSMA6	0,41
APOA1	0,69	MSN	0,69	SNRPGP15	0,69	PHGDH	0,69	POTEKP	0,41
PSMB4	0,69	SPAG17	0,69	RPL13AP3	0,69	IK	0,69	PGM1	0,41
RPL31	0,69	EIF4A3	0,69	EIF4H	0,69	BDP1	0,69		
RPS17L	0,69	BSG	0,69	PRPF6	0,69	HNRNPU	0,64		

Table S3. Antibodies used in this study.

Antibody	Reference	Supplier
MDC1 (WB, IP)	A300-051A	Bethyl laboratories
MDC1 (ChIP)	A300-051A and A301-053A	Bethyl laboratories
MRE11 (ChIP, ChIP-seq)	A301-998A	Bethyl laboratories
NBS1 (ChIP)	A301-189A and A301-289A	Bethyl laboratories
NBS1 (ChIP-seq)	A301-289A	Bethyl laboratories
RNAPII (WB)	CTD4H8	Millipore
RNAPII (ChIP, ChIP-seq)	F-12, sc-55492	SCBT
γH2A.X (ChIP, ChIP-seq)	Ab2893	Abcam
RNAPII-ser5 (WB)	Ab5131	Abcam
RNAPII-ser2 (WB and ChIP)	Ab5095	Abcam
RNAPII-ser2 (ChIP-seq)	Ab5055	Abcam
SNRNP200	Sc-393170	SCBT
PRP8	Sc-55533	SCBT
HNRNPM	Sc-20002	SCBT
EFTUD2	A300-957A	Bethyl laboratories
CCNT1	A303-499A	Bethyl laboratories
CDK9	A303-493A	Bethyl laboratories
TUBULIN	DM1A clone, T6199	Sigma-Aldrich
Flag	M2 clone, F1804	Sigma-Aldrich
HA	Sc-7392	SCBT
Ku80 (ChIP)	Ab80592	Abcam
IgG (mouse)	Sc-2025	SCBT
IgG (rabbit)	Sc-2027	SCBT

Table S4. Sequences of deoxyoligonucleotides used in this study.

Primer	Forward (5’ to 3’)	Reverse (5’ to 3’)
ACADM	GAATACGCGGAAAGGTCAAA	TGCTCTCAAATGCCCTACCT
ATXN7	TACTTCGTCCTGACACCTGG	AAAGCATGAATGGGAAGCCG
Bambi	GGATTGACAGGAGCCCGG	CTCCGCGCCGTAAATAGC
CCN1	CCCCGAACTGTTTTGTTGAC	TCCATTGCACCTTTGTGTGT
FOS_TSS	ACTACCACTCACCCGCAGAC	GGGAGCCCCCTACTCATCTA
FOS_MID	CAAGCCCATTCCATCCCAAC	AGAGCCCAGTCTGAAAGCAT
GAPDH	CACATCGCTCAGACACCAT	GAGGTCAATGAAGGGGTCAT
GNG12	GTGGACGTCTGGTCGTCTTT	GAAACTTGCGAGGAGTCAGG
HIV-1 LTR	GGGTCTCTCTGGTTAGA	GGGTTCCCTAGTTAGGC
HIV-1_MID	ACCCTATTTTCATTCTTCGCC	AGCCCATATCCTTGTCGTATC
HSPA1A_TSS	CTGCTGCGACAGTCCACTAC	CTGGAAACGGAACACTGGAT
HSP1A_MID	AGATTCACGACCTGGTCCTG	CGGGGTTGATGCTCTTGTTC
HSP1B_TSS	CCGACTGGAGGAGCATAAAA	ATGCTGGCTGACTCTGCTCT
HSP1B_MID	AGATTCACGACCTGGTCCTG	CGGGGTTGATGCTCTTGTTC
HSPD1_TSS	GGTTCCCTCACGTACACCAC	ACTCGTGGGGTTTGAATTTG
HSPD1_MID	TTGATCACAGCCAAGGGTAGT	GATTCTCCCACCTCAGCCTC
PDCD6	GGCAAACTAAGTCCGTGCTC	GAAAGCGCTTGGAGAAAATG
RBM17	TCCCACTTTCCCTTTCCCTC	CAAGGAGCACGAGGTTAAGC
RWDD1	TGTAGGTGTCCTTGGCGAAT	TCCCTACTTCCACGTTCGAC
TNP03	CTGTTCCTCGCTAGCAGCTT	AGTCGGAAGGGGTAGGAAAA
ZNF644	GGGCCAATTAGACAAACGAG	GGAACCTCCTTGGGAGTGAT
ssODN MDC1	GATGCTGTTTGTGTGAACCAATAATGAGGATAATTGATAATGTGTATCCTTCCCAGATCATGCAG GACTACAAGGACGACGATGACAAGCTCGATGGAGGATACCCCTACGACGTGCCCGACTACGCCG GAGGACTCGAGGAGGACACCCAGGCTATTGACTGGGATGTTGAAGAAGAGGAGGAGACAGAGC AATCCAGT	
PICh probe	ATGTCGACCTGCAGACATGGG (MS2.1)	
PICh probe	TGATCCTCATGTTTTCTAGGCAATTA (MS2.2)	

Table S5. Double stranded shRNAs used in this study.

shRNA	Sequence (5’ to 3’)
Con	CCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGG
MRE11	CCGGACGACTGCGAGTGGACTATAGCTCGAGCTATAGTCCACTCGCAGTCGTTTTTTG
NBS1	CCGGGCAAGCAGATACATGGGATTTCTCGAGAAATCCCATGTATCTGCTTGCTTTTTG

Bibliography

Abbott, A. L., Alvarez-Saavedra, E., Miska, E. A., Lau, N. C., Bartel, D. P., Horvitz, H. R., & Ambros, V. (2005). The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. *Developmental Cell*, *9*(3), 403–414. https://doi.org/10.1016/j.devcel.2005.07.009

Acunzo, M., Romano, G., Wernicke, D., & Croce, C. M. (2015). MicroRNA and cancer – A brief overview. *Advances in Biological Regulation*, *57*, 1–9. https://doi.org/10.1016/j.jbior.2014.09.013

Adams, B. D., Claffey, K. P., & White, B. A. (2009). Argonaute-2 expression is regulated by epidermal growth factor receptor and mitogen-activated protein kinase signaling and correlates with a transformed phenotype in breast cancer cells. *Endocrinology*, *150*(1), 14–23. https://doi.org/10.1210/en.2008-0984

Agranat-Tamir, L., Shomron, N., Sperling, J., & Sperling, R. (2014). Interplay between pre-mRNA splicing and microRNA biogenesis within the supraspliceosome. *Nucleic Acids Research*, *42*(7), 4640–4651. https://doi.org/10.1093/nar/gkt1413

Alagia, A., Jorge, A. F., Aviñó, A., Cova, T. F. G. G., Crehuet, R., Grijalvo, S., Pais, A. A. C. C., & Eritja, R. (2018). Exploring PAZ/3'-overhang interaction to improve siRNA specificity. A combined experimental and modeling study. *Chemical Science*, *9*(8), 2074–2086. https://doi.org/10.1039/c8sc00010g

Alexander, R. P., Fang, G., Rozowsky, J., Snyder, M., & Gerstein, M. B. (2010). Annotating non-coding regions of the genome [Number: 8 Publisher: Nature Publishing Group]. *Nature Reviews Genetics*, *11*(8), 559–571. https://doi.org/10.1038/nrg2814

Alló, M., Agirre, E., Bessonov, S., Bertucci, P., Gómez Acuña, L., Buggiano, V., Bellora, N., Singh, B., Petrillo, E., Blaustein, M., Miñana, B., Dujardin, G., Pozzi, B., Pelisch, F., Bechara, E., Agafonov, D. E., Srebrow, A., Lührmann, R., Valcárcel, J., … Kornblihtt, A. R. (2014). Argonaute-1 binds transcriptional enhancers and controls constitutive and alternative splicing in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(44), 15622–15629. https://doi.org/10.1073/pnas.1416858111

Amarzguioui, M., Holen, T., Babaie, E., & Prydz, H. (2003). Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Research*, *31*(2), 589–595. https://doi.org/10.1093/nar/gkg147

Ambros, V. (2004). The functions of animal microRNAs [Number: 7006 Publisher: Nature Publishing Group]. *Nature*, *431*(7006), 350–355. https://doi.org/10.1038/nature02871

Ameres, S. L., Horwich, M. D., Hung, J.-H., Xu, J., Ghildiyal, M., Weng, Z., & Zamore, P. D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. *Science (New York, N.Y.)*, *328*(5985), 1534–1539. https://doi.org/10.1126/science.1187058

Ameres, S. L., Martinez, J., & Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. *Cell*, *130*(1), 101–112. https://doi.org/10.1016/j.cell.2007.04.037

Audic, Y., & Hartley, R. S. (2004). Post-transcriptional regulation in cancer. *Biology of the Cell*, *96*(7), 479–498. https://doi.org/10.1016/j.biolcel.2004.05.002

Auyeung, V. C., Ulitsky, I., McGeary, S. E., & Bartel, D. P. (2013). Beyond secondary structure: Primary-sequence determinants license pri-miRNA hairpins for processing. *Cell*, *152*(4), 844–858. https://doi.org/10.1016/j.cell.2013.01.031

Azubel, M., Wolf, S. G., Sperling, J., & Sperling, R. (2004). Three-dimensional structure of the native spliceosome by cryo-electron microscopy. *Molecular Cell*, *15*(5), 833–839. https://doi.org/10.1016/j.molcel.2004.07.022

Barbier, J., Chen, X., Sanchez, G., Cai, M., Helsmoortel, M., Higuchi, T., Giraud, P., Contreras, X., Yuan, G., Feng, Z., Nait-Saidi, R., Deas, O., Bluy, L., Judde, J.-G., Rouquier, S., Ritchie, W., Sakamoto, S., Xie, D., & Kiernan, R.

(2018). An NF90/NF110-mediated feedback amplification loop regulates dicer expression and controls ovarian carcinoma progression. *Cell Research*, *28*(5), 556–571. https://doi.org/10.1038/s41422-018-0016-8

Bartel, D. P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell*, *136*(2), 215–233. https://doi.org/10.1016/j.cell.2009.01.002

Bartel, D. P. (2018). Metazoan MicroRNAs. *Cell*, *173*(1), 20–51. https://doi.org/10.1016/j.cell.2018.03.006

Barth, S., Pfuhl, T., Mamiani, A., Ehses, C., Roemer, K., Kremmer, E., Jäker, C., Höck, J., Meister, G., & Grässer, F. A. (2008). Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic Acids Research*, *36*(2), 666–675. https://doi.org/10.1093/nar/gkm1080

Bennasser, Y., Chable-Bessia, C., Triboulet, R., Gibbings, D., Gwizdek, C., Dargemont, C., Kremer, E. J., Voinnet, O., & Benkirane, M. (2011). Competition for XPO5 binding between Dicer mRNA, pre-miRNA and viral RNA regulates human Dicer levels. *Nature Structural & Molecular Biology*, *18*(3), 323–327. https://doi.org/10.1038/nsmb.1987

Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., Sharon, E., Spector, Y., & Bentwich, Z. (2005). Identification of hundreds of conserved and nonconserved human microRNAs. *Nature Genetics*, *37*(7), 766–770. https://doi.org/10.1038/ng1590

Berezikov, E. (2011). Evolution of microRNA diversity and regulation in animals. *Nature Reviews. Genetics*, *12*(12), 846–860. https://doi.org/10.1038/nrg3079

Berezikov, E., Chung, W.-J., Willis, J., Cuppen, E., & Lai, E. C. (2007). Mammalian mirtron genes. *Molecular Cell*, *28*(2), 328–336. https://doi.org/10.1016/j.molcel.2007.09.028

Berezikov, E., Thuemmler, F., van Laake, L. W., Kondova, I., Bontrop, R., Cuppen, E., & Plasterk, R. H. A. (2006). Diversity of microRNAs in human and chimpanzee brain. *Nature Genetics*, *38*(12), 1375–1377. https://doi.org/10.1038/ng1914

Berkhout, B. (2018). RNAi-mediated antiviral immunity in mammals. *Current Opinion in Virology*, *32*, 9–14. https://doi.org/10.1016/j.coviro.2018.07.008

Betancur, J. G., & Tomari, Y. (2012). Dicer is dispensable for asymmetric RISC loading in mammals. *RNA (New York, N.Y.)*, *18*(1), 24–30. https://doi.org/10.1261/rna.029785.111

Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., & Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell*, *125*(6), 1111–1124. https://doi.org/10.1016/j.cell.2006.04.031

Boland, A., Huntzinger, E., Schmidt, S., Izaurralde, E., & Weichenrieder, O. (2011). Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(26), 10466–10471. https://doi.org/10.1073/pnas.1103946108

Bommer, G. T., Gerin, I., Feng, Y., Kaczorowski, A. J., Kuick, R., Love, R. E., Zhai, Y., Giordano, T. J., Qin, Z. S., Moore, B. B., MacDougald, O. A., Cho, K. R., & Fearon, E. R. (2007). P53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Current biology: CB*, *17*(15), 1298–1307. https://doi.org/10.1016/j.cub.2007.06.068

Boulon, S., Verheggen, C., Jady, B. E., Girard, C., Pescia, C., Paul, C., Ospina, J. K., Kiss, T., Matera, A. G., Bordonné, R., & Bertrand, E. (2004). PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Molecular Cell*, *16*(5), 777–787. https://doi.org/10.1016/j.molcel.2004.11.013

Bronisz, A., Godlewski, J., Wallace, J. A., Merchant, A. S., Nowicki, M. O., Mathsyaraja, H., Srinivasan, R., Trimboli, A. J., Martin, C. K., Li, F., Yu, L., Fernandez, S. A., Pécot, T., Rosol, T. J., Cory, S., Hallett, M., Park, M., Piper, M. G., Marsh, C. B., … Ostrowski, M. C. (2011). Reprogramming of the tumour microenvironment by stromal PTEN-regulated miR-320. *Nature Cell Biology*, *14*(2), 159–167. https://doi.org/10.1038/ncb2396

Burger, K., Schlackow, M., Potts, M., Hester, S., Mohammed, S., & Gullerova, M. (2017). Nuclear phosphorylated Dicer processes double-stranded RNA in response to DNA damage. *The Journal of Cell Biology*, *216*(8), 2373–2389. https://doi.org/10.1083/jcb.201612131

Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M.,

Bullrich, F., & Croce, C. M. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15524–15529. https://doi.org/10.1073/pnas.242606799

Campo-Paysaa, F., Sémon, M., Cameron, R. A., Peterson, K. J., & Schubert, M. (2011). microRNA complements in deuterostomes: Origin and evolution of microRNAs. *Evolution & Development*, *13*(1), 15–27. https://doi.org/10.1111/j.1525-142X.2010.00452.x

Caramuta, S., Lee, L., Ozata, D. M., Akçakaya, P., Xie, H., Höög, A., Zedenius, J., Bäckdahl, M., Larsson, C., & Lui, W.-O. (2013). Clinical and functional impact of TARBP2 over-expression in adrenocortical carcinoma. *Endocrine-Related Cancer*, *20*(4), 551–564. https://doi.org/10.1530/ERC-13-0098

Castella, S., Bernard, R., Corno, M., Fradin, A., & Larcher, J.-C. (2015). Ilf3 and NF90 functions in RNA biology [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wrna.1270]. *WIREs RNA*, *6*(2), 243–256. https://doi.org/https://doi.org/10.1002/wrna.1270

Castilla-Llorente, V., Nicastro, G., & Ramos, A. (2013). Terminal loop-mediated regulation of miRNA biogenesis: Selectivity and mechanisms. *Biochemical Society Transactions*, *41*(4), 861–865. https://doi.org/10.1042/BST20130058

Chakravarthy, S., Sternberg, S. H., Kellenberger, C. A., & Doudna, J. A. (2010). Substrate-specific kinetics of Dicer-catalyzed RNA processing. *Journal of Molecular Biology*, *404*(3), 392–402. https://doi.org/10.1016/j.jmb.2010.09.030

Chandradoss, S. D., Schirle, N. T., Szczepaniak, M., MacRae, I. J., & Joo, C. (2015). A Dynamic Search Process Underlies MicroRNA Targeting. *Cell*, *162*(1), 96–107. https://doi.org/10.1016/j.cell.2015.06.032

Chaulk, S. G., Ebhardt, H. A., & Fahlman, R. P. (2015). Correlations of microRNA:microRNA expression patterns reveal insights into microRNA clusters and global microRNA expression patterns [Publisher: The Royal Society of Chemistry]. *Molecular BioSystems*, *12*(1), 110–119. https://doi.org/10.1039/C5MB00415B

Chen, J.-S., Su, I.-J., Leu, Y.-W., Young, K.-C., & Sun, H. S. (2012). Expression of T-cell lymphoma invasion and metastasis 2 (TIAM2) promotes proliferation and invasion of liver cancer. *International Journal of Cancer*, *130*(6), 1302–1313. https://doi.org/10.1002/ijc.26117

Chen, Y., Boland, A., Kuzuoğlu-Öztürk, D., Bawankar, P., Loh, B., Chang, C.-T., Weichenrieder, O., & Izaurralde, E. (2014). A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Molecular Cell*, *54*(5), 737–750. https://doi.org/10.1016/j.molcel.2014.03.034

Chen, Z., Lin, J., Wu, S., Xu, C., Chen, F., & Huang, Z. (2018). Up-regulated miR-548k promotes esophageal squamous cell carcinoma progression via targeting long noncoding RNA-LET. *Experimental Cell Research*, *362*(1), 90–101. https://doi.org/10.1016/j.yexcr.2017.11.006

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., & Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, *436*(7051), 740–744. https://doi.org/10.1038/nature03868

Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeung, V. C., Spies, N., Baek, D., Johnston, W. K., Russ, C., Luo, S., Babiarz, J. E., Blelloch, R., Schroth, G. P., Nusbaum, C., & Bartel, D. P. (2010). Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Genes & Development*, *24*(10), 992–1009. https://doi.org/10.1101/gad.1884710

Chien, C.-H., Sun, Y.-M., Chang, W.-C., Chiang-Hsieh, P.-Y., Lee, T.-Y., Tsai, W.-C., Horng, J.-T., Tsou, A.-P., & Huang, H.-D. (2011). Identifying transcriptional start sites of human microRNAs based on high-throughput sequencing data. *Nucleic Acids Research*, *39*(21), 9345–9356. https://doi.org/10.1093/nar/gkr604

Chou, C.-H., Shrestha, S., Yang, C.-D., Chang, N.-W., Lin, Y.-L., Liao, K.-W., Huang, W.-C., Sun, T.-H., Tu, S.-J., Lee, W.-H., Chiew, M.-Y., Tai, C.-S., Wei, T.-Y., Tsai, T.-R., Huang, H.-T., Wang, C.-Y., Wu, H.-Y., Ho, S.-Y., Chen, P.-R., … Huang, H.-D. (2018). miRTarBase update 2018: A resource

for experimentally validated microRNA-target interactions. *Nucleic Acids Research*, *46*(D1), D296–D302. https://doi.org/10.1093/nar/gkx1067

Choudhury, N. R., de Lima Alves, F., de Andrés-Aguayo, L., Graf, T., Cáceres, J. F., Rappsilber, J., & Michlewski, G. (2013). Tissue-specific control of brain-enriched miR-7 biogenesis. *Genes & Development*, *27*(1), 24–38. https://doi.org/10.1101/gad.199190.112

Choudhury, Y., Tay, F. C., Lam, D. H., Sandanaraj, E., Tang, C., Ang, B.-T., & Wang, S. (2012). Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells. *The Journal of Clinical Investigation*, *122*(11), 4059–4076. https://doi.org/10.1172/JCI62925

Christou-Kent, M., Dhellemmes, M., Lambert, E., Ray, P. F., & Arnoult, C. (2020). Diversity of RNA-Binding Proteins Modulating Post-Transcriptional Regulation of Protein Expression in the Maturing Mammalian Oocyte. *Cells*, *9*(3). https://doi.org/10.3390/cells9030662

Ciechanowska, K., Pokornowska, M., & Kurzyńska-Kokorniak, A. (2021). Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases [Number: 2 Publisher: Multidisciplinary Digital Publishing Institute]. *International Journal of Molecular Sciences*, *22*(2), 616. https://doi.org/10.3390/ijms22020616

Cifuentes, D., Xue, H., Taylor, D. W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G. J., Lawson, N. D., Wolfe, S. A., & Giraldez, A. J. (2010). A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science (New York, N.Y.)*, *328*(5986), 1694–1698. https://doi.org/10.1126/science.1190809

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., Chaves, D. A., Shirayama, M., Mitani, S., Ketting, R. F., Conte, D., & Mello, C. C. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. *Cell*, *139*(1), 123–134. https://doi.org/10.1016/j.cell.2009.09.014

Cock, J. M., Liu, F., Duan, D., Bourdareau, S., Lipinska, A. P., Coelho, S. M., & Tarver, J. E. (2017). Rapid Evolution of microRNA Loci in the Brown Algae. *Genome Biology and Evolution*, *9*(3), 740–749. https://doi.org/10.1093/gbe/evx038

Cole, C., Sobala, A., Lu, C., Thatcher, S. R., Bowman, A., Brown, J. W. S., Green, P. J., Barton, G. J., & Hutvagner, G. (2009). Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA (New York, N.Y.)*, *15*(12), 2147–2160. https://doi.org/10.1261/rna.1738409

Contreras, X., Salifou, K., Sanchez, G., Helsmoortel, M., Beyne, E., Bluy, L., Pelletier, S., Rousset, E., Rouquier, S., & Kiernan, R. (2018). Nuclear RNA surveillance complexes silence HIV-1 transcription (B. R. Cullen, Ed.). *PLOS Pathogens*, *14*(3), e1006950. https://doi.org/10.1371/journal.ppat.1006950

Corbett, A. H. (2018). Post-transcriptional regulation of gene expression and human disease. *Current Opinion in Cell Biology*, *52*, 96–104. https://doi.org/10.1016/j.ceb.2018.02.011

Correia de Sousa, M., Gjorgjieva, M., Dolicka, D., Sobolewski, C., & Foti, M. (2019). Deciphering miRNAs' Action through miRNA Editing. *International Journal of Molecular Sciences*, *20*(24). https://doi.org/10.3390/ijms20246249

Cortez, M. A., Ivan, C., Valdecanas, D., Wang, X., Peltier, H. J., Ye, Y., Araujo, L., Carbone, D. P., Shilo, K., Giri, D. K., Kelnar, K., Martin, D., Komaki, R., Gomez, D. R., Krishnan, S., Calin, G. A., Bader, A. G., & Welsh, J. W. (2016). PDL1 Regulation by p53 via miR-34. *Journal of the National Cancer Institute*, *108*(1). https://doi.org/10.1093/jnci/djv303

Corthésy, B., & Kao, P. N. (1994). Purification by DNA affinity chromatography of two polypeptides that contact the NF-AT DNA binding site in the interleukin 2 promoter. *The Journal of Biological Chemistry*, *269*(32), 20682–20690.

Courel, M., Clément, Y., Bossevain, C., Foretek, D., Vidal Cruchez, O., Yi, Z., Bénard, M., Benassy, M.-N., Kress, M., Vindry, C., Ernoult-Lange, M., Antoniewski, C., Morillon, A., Brest, P., Hubstenberger, A., Roest Crollius, H., Standart, N., & Weil, D. (2019). GC content shapes mRNA storage and decay in human cells. *eLife*, *8*, e49708. https://doi.org/10.7554/eLife.49708

Creugny, A., Fender, A., & Pfeffer, S. (2018). Regulation of primary microRNA processing. *FEBS letters*, *592*(12), 1980–1996. https://doi.org/10.1002/1873-3468.13067

Cullen, B. R. (2006). Is RNA interference involved in intrinsic antiviral immunity in mammals? [Number: 6 Publisher: Nature Publishing Group]. *Nature Immunology*, *7*(6), 563–567. https://doi.org/10.1038/ni1352

Curtis, H. J., Sibley, C. R., & Wood, M. J. A. (2012). Mirtrons, an emerging class of atypical miRNA. *Wiley interdisciplinary reviews. RNA*, *3*(5), 617–632. https://doi.org/10.1002/wrna.1122

Daniels, S. M., Melendez-Peña, C. E., Scarborough, R. J., Daher, A., Christensen, H. S., El Far, M., Purcell, D. F. J., Lainé, S., & Gatignol, A. (2009). Characterization of the TRBP domain required for dicer interaction and function in RNA interference. *BMC molecular biology*, *10*, 38. https://doi.org/10.1186/1471-2199-10-38

Dardenne, E., Polay Espinoza, M., Fattet, L., Germann, S., Lambert, M.-P., Neil, H., Zonta, E., Mortada, H., Gratadou, L., Deygas, M., Chakrama, F. Z., Samaan, S., Desmet, F.-O., Tranchevent, L.-C., Dutertre, M., Rimokh, R., Bourgeois, C. F., & Auboeuf, D. (2014). RNA helicases DDX5 and DDX17 dynamically orchestrate transcription, miRNA, and splicing programs in cell differentiation. *Cell Reports*, *7*(6), 1900–1913. https://doi.org/10.1016/j.celrep.2014.05.010

Davis, B. N., Hilyard, A. C., Lagna, G., & Hata, A. (2008). SMAD proteins control DROSHA-mediated microRNA maturation. *Nature*, *454*(7200), 56–61. https://doi.org/10.1038/nature07086

Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F., & Hannon, G. J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature*, *432*(7014), 231–235. https://doi.org/10.1038/nature03049

Di Leva, G., Garofalo, M., & Croce, C. M. (2014). microRNAs in cancer. *Annual review of pathology*, *9*, 287–314. https://doi.org/10.1146/annurev-pathol-012513-104715

Diederichs, S., & Haber, D. A. (2007). Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell*, *131*(6), 1097–1108. https://doi.org/10.1016/j.cell.2007.10.032

Ding, D., Huang, H., Li, Q., Yu, W., Wang, C., Ma, H., Wu, J., Dang, Y., Yu, L., & Jiang, W. (2020). NF90 stabilizes cyclin E1 mRNA through phosphorylation

of NF90-Ser382 by CDK2. *Cell Death Discovery*, *6*, 3. https://doi.org/10.1038/s41420-020-0236-9

Dlakić, M. (2006). DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. *Bioinformatics (Oxford, England)*, *22*(22), 2711–2714. https://doi.org/10.1093/bioinformatics/btl468

Drake, M., Furuta, T., Suen, K. M., Gonzalez, G., Liu, B., Kalia, A., Ladbury, J. E., Fire, A. Z., Skeath, J. B., & Arur, S. (2014). A requirement for ERK-dependent Dicer phosphorylation in coordinating oocyte-to-embryo transition in C. elegans. *Developmental Cell*, *31*(5), 614–628. https://doi.org/10.1016/j.devcel.2014.11.004

Drusco, A., Nuovo, G. J., Zanesi, N., Di Leva, G., Pichiorri, F., Volinia, S., Fernandez, C., Antenucci, A., Costinean, S., Bottoni, A., Rosito, I. A., Liu, C.-G., Burch, A., Acunzo, M., Pekarsky, Y., Alder, H., Ciardi, A., & Croce, C. M. (2014). MicroRNA profiles discriminate among colon cancer metastasis. *PloS One*, *9*(6), e96670. https://doi.org/10.1371/journal.pone.0096670

Du, Z., Lee, J. K., Tjhen, R., Stroud, R. M., & James, T. L. (2008). Structural and biochemical insights into the dicing mechanism of mouse Dicer: A conserved lysine is critical for dsRNA cleavage. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(7), 2391–2396. https://doi.org/10.1073/pnas.0711506105

Duchange, N., Pidoux, J., Camus, E., & Sauvaget, D. (2000). Alternative splicing in the human interleukin enhancer binding factor 3 (ILF3) gene. *Gene*, *261*(2), 345–353. https://doi.org/10.1016/s0378-1119(00)00495-9

Duret, L., Dorkeld, F., & Gautier, C. (1993). Strong conservation of non-coding sequences during vertebrates evolution: Potential involvement in post-transcriptional regulation of gene expression. *Nucleic Acids Research*, *21*(10), 2315–2322. https://doi.org/10.1093/nar/21.10.2315

Ebert, M. S., & Sharp, P. A. (2012). Roles for microRNAs in conferring robustness to biological processes. *Cell*, *149*(3), 515–524. https://doi.org/10.1016/j.cell.2012.04.005

Elbashir, S. M., Lendeckel, W., & Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development*, *15*(2), 188–200. https://doi.org/10.1101/gad.862301

Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W., & Tuschl, T. (2001). Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *The EMBO journal*, *20*(23), 6877–6888. https://doi.org/10.1093/emboj/20.23.6877

Fabian, M. R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J. T., Svitkin, Y. V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J. A., Chen, C.-Y. A., Shyu, A.-B., Yates, J. R., Hannon, G. J., Filipowicz, W., Duchaine, T. F., & Sonenberg, N. (2009). Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Molecular Cell*, *35*(6), 868–880. https://doi.org/10.1016/j.molcel.2009.08.004

Fang, W., & Bartel, D. P. (2015). The Menu of Features that Define Primary MicroRNAs and Enable De Novo Design of MicroRNA Genes. *Molecular Cell*, *60*(1), 131–145. https://doi.org/10.1016/j.molcel.2015.08.015

Felippes, F. F. d., Schneeberger, K., Dezulian, T., Huson, D. H., & Weigel, D. (2008). Evolution of Arabidopsis thaliana microRNAs from random sequences. *RNA (New York, N.Y.)*, *14*(12), 2455–2459. https://doi.org/10.1261/rna.1149408

Fernandez-Valverde, S. L., Taft, R. J., & Mattick, J. S. (2010). Dynamic isomiR regulation in Drosophila development. *RNA (New York, N.Y.)*, *16*(10), 1881–1888. https://doi.org/10.1261/rna.2379610

Finnegan, E. F., & Pasquinelli, A. E. (2013). MicroRNA biogenesis: Regulating the regulators. *Critical Reviews in Biochemistry and Molecular Biology*, *48*(1), 51–68. https://doi.org/10.3109/10409238.2012.738643

Flores, O., Kennedy, E. M., Skalsky, R. L., & Cullen, B. R. (2014). Differential RISC association of endogenous human microRNAs predicts their inhibitory potential. *Nucleic Acids Research*, *42*(7), 4629–4639. https://doi.org/10.1093/nar/gkt1393

Flynt, A. S., Greimann, J. C., Chung, W.-J., Lima, C. D., & Lai, E. C. (2010). MicroRNA biogenesis via splicing and exosome-mediated trimming in

Drosophila. *Molecular Cell*, *38*(6), 900–907. https://doi.org/10.1016/j.molcel.2010.06.014

Földes-Papp, Z., König, K., Studier, H., Bückle, R., Breunig, H. G., Uchugonova, A., & Kostner, G. M. (2009). Trafficking of mature miRNA-122 into the nucleus of live liver cells. *Current Pharmaceutical Biotechnology*, *10*(6), 569–578. https://doi.org/10.2174/138920109789069332

Francia, S., Michelini, F., Saxena, A., Tang, D., de Hoon, M., Anelli, V., Mione, M., Carninci, P., & d'Adda di Fagagna, F. (2012). Site-specific DICER and DROSHA RNA products control the DNA-damage response. *Nature*, *488*(7410), 231–235. https://doi.org/10.1038/nature11179

Fritz, J., Strehblow, A., Taschner, A., Schopoff, S., Pasierbek, P., & Jantsch, M. F. (2009). RNA-regulated interaction of transportin-1 and exportin-5 with the double-stranded RNA-binding domain regulates nucleocytoplasmic shuttling of ADAR1. *Molecular and Cellular Biology*, *29*(6), 1487–1497. https://doi.org/10.1128/MCB.01519-08

Fuchs Wightman, F., Giono, L. E., Fededa, J. P., & de la Mata, M. (2018). Target RNAs Strike Back on MicroRNAs. *Frontiers in Genetics*, *9*. https://doi.org/10.3389/fgene.2018.00435

Gantier, M. P., McCoy, C. E., Rusinova, I., Saulep, D., Wang, D., Xu, D., Irving, A. T., Behlke, M. A., Hertzog, P. J., Mackay, F., & Williams, B. R. G. (2011). Analysis of microRNA turnover in mammalian cells following Dicer1 ablation. *Nucleic Acids Research*, *39*(13), 5692–5703. https://doi.org/10.1093/nar/gkr148

Geisler, S., & Coller, J. (2013). RNA in unexpected places: Long non-coding RNA functions in diverse cellular contexts. *Nature reviews. Molecular cell biology*, *14*(11), 699–712. https://doi.org/10.1038/nrm3679

Gerstberger, S., Hafner, M., & Tuschl, T. (2014). A census of human RNA-binding proteins [Number: 12 Publisher: Nature Publishing Group]. *Nature Reviews Genetics*, *15*(12), 829–845. https://doi.org/10.1038/nrg3813

Giuliani, A., Londin, E., Ferracin, M., Mensà, E., Prattichizzo, F., Ramini, D., Marcheselli, F., Recchioni, R., Rippo, M. R., Bonafè, M., Rigoutsos, I., Olivieri, F., & Sabbatinelli, J. (2020). Long-term exposure of human

endothelial cells to metformin modulates miRNAs and isomiRs. *Scientific Reports*, *10*(1), 21782. https://doi.org/10.1038/s41598-020-78871-5

Gollavilli, P. N., Parma, B., Siddiqui, A., Yang, H., Ramesh, V., Napoli, F., Schwab, A., Natesan, R., Mielenz, D., Asangani, I. A., Brabletz, T., Pilarsky, C., & Ceppi, P. (2021). The role of miR-200b/c in balancing EMT and proliferation revealed by an activity reporter. *Oncogene*, *40*(12), 2309–2322. https://doi.org/10.1038/s41388-021-01708-6

Gomila, R. C., Martin, G. W., & Gehrke, L. (2011). NF90 binds the dengue virus RNA 3′ terminus and is a positive regulator of dengue virus replication. *PloS One*, *6*(2), e16687. https://doi.org/10.1371/journal.pone.0016687

Grasso, G., Higuchi, T., Mac, V., Barbier, J., Helsmoortel, M., Lorenzi, C., Sanchez, G., Bello, M., Ritchie, W., Sakamoto, S., & Kiernan, R. (2020). NF90 modulates processing of a subset of human pri-miRNAs. *Nucleic Acids Research*, *48*(12), 6874–6888. https://doi.org/10.1093/nar/gkaa386

Gregersen, L. H., Schueler, M., Munschauer, M., Mastrobuoni, G., Chen, W., Kempa, S., Dieterich, C., & Landthaler, M. (2014). MOV10 Is a 5′ to 3′ RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3′ UTRs. *Molecular Cell*, *54*(4), 573–585. https://doi.org/10.1016/j.molcel.2014.03.017

Gregory, R. I., Yan, K.-P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., & Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature*, *432*(7014), 235–240. https://doi.org/10.1038/nature03120

Grimson, A., Farh, K. K.-H., Johnston, W. K., Garrett-Engele, P., Lim, L. P., & Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Molecular Cell*, *27*(1), 91–105. https://doi.org/10.1016/j.molcel.2007.06.017

Groebe, D. R., & Uhlenbeck, O. C. (1988). Characterization of RNA hairpin loop stability. *Nucleic Acids Research*, *16*(24), 11725–11735. https://doi.org/10.1093/nar/16.24.11725

Gross, C. H., & Shuman, S. (1996). Vaccinia virus RNA helicase: Nucleic acid specificity in duplex unwinding. *Journal of Virology*, *70*(4), 2615–2619. https://doi.org/10.1128/JVI.70.4.2615-2619.1996

Grosswendt, S., Filipchyk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E., & Rajewsky, N. (2014). Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Molecular Cell*, *54*(6), 1042–1054. https://doi.org/10.1016/j.molcel.2014.03.049

Gu, S., Jin, L., Huang, Y., Zhang, F., & Kay, M. A. (2012). Slicing-independent RISC activation requires the argonaute PAZ domain. *Current biology: CB*, *22*(16), 1536–1542. https://doi.org/10.1016/j.cub.2012.06.040

Guan, D., Altan-Bonnet, N., Parrott, A. M., Arrigo, C. J., Li, Q., Khaleduzzaman, M., Li, H., Lee, C.-G., Pe'ery, T., & Mathews, M. B. (2008). Nuclear Factor 45 (NF45) Is a Regulatory Subunit of Complexes with NF90/110 Involved in Mitotic Control. *Molecular and Cellular Biology*, *28*(14), 4629–4641. https://doi.org/10.1128/MCB.00120-08

Guil, S., & Cáceres, J. F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nature Structural & Molecular Biology*, *14*(7), 591–596. https://doi.org/10.1038/nsmb1250

Guo, Y., Liu, J., Elfenbein, S. J., Ma, Y., Zhong, M., Qiu, C., Ding, Y., & Lu, J. (2015). Characterization of the mammalian miRNA turnover landscape. *Nucleic Acids Research*, *43*(4), 2326–2341. https://doi.org/10.1093/nar/gkv057

Gwizdek, C., Ossareh-Nazari, B., Brownawell, A. M., Doglio, A., Bertrand, E., Macara, I. G., & Dargemont, C. (2003). Exportin-5 mediates nuclear export of minihelix-containing RNAs. *The Journal of Biological Chemistry*, *278*(8), 5505–5508. https://doi.org/10.1074/jbc.C200668200

Gwizdek, C., Ossareh-Nazari, B., Brownawell, A. M., Evers, S., Macara, I. G., & Dargemont, C. (2004). Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3. *The Journal of Biological Chemistry*, *279*(2), 884–891. https://doi.org/10.1074/jbc.M306808200

Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nature Reviews. Molecular Cell Biology*, *15*(8), 509–524. https://doi.org/10.1038/nrm3838

Hamilton, A. J., & Baulcombe, D. C. (1999). A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants [Publisher: American Association for the Advancement of Science Section: Report]. *Science*, *286*(5441), 950–952. https://doi.org/10.1126/science.286.5441.950

Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S. Y., Cho, Y., Zhang, B.-T., & Kim, V. N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, *125*(5), 887–901. https://doi.org/10.1016/j.cell.2006.03.043

Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H., & Kim, V. N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development*, *18*(24), 3016–3027. https://doi.org/10.1101/gad.1262504

Han, J., Pedersen, J. S., Kwon, S. C., Belair, C. D., Kim, Y.-K., Yeom, K.-H., Yang, W.-Y., Haussler, D., Blelloch, R., & Kim, V. N. (2009). Posttranscriptional crossregulation between Drosha and DGCR8. *Cell*, *136*(1), 75–84. https://doi.org/10.1016/j.cell.2008.10.053

Hansen, S. R., Aderounmu, A. M., Donelick, H. M., & Bass, B. L. (2019). Dicer's Helicase Domain: A Meeting Place for Regulatory Proteins. *Cold Spring Harbor Symposia on Quantitative Biology*, *84*, 185–193. https://doi.org/10.1101/sqb.2019.84.039750

Harashima, A., Guettouche, T., & Barber, G. N. (2010). Phosphorylation of the NFAR proteins by the dsRNA-dependent protein kinase PKR constitutes a novel mechanism of translational regulation and cellular defense. *Genes & Development*, *24*(23), 2640–2653. https://doi.org/10.1101/gad.1965010

Harborth, J., Elbashir, S. M., Vandenburgh, K., Manninga, H., Scaringe, S. A., Weber, K., & Tuschl, T. (2003). Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense & Nucleic Acid Drug Development*, *13*(2), 83–105. https://doi.org/10.1089/108729003321629638

Hausser, J., Syed, A. P., Selevsek, N., van Nimwegen, E., Jaskiewicz, L., Aebersold, R., & Zavolan, M. (2013). Timescales and bottlenecks in miRNA-dependent gene

regulation. *Molecular Systems Biology*, *9*, 711. https://doi.org/10.1038/msb.2013.68

Havens, M. A., Reich, A. A., Duelli, D. M., & Hastings, M. L. (2012). Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Research*, *40*(10), 4626–4640. https://doi.org/10.1093/nar/gks026

Havens, M. A., Reich, A. A., & Hastings, M. L. (2014). Drosha promotes splicing of a pre-microRNA-like alternative exon. *PLoS genetics*, *10*(5), e1004312. https://doi.org/10.1371/journal.pgen.1004312

Heimberg, A. M., Sempere, L. F., Moy, V. N., Donoghue, P. C. J., & Peterson, K. J. (2008). MicroRNAs and the advent of vertebrate morphological complexity. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(8), 2946–2950. https://doi.org/10.1073/pnas.0712259105

Heller, G., Altenberger, C., Steiner, I., Topakian, T., Ziegler, B., Tomasich, E., Lang, G., End-Pfützenreuter, A., Zehetmayer, S., Döme, B., Arns, B.-M., Klepetko, W., Zielinski, C. C., & Zöchbauer-Müller, S. (2018). DNA methylation of microRNA-coding genes in non-small-cell lung cancer patients. *The Journal of Pathology*, *245*(4), 387–398. https://doi.org/10.1002/path.5079

Heyam, A., Lagos, D., & Plevin, M. (2015). Dissecting the roles of TRBP and PACT in double-stranded RNA recognition and processing of noncoding RNAs. *Wiley interdisciplinary reviews. RNA*, *6*(3), 271–289. https://doi.org/10.1002/wrna.1272

Higuchi, T., Morisawa, K., Todaka, H., Lai, S., Chi, E., Matsukawa, K., Sugiyama, Y., & Sakamoto, S. (2018). A negative feedback loop between nuclear factor 90 (NF90) and an anti-oncogenic microRNA, miR-7. *Biochemical and Biophysical Research Communications*, *503*(3), 1819–1824. https://doi.org/10.1016/j.bbrc.2018.07.119

Higuchi, T., Todaka, H., Sugiyama, Y., Ono, M., Tamaki, N., Hatano, E., Takezaki, Y., Hanazaki, K., Miwa, T., Lai, S., Morisawa, K., Tsuda, M., Taniguchi, T., & Sakamoto, S. (2016). Suppression of MicroRNA-7 (miR-7) Biogenesis by Nuclear Factor 90-Nuclear Factor 45 Complex (NF90-NF45) Controls Cell

Proliferation in Hepatocellular Carcinoma. *The Journal of Biological Chemistry*, *291*(40), 21074–21084. https://doi.org/10.1074/jbc.M116.748210

Höck, J., Weinmann, L., Ender, C., Rüdel, S., Kremmer, E., Raabe, M., Urlaub, H., & Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA–protein complexes in human cells. *EMBO Reports*, *8*(11), 1052–1060. https://doi.org/10.1038/sj.embor.7401088

Hombach, S., & Kretz, M. (2016). Non-coding RNAs: Classification, Biology and Functioning. In O. Slaby & G. A. Calin (Eds.), *Non-coding RNAs in Colorectal Cancer* (pp. 3–17). Springer International Publishing. https://doi.org/10.1007/978-3-319-42059-2_1

Hong, S., Noh, H., Chen, H., Padia, R., Pan, Z. K., Su, S.-B., Jing, Q., Ding, H.-F., & Huang, S. (2013). Signaling by p38 MAPK stimulates nuclear localization of the microprocessor component p68 for processing of selected primary microRNAs. *Science Signaling*, *6*(266), ra16. https://doi.org/10.1126/scisignal.2003706

Hoque, M., Shamanna, R. A., Guan, D., Pe'ery, T., & Mathews, M. B. (2011). HIV-1 replication and latency are regulated by translational control of cyclin T1. *Journal of Molecular Biology*, *410*(5), 917–932. https://doi.org/10.1016/j.jmb.2011.03.060

Hsu, S.-H., Wang, B., Kota, J., Yu, J., Costinean, S., Kutay, H., Yu, L., Bai, S., La Perle, K., Chivukula, R. R., Mao, H., Wei, M., Clark, K. R., Mendell, J. R., Caligiuri, M. A., Jacob, S. T., Mendell, J. T., & Ghoshal, K. (2012). Essential metabolic, anti-inflammatory, and anti-tumorigenic functions of miR-122 in liver. *The Journal of Clinical Investigation*, *122*(8), 2871–2883. https://doi.org/10.1172/JCI63539

Hua, X., Chen, L., Wang, J., Li, J., & Wingender, E. (2016). Identifying cell-specific microRNA transcriptional start sites. *Bioinformatics (Oxford, England)*, *32*(16), 2403–2410. https://doi.org/10.1093/bioinformatics/btw171

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, V., Zheng, J., Qi, Z., Wang, J., Place, R. F., Yu, J., Li, H., & Li, L.-C. (2013). Ago1 Interacts with RNA polymerase II and binds to the promoters of actively transcribed genes in human cancer cells. *PLoS genetics*, *9*(9), e1003821. https://doi.org/10.1371/journal.pgen.1003821

Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.-B., Munier, A., Fradet, M., Daunesse, M., Bertrand, E., Pierron, G., Mozziconacci, J., Kress, M., & Weil, D. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Molecular Cell*, *68*(1), 144–157.e5. https://doi.org/10.1016/j.molcel.2017.09.003

Idda, M. L., Lodde, V., McClusky, W. G., Martindale, J. L., Yang, X., Munk, R., Steri, M., Orrù, V., Mulas, A., Cucca, F., Abdelmohsen, K., & Gorospe, M. (2018). Cooperative translational control of polymorphic BAFF by NF90 and miR-15a. *Nucleic Acids Research*, *46*(22), 12040–12051. https://doi.org/10.1093/nar/gky866

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., & Tomari, Y. (2010). Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Molecular Cell*, *39*(2), 292–299. https://doi.org/10.1016/j.molcel.2010.05.015

Janas, M. M., Khaled, M., Schubert, S., Bernstein, J. G., Golan, D., Veguilla, R. A., Fisher, D. E., Shomron, N., Levy, C., & Novina, C. D. (2011). Feed-forward microprocessing and splicing activities at a microRNA-containing intron. *PLoS genetics*, *7*(10), e1002330. https://doi.org/10.1371/journal.pgen.1002330

Janas, M. M., Wang, B., Harris, A. S., Aguiar, M., Shaffer, J. M., Subrahmanyam, Y. V. B. K., Behlke, M. A., Wucherpfennig, K. W., Gygi, S. P., Gagnon, E., & Novina, C. D. (2012). Alternative RISC assembly: Binding and repression of microRNA-mRNA duplexes by human Ago proteins. *RNA (New York, N.Y.)*, *18*(11), 2041–2055. https://doi.org/10.1261/rna.035675.112

Janowski, B. A., Huffman, K. E., Schwartz, J. C., Ram, R., Nordsell, R., Shames, D. S., Minna, J. D., & Corey, D. R. (2006). Involvement of AGO1 and

AGO2 in mammalian transcriptional silencing. *Nature Structural & Molecular Biology*, *13*(9), 787–792. https://doi.org/10.1038/nsmb1140

Jayachandran, U., Grey, H., & Cook, A. G. (2016). Nuclear factor 90 uses an ADAR2-like binding mode to recognize specific bases in dsRNA. *Nucleic Acids Research*, *44*(4), 1924–1936. https://doi.org/10.1093/nar/gkv1508

Jiang, W., Huang, H., Ding, L., Zhu, P., Saiyin, H., Ji, G., Zuo, J., Han, D., Pan, Y., Ding, D., Ma, X., Zhang, Y., Wu, J., Yi, Q., Liu, J. O., Huang, H., Dang, Y., & Yu, L. (2015). Regulation of cell cycle of hepatocellular carcinoma by NF90 through modulation of cyclin E1 mRNA stability. *Oncogene*, *34*(34), 4460–4470. https://doi.org/10.1038/onc.2014.373

Jinek, M., & Doudna, J. A. (2009). A three-dimensional view of the molecular machinery of RNA interference. *Nature*, *457*(7228), 405–412. https://doi.org/10.1038/nature07755

Johanson, T. M., Keown, A. A., Cmero, M., Yeo, J. H. C., Kumar, A., Lew, A. M., Zhan, Y., & Chong, M. M. W. (2015). Drosha controls dendritic cell development by cleaving messenger RNAs encoding inhibitors of myelopoiesis. *Nature Immunology*, *16*(11), 1134–1141. https://doi.org/10.1038/ni.3293

Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nature Reviews. Genetics*, *16*(7), 421–433. https://doi.org/10.1038/nrg3965

Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M., & Sarnow, P. (2005). Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science (New York, N.Y.)*, *309*(5740), 1577–1581. https://doi.org/10.1126/science.1113329

Kanakamani, S., Suresh, P. S., & Venkatesh, T. (2021). Regulation of processing bodies: From viruses to cancer epigenetic machinery. *Cell Biology International*, *45*(4), 708–719. https://doi.org/10.1002/cbin.11527

Kataoka, N., Fujita, M., & Ohno, M. (2009). Functional association of the Microprocessor complex with the spliceosome. *Molecular and Cellular Biology*, *29*(12), 3243–3254. https://doi.org/10.1128/MCB.00360-09

Kawamata, T., & Tomari, Y. (2010). Making RISC. *Trends in Biochemical Sciences*, *35*(7), 368–376. https://doi.org/10.1016/j.tibs.2010.03.009

Kenny, P. J., Kim, M., Skariah, G., Nielsen, J., Lannom, M. C., & Ceman, S. (2020). The FMRP-MOV10 complex: A translational regulatory switch modulated by G-Quadruplexes. *Nucleic Acids Research*, *48*(2), 862–878. https://doi.org/10.1093/nar/gkz1092

Kenny, P. J., Zhou, H., Kim, M., Skariah, G., Khetani, R. S., Drnevich, J., Arcila, M. L., Kosik, K. S., & Ceman, S. (2014). MOV10 and FMRP Regulate AGO2 Association with MicroRNA Recognition Elements. *Cell Reports*, *9*(5), 1729–1741. https://doi.org/10.1016/j.celrep.2014.10.054

Khongnomnan, K., Makkoch, J., Poomipak, W., Poovorawan, Y., & Payungporn, S. (2015). Human miR-3145 inhibits influenza A viruses replication by targeting and silencing viral PB1 gene. *Experimental Biology and Medicine (Maywood, N.J.)*, *240*(12), 1630–1639. https://doi.org/10.1177/1535370215589051

Khvorova, A., Reynolds, A., & Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell*, *115*(2), 209–216. https://doi.org/10.1016/s0092-8674(03)00801-8

Kiesler, P., Haynes, P. A., Shi, L., Kao, P. N., Wysocki, V. H., & Vercelli, D. (2010). NF45 and NF90 Regulate HS4-dependent Interleukin-13 Transcription in T Cells. *The Journal of Biological Chemistry*, *285*(11), 8256–8267. https://doi.org/10.1074/jbc.M109.041004

Kim, B., Jeong, K., & Kim, V. N. (2017). Genome-wide Mapping of DROSHA Cleavage Sites on Primary MicroRNAs and Noncanonical Substrates. *Molecular Cell*, *66*(2), 258–269.e5. https://doi.org/10.1016/j.molcel.2017.03.013

Kim, K., Nguyen, T. D., Li, S., & Nguyen, T. A. (2018). SRSF3 recruits DROSHA to the basal junction of primary microRNAs. *RNA (New York, N.Y.)*, *24*(7), 892–898. https://doi.org/10.1261/rna.065862.118

Kim, V. N., Han, J., & Siomi, M. C. (2009). Biogenesis of small RNAs in animals. *Nature Reviews. Molecular Cell Biology*, *10*(2), 126–139. https://doi.org/10.1038/nrm2632

Kim, Y., Yeo, J., Lee, J. H., Cho, J., Seo, D., Kim, J.-S., & Kim, V. N. (2014). Deletion of human tarbp2 reveals cellular microRNA targets and cell-cycle function of

TRBP. *Cell Reports*, *9*(3), 1061–1074. https://doi.org/10.1016/j.celrep.2014.09.039

Kim, Y.-K., & Kim, V. N. (2007). Processing of intronic microRNAs. *The EMBO journal*, *26*(3), 775–783. https://doi.org/10.1038/sj.emboj.7601512

Kingston, E. R., & Bartel, D. P. (2021). Ago2 protects Drosophila siRNAs and microRNAs from target-directed degradation, even in the absence of 2-O-methylation [Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab]. *RNA*, rna.078746.121. https://doi.org/10.1261/rna.078746.121

Kleaveland, B., Shi, C. Y., Stefano, J., & Bartel, D. P. (2018). A Network of Noncoding Regulatory RNAs Acts in the Mammalian Brain. *Cell*, *174*(2), 350–362.e17. https://doi.org/10.1016/j.cell.2018.05.022

Knuckles, P., Vogt, M. A., Lugert, S., Milo, M., Chong, M. M. W., Hautbergue, G. M., Wilson, S. A., Littman, D. R., & Taylor, V. (2012). Drosha regulates neurogenesis by controlling neurogenin 2 expression independent of microRNAs. *Nature Neuroscience*, *15*(7), 962–969. https://doi.org/10.1038/nn.3139

Kok, K. H., Ng, M.-H. J., Ching, Y.-P., & Jin, D.-Y. (2007). Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. *The Journal of Biological Chemistry*, *282*(24), 17649–17657. https://doi.org/10.1074/jbc.M611768200

Komori, C., Takahashi, T., Nakano, Y., & Ui-Tei, K. (2020). TRBP-Dicer interaction may enhance HIV-1 TAR RNA translation via TAR RNA processing, repressing host-cell apoptosis. *Biology Open*, *9*(2). https://doi.org/10.1242/bio.050435

Kooshapur, H., Choudhury, N. R., Simon, B., Mühlbauer, M., Jussupow, A., Fernandez, N., Jones, A. N., Dallmann, A., Gabel, F., Camilloni, C., Michlewski, G., Caceres, J. F., & Sattler, M. (2018). Structural basis for terminal loop recognition and stimulation of pri-miRNA-18a processing by

hnRNP A1. *Nature Communications*, *9*(1), 2479. https://doi.org/10.1038/s41467-018-04871-9

Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: From microRNA sequences to function. *Nucleic Acids Research*, *47*(D1), D155–D162. https://doi.org/10.1093/nar/gky1141

Krek, A., Grün, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., & Rajewsky, N. (2005). Combinatorial microRNA target predictions. *Nature Genetics*, *37*(5), 495–500. https://doi.org/10.1038/ng1536

Krol, J., Busskamp, V., Markiewicz, I., Stadler, M. B., Ribi, S., Richter, J., Duebel, J., Bicker, S., Fehling, H. J., Schübeler, D., Oertner, T. G., Schratt, G., Bibel, M., Roska, B., & Filipowicz, W. (2010). Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell*, *141*(4), 618–631. https://doi.org/10.1016/j.cell.2010.03.039

Krol, J., Loedige, I., & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews. Genetics*, *11*(9), 597–610. https://doi.org/10.1038/nrg2843

Kurzynska-Kokorniak, A., Pokornowska, M., Koralewska, N., Hoffmann, W., Bienkowska-Szewczyk, K., & Figlerowicz, M. (2016). Revealing a new activity of the human Dicer DUF283 domain in vitro. *Scientific Reports*, *6*, 23989. https://doi.org/10.1038/srep23989

Kuwano, Y., Pullmann, R., Marasa, B. S., Abdelmohsen, K., Lee, E. K., Yang, X., Martindale, J. L., Zhan, M., & Gorospe, M. (2010). NF90 selectively represses the translation of target mRNAs bearing an AU-rich signature motif. *Nucleic Acids Research*, *38*(1), 225–238. https://doi.org/10.1093/nar/gkp861

Kwak, P. B., & Tomari, Y. (2012). The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nature Structural & Molecular Biology*, *19*(2), 145–151. https://doi.org/10.1038/nsmb.2232

Kwon, S. C., Baek, S. C., Choi, Y.-G., Yang, J., Lee, Y.-S., Woo, J.-S., & Kim, V. N. (2019). Molecular Basis for the Single-Nucleotide Precision of Primary microRNA Processing. *Molecular Cell*, *73*(3), 505–518.e5. https://doi.org/10.1016/j.molcel.2018.11.005

Kwon, S. C., Nguyen, T. A., Choi, Y.-G., Jo, M. H., Hohng, S., Kim, V. N., & Woo, J.-S. (2016). Structure of Human DROSHA. *Cell*, *164*(1-2), 81–90. https://doi.org/10.1016/j.cell.2015.12.019

La, T., Liu, G. Z., Farrelly, M., Cole, N., Feng, Y. C., Zhang, Y. Y., Sherwin, S. K., Yari, H., Tabatabaee, H., Yan, X. G., Guo, S. T., Liu, T., Thorne, R. F., Jin, L., & Zhang, X. D. (2018). A p53-Responsive miRNA Network Promotes Cancer Cell Quiescence. *Cancer Research*, *78*(23), 6666–6679. https://doi.org/10.1158/0008-5472.CAN-18-1886

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A. O., Landthaler, M., Lin, C., Socci, N. D., Hermida, L., Fulci, V., Chiaretti, S., Foà, R., Schliwka, J., Fuchs, U., Novosel, A., … Tuschl, T. (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, *129*(7), 1401–1414. https://doi.org/10.1016/j.cell.2007.04.040

Lee, D., Nam, J.-W., & Shin, C. (2017). DROSHA targets its own transcript to modulate alternative splicing. *RNA (New York, N.Y.)*, *23*(7), 1035–1047. https://doi.org/10.1261/rna.059808.116

Lee, H. Y., & Doudna, J. A. (2012). TRBP alters human precursor microRNA processing in vitro. *RNA (New York, N.Y.)*, *18*(11), 2012–2019. https://doi.org/10.1261/rna.035501.112

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, *75*(5), 843–854. https://doi.org/10.1016/0092-8674(93)90529-y

Lee, Y. S., & Dutta, A. (2009). MicroRNAs in cancer. *Annual review of pathology*, *4*, 199–227. https://doi.org/10.1146/annurev.pathol.4.110807.092222

Lee, Y., Hur, I., Park, S.-Y., Kim, Y.-K., Suh, M. R., & Kim, V. N. (2006). The role of PACT in the RNA silencing pathway. *The EMBO journal*, *25*(3), 522–532. https://doi.org/10.1038/sj.emboj.7600942

Lee, Y. S., Nakahara, K., Pham, J. W., Kim, K., He, Z., Sontheimer, E. J., & Carthew, R. W. (2004). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell*, *117*(1), 69–81. https://doi.org/10.1016/s0092-8674(04)00261-2

Lewis, B. P., Burge, C. B., & Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, *120*(1), 15–20. https://doi.org/10.1016/j.cell.2004.12.035

Li, T., Li, X., Zhu, W., Wang, H., Mei, L., Wu, S., Lin, X., & Han, X. (2016). NF90 is a novel influenza A virus NS1-interacting protein that antagonizes the inhibitory role of NS1 on PKR phosphorylation [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/1873-3468.12311]. *FEBS Letters*, *590*(16), 2797–2810. https://doi.org/https://doi.org/10.1002/1873-3468.12311

Li, X., Liu, C.-X., Xue, W., Zhang, Y., Jiang, S., Yin, Q.-F., Wei, J., Yao, R.-W., Yang, L., & Chen, L.-L. (2017). Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection. *Molecular Cell*, *67*(2), 214–227.e7. https://doi.org/10.1016/j.molcel.2017.05.023

Li, Y., & Belshan, M. (2016). NF45 and NF90 Bind HIV-1 RNA and Modulate HIV Gene Expression. *Viruses*, *8*(2). https://doi.org/10.3390/v8020047

Li, Y., Masaki, T., Shimakami, T., & Lemon, S. M. (2014). hnRNP L and NF90 interact with hepatitis C virus 5′-terminal untranslated RNA and promote efficient replication. *Journal of Virology*, *88*(13), 7199–7209. https://doi.org/10.1128/JVI.00225-14

Liang, T., Han, L., & Guo, L. (2020). Rewired functional regulatory networks among miRNA isoforms (isomiRs) from let-7 and miR-10 gene families in cancer. *Computational and Structural Biotechnology Journal*, *18*, 1238–1248. https://doi.org/10.1016/j.csbj.2020.05.001

Lin, J., Chen, Z., Wu, S., Huang, W., Chen, F., & Huang, Z. (2019). An NF90/long noncoding RNA-LET/miR-548k feedback amplification loop controls esophageal squamous cell carcinoma progression. *Journal of Cancer*, *10*(21), 5139–5152. https://doi.org/10.7150/jca.30816

Lin, S., & Gregory, R. I. (2015). MicroRNA biogenesis pathways in cancer. *Nature reviews. Cancer*, *15*(6), 321–333. https://doi.org/10.1038/nrc3932

Ling, H.-Y., Ou, H.-S., Feng, S.-D., Zhang, X.-Y., Tuo, Q.-H., Chen, L.-X., Zhu, B.-Y., Gao, Z.-P., Tang, C.-K., Yin, W.-D., Zhang, L., & Liao, D.-F. (2009). CHANGES

IN microRNA (miR) profile and effects of miR-320 in insulin-resistant 3T3-L1 adipocytes. *Clinical and Experimental Pharmacology & Physiology*, *36*(9), e32–39. https://doi.org/10.1111/j.1440-1681.2009.05207.x

Link, S., Grund, S. E., & Diederichs, S. (2016). Alternative splicing affects the subcellular localization of Drosha. *Nucleic Acids Research*, *44*(11), 5330–5343. https://doi.org/10.1093/nar/gkw400

Liu, F., Ou, Y., Lin, Q. S., Qiu, C., Luo, H. L., & Zhu, P. Q. (2016). Low expression of long non-coding RNA-LET can indicate metastasis and a poor prognosis: A meta-analysis. *Minerva Medica*.

Liu, H., Liang, C., Kollipara, R. K., Matsui, M., Ke, X., Jeong, B.-C., Wang, Z., Yoo, K. S., Yadav, G. P., Kinch, L. N., Grishin, N. V., Nam, Y., Corey, D. R., Kittler, R., & Liu, Q. (2016). HP1BP3, a Chromatin Retention Factor for Co-transcriptional MicroRNA Processing. *Molecular Cell*, *63*(3), 420–432. https://doi.org/10.1016/j.molcel.2016.06.014

Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J.-J., Hammond, S. M., Joshua-Tor, L., & Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science (New York, N.Y.)*, *305*(5689), 1437–1441. https://doi.org/10.1126/science.1102513

Liu, J., Valencia-Sanchez, M. A., Hannon, G. J., & Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nature Cell Biology*, *7*(7), 719–723. https://doi.org/10.1038/ncb1274

Liu, Q., Wang, J., Zhao, Y., Li, C.-I., Stengel, K. R., Acharya, P., Johnston, G., Hiebert, S. W., & Shyr, Y. (2017). Identification of active miRNA promoters from nuclear run-on RNA sequencing. *Nucleic Acids Research*, *45*(13), e121. https://doi.org/10.1093/nar/gkx318

Liu, Q., & Paroo, Z. (2010). Biochemical principles of small RNA pathways. *Annual Review of Biochemistry*, *79*, 295–319. https://doi.org/10.1146/annurev.biochem.052208.151733

Liu, X., Jin, D.-Y., McManus, M. T., & Mourelatos, Z. (2012). Precursor microRNA-programmed silencing complex assembly pathways in mammals. *Molecular Cell*, *46*(4), 507–517. https://doi.org/10.1016/j.molcel.2012.03.010

Liu, Y., Dou, M., Song, X., Dong, Y., Liu, S., Liu, H., Tao, J., Li, W., Yin, X., & Xu, W. (2019). The emerging role of the piRNA/piwi complex in cancer. *Molecular Cancer*, *18*(1), 123. https://doi.org/10.1186/s12943-019-1052-9

Liu, Z., Wang, J., Cheng, H., Ke, X., Sun, L., Zhang, Q. C., & Wang, H.-W. (2018). Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell*, *173*(5), 1191–1203.e12. https://doi.org/10.1016/j.cell.2018.03.080

Long, Y.-J., Liu, X.-P., Chen, S.-S., Zong, D.-D., Chen, Y., & Chen, P. (2018). miR-34a is involved in CSE-induced apoptosis of human pulmonary microvascular endothelial cells by targeting Notch-1 receptor protein. *Respiratory Research*, *19*(1), 21. https://doi.org/10.1186/s12931-018-0722-2

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., & Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. *Nature*, *435*(7043), 834–838. https://doi.org/10.1038/nature03702

Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E., & Kutay, U. (2004). Nuclear export of microRNA precursors. *Science (New York, N.Y.)*, *303*(5654), 95–98. https://doi.org/10.1126/science.1090599

Luo, J., Wang, Y., Yuan, J., Zhao, Z., & Lu, J. (2018). MicroRNA duplication accelerates the recruitment of new targets during vertebrate evolution. *RNA (New York, N.Y.)*, *24*(6), 787–802. https://doi.org/10.1261/rna.062752.117

Ma, E., MacRae, I. J., Kirsch, J. F., & Doudna, J. A. (2008). Autoinhibition of human dicer by its internal helicase domain. *Journal of Molecular Biology*, *380*(1), 237–243. https://doi.org/10.1016/j.jmb.2008.05.005

Ma, E., Zhou, K., Kidwell, M. A., & Doudna, J. A. (2012). Coordinated activities of human dicer domains in regulatory RNA processing. *Journal of Molecular Biology*, *422*(4), 466–476. https://doi.org/10.1016/j.jmb.2012.06.009

Ma, H., Wu, Y., Choi, J.-G., & Wu, H. (2013). Lower and upper stem-single-stranded RNA junctions together determine the Drosha cleavage site. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(51), 20687–20692. https://doi.org/10.1073/pnas.1311639110

Ma, J.-B., Ye, K., & Patel, D. J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*, *429*(6989), 318–322. https://doi.org/10.1038/nature02519

MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V., & Doudna, J. A. (2008). In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(2), 512–517. https://doi.org/10.1073/pnas.0710869105

Macrae, I. J., Zhou, K., Li, F., Repic, A., Brooks, A. N., Cande, W. Z., Adams, P. D., & Doudna, J. A. (2006). Structural basis for double-stranded RNA processing by Dicer. *Science (New York, N.Y.)*, *311*(5758), 195–198. https://doi.org/10.1126/science.1121638

Maniataki, E., & Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes & Development*, *19*(24), 2979–2990. https://doi.org/10.1101/gad.1384005

Maroney, P. A., Yu, Y., Fisher, J., & Nilsen, T. W. (2006). Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nature Structural & Molecular Biology*, *13*(12), 1102–1107. https://doi.org/10.1038/nsmb1174

Marsico, A., Huska, M. R., Lasserre, J., Hu, H., Vucicevic, D., Musahl, A., Orom, U., & Vingron, M. (2013). PROmiRNA: A new miRNA promoter recognition method uncovers the complex regulation of intronic miRNAs. *Genome Biology*, *14*(8), R84. https://doi.org/10.1186/gb-2013-14-8-r84

Martí, E., Pantano, L., Bañez-Coronel, M., Llorens, F., Miñones-Moyano, E., Porta, S., Sumoy, L., Ferrer, I., & Estivill, X. (2010). A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. *Nucleic Acids Research*, *38*(20), 7219–7235. https://doi.org/10.1093/nar/gkq575

Martinez, I., Hayes, K. E., Barr, J. A., Harold, A. D., Xie, M., Bukhari, S. I. A., Vasudevan, S., Steitz, J. A., & DiMaio, D. (2017). An Exportin-1-dependent microRNA biogenesis pathway during human cell quiescence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(25), E4961–E4970. https://doi.org/10.1073/pnas.1618732114

Marzec, M. (2020). New insights into the function of mammalian Argonaute2. *PLoS genetics*, *16*(11), e1009058. https://doi.org/10.1371/journal.pgen.1009058

Masuda, K., Kuwano, Y., Nishida, K., Rokutan, K., & Imoto, I. (2013). NF90 in Posttranscriptional Gene Regulation and MicroRNA Biogenesis. *International Journal of Molecular Sciences*, *14*(8), 17111–17121. https://doi.org/10.3390/ijms140817111

Matsui, M., Chu, Y., Zhang, H., Gagnon, K. T., Shaikh, S., Kuchimanchi, S., Manoharan, M., Corey, D. R., & Janowski, B. A. (2013). Promoter RNA links transcriptional regulation of inflammatory pathway genes. *Nucleic Acids Research*, *41*(22), 10086–10109. https://doi.org/10.1093/nar/gkt777

Medina, P. P., Nolde, M., & Slack, F. J. (2010). OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. *Nature*, *467*(7311), 86–90. https://doi.org/10.1038/nature09284

Meister, G. (2013). Argonaute proteins: Functional insights and emerging roles. *Nature Reviews. Genetics*, *14*(7), 447–459. https://doi.org/10.1038/nrg3462

Meister, G., Landthaler, M., Peters, L., Chen, P. Y., Urlaub, H., Lührmann, R., & Tuschl, T. (2005). Identification of novel argonaute-associated proteins. *Current biology: CB*, *15*(23), 2149–2155. https://doi.org/10.1016/j.cub.2005.10.048

Melo, S. A., Moutinho, C., Ropero, S., Calin, G. A., Rossi, S., Spizzo, R., Fernandez, A. F., Davalos, V., Villanueva, A., Montoya, G., Yamamoto, H., Schwartz, S., & Esteller, M. (2010). A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer Cell*, *18*(4), 303–315. https://doi.org/10.1016/j.ccr.2010.09.007

Mendell, J. T., & Olson, E. N. (2012). MicroRNAs in stress signaling and human disease. *Cell*, *148*(6), 1172–1187. https://doi.org/10.1016/j.cell.2012.02.005

Meroueh, M., & Chow, C. S. (1999). Thermodynamics of RNA hairpins containing single internal mismatches. *Nucleic Acids Research*, *27*(4), 1118–1125. https://doi.org/10.1093/nar/27.4.1118

Miao, L., Yao, H., Li, C., Pu, M., Yao, X., Yang, H., Qi, X., Ren, J., & Wang, Y. (2016). A dual inhibition: microRNA-552 suppresses both transcription and translation

of cytochrome P450 2E1. *Biochimica Et Biophysica Acta*, *1859*(4), 650–662. https://doi.org/10.1016/j.bbagrm.2016.02.016

Michlewski, G., & Cáceres, J. F. (2010). Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nature Structural & Molecular Biology*, *17*(8), 1011–1018. https://doi.org/10.1038/nsmb.1874

Michlewski, G., & Cáceres, J. F. (2019). Post-transcriptional control of miRNA biogenesis. *RNA (New York, N.Y.)*, *25*(1), 1–16. https://doi.org/10.1261/rna.068692.118

Michlewski, G., Guil, S., & Cáceres, J. F. (2010). Stimulation of pri-miR-18a processing by hnRNP A1. *Advances in Experimental Medicine and Biology*, *700*, 28–35. https://doi.org/10.1007/978-1-4419-7823-3_3

Michlewski, G., Guil, S., Semple, C. A., & Cáceres, J. F. (2008). Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Molecular Cell*, *32*(3), 383–393. https://doi.org/10.1016/j.molcel.2008.10.013

Min, K.-W., Jo, M. H., Shin, S., Davila, S., Zealy, R. W., Kang, S. I., Lloyd, L. T., Hohng, S., & Yoon, J.-H. (2017). AUF1 facilitates microRNA-mediated gene silencing. *Nucleic Acids Research*, *45*(10), 6064–6073. https://doi.org/10.1093/nar/gkx149

Mollaei, H., Safaralizadeh, R., & Rostami, Z. (2019). MicroRNA replacement therapy in cancer. *Journal of Cellular Physiology*, *234*(8), 12369–12384. https://doi.org/10.1002/jcp.28058

Monteys, A. M., Spengler, R. M., Wan, J., Tecedor, L., Lennox, K. A., Xing, Y., & Davidson, B. L. (2010). Structure and activity of putative intronic miRNA promoters. *RNA (New York, N.Y.)*, *16*(3), 495–505. https://doi.org/10.1261/rna.1731910

Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., & Proudfoot, N. J. (2008). Primary microRNA transcripts are processed co-transcriptionally. *Nature Structural & Molecular Biology*, *15*(9), 902–909. https://doi.org/10.1038/nsmb.1475

Morlando, M., Dini Modigliani, S., Torrelli, G., Rosa, A., Di Carlo, V., Caffarelli, E., & Bozzoni, I. (2012). FUS stimulates microRNA biogenesis by facilitating

co-transcriptional Drosha recruitment. *The EMBO journal*, *31*(24), 4502–4510. https://doi.org/10.1038/emboj.2012.319

Mukherjee, K., Ghoshal, B., Ghosh, S., Chakrabarty, Y., Shwetha, S., Das, S., & Bhattacharyya, S. N. (2016). Reversible HuR-microRNA binding controls extracellular export of miR-122 and augments stress response. *EMBO reports*, *17*(8), 1184–1203. https://doi.org/10.15252/embr.201541930

Mukherjee, K., Campos, H., & Kolaczkowski, B. (2013). Evolution of animal and plant dicers: Early parallel duplications and recurrent adaptation of antiviral RNA binding in plants. *Molecular Biology and Evolution*, *30*(3), 627–641. https://doi.org/10.1093/molbev/mss263

Nakadai, T., Fukuda, A., Shimada, M., Nishimura, K., & Hisatake, K. (2015). The RNA binding complexes NF45-NF90 and NF45-NF110 associate dynamically with the c-fos gene and function as transcriptional coactivators. *The Journal of Biological Chemistry*, *290*(44), 26832–26845. https://doi.org/10.1074/jbc.M115.688317

Newman, M. A., Thomson, J. M., & Hammond, S. M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA (New York, N.Y.)*, *14*(8), 1539–1549. https://doi.org/10.1261/rna.1155108

Ngo, T. D., Partin, A. C., & Nam, Y. (2019). RNA Specificity and Autoregulation of DDX17, a Modulator of MicroRNA Biogenesis. *Cell Reports*, *29*(12), 4024–4035.e5. https://doi.org/10.1016/j.celrep.2019.11.059

Nguyen, T. A., Jo, M. H., Choi, Y.-G., Park, J., Kwon, S. C., Hohng, S., Kim, V. N., & Woo, J.-S. (2015). Functional Anatomy of the Human Microprocessor. *Cell*, *161*(6), 1374–1387. https://doi.org/10.1016/j.cell.2015.05.010

Nguyen, T. A., Park, J., Dang, T. L., Choi, Y.-G., & Kim, V. N. (2018). Microprocessor depends on hemin to recognize the apical loop of primary microRNA. *Nucleic Acids Research*, *46*(11), 5726–5736. https://doi.org/10.1093/nar/gky248

Nicholson, A. L., & Pasquinelli, A. E. (2019). Tales of Detailed Poly(A) Tails. *Trends in cell biology*, *29*(3), 191–200. https://doi.org/10.1016/j.tcb.2018.11.002

Noland, C. L., & Doudna, J. A. (2013). Multiple sensors ensure guide strand selection in human RNAi pathways. *RNA (New York, N.Y.)*, *19*(5), 639–648. https://doi.org/10.1261/rna.037424.112

Nussbacher, J. K., & Yeo, G. W. (2018). Systematic Discovery of RNA Binding Proteins that Regulate MicroRNA Levels. *Molecular Cell*, *69*(6), 1005–1016.e7. https://doi.org/10.1016/j.molcel.2018.02.012

Obbard, D. J., Gordon, K. H. J., Buck, A. H., & Jiggins, F. M. (2009). The evolution of RNAi as a defence against viruses and transposable elements. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1513), 99–115. https://doi.org/10.1098/rstb.2008.0168

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Frontiers in Endocrinology*, *9*, 402. https://doi.org/10.3389/fendo.2018.00402

Okada, C., Yamashita, E., Lee, S. J., Shibata, S., Katahira, J., Nakagawa, A., Yoneda, Y., & Tsukihara, T. (2009). A high-resolution structure of the pre-microRNA nuclear export machinery. *Science (New York, N.Y.)*, *326*(5957), 1275–1279. https://doi.org/10.1126/science.1178705

Ozsolak, F., Poling, L. L., Wang, Z., Liu, H., Liu, X. S., Roeder, R. G., Zhang, X., Song, J. S., & Fisher, D. E. (2008). Chromatin structure analyses identify miRNA promoters. *Genes & Development*, *22*(22), 3172–3183. https://doi.org/10.1101/gad.1706508

Pagani, M., Rossetti, G., Panzeri, I., Candia, P. d., Bonnal, R. J. P., Rossi, R. L., Geginat, J., & Abrignani, S. (2013). Role of microRNAs and long-non-coding RNAs in CD4+ T-cell differentiation [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/imr.12055]. *Immunological Reviews*, *253*(1), 82–96. https://doi.org/https://doi.org/10.1111/imr.12055

Papapetrou, E. P., Korkola, J. E., & Sadelain, M. (2010). A genetic strategy for single and combinatorial analysis of miRNA function in mammalian hematopoietic stem cells. *Stem Cells (Dayton, Ohio)*, *28*(2), 287–296. https://doi.org/10.1002/stem.257

Park, J.-E., Heo, I., Tian, Y., Simanshu, D. K., Chang, H., Jee, D., Patel, D. J., & Kim, V. N. (2011). Dicer recognizes the 5′ end of RNA for efficient and accurate processing. *Nature*, *475*(7355), 201–205. https://doi.org/10.1038/nature10198

Park, M. S., Phan, H.-D., Busch, F., Hinckley, S. H., Brackbill, J. A., Wysocki, V. H., & Nakanishi, K. (2017). Human Argonaute3 has slicer activity. *Nucleic Acids Research*, *45*(20), 11867–11877. https://doi.org/10.1093/nar/gkx916

Parker, J. S., Roe, S. M., & Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature*, *434*(7033), 663–666. https://doi.org/10.1038/nature03462

Paroo, Z., Ye, X., Chen, S., & Liu, Q. (2009). Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. *Cell*, *139*(1), 112–122. https://doi.org/10.1016/j.cell.2009.06.044

Parrott, A. M., Walsh, M. R., Reichman, T. W., & Mathews, M. B. (2005). RNA binding and phosphorylation determine the intracellular distribution of nuclear factors 90 and 110. *Journal of Molecular Biology*, *348*(2), 281–293. https://doi.org/10.1016/j.jmb.2005.02.047

Partin, A. C., Ngo, T. D., Herrell, E., Jeong, B.-C., Hon, G., & Nam, Y. (2017). Heme enables proper positioning of Drosha and DGCR8 on primary microRNAs. *Nature Communications*, *8*(1), 1737. https://doi.org/10.1038/s41467-017-01713-y

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Müller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E., & Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA [Number: 6808 Publisher: Nature Publishing Group]. *Nature*, *408*(6808), 86–89. https://doi.org/10.1038/35040556

Patel, R. C., & Sen, G. C. (1998). PACT, a protein activator of the interferon-induced protein kinase, PKR. *The EMBO journal*, *17*(15), 4379–4390. https://doi.org/10.1093/emboj/17.15.4379

Patel, V. D., & Capra, J. A. (2017). Ancient human miRNAs are more likely to have broad functions and disease associations than young miRNAs. *BMC genomics*, *18*(1), 672. https://doi.org/10.1186/s12864-017-4073-z

Patiño, C., Haenni, A.-L., & Urcuqui-Inchima, S. (2015). NF90 isoforms, a new family of cellular proteins involved in viral replication? *Biochimie*, *108*, 20–24. https://doi.org/10.1016/j.biochi.2014.10.022

Patterson, D. G., Roberts, J. T., King, V. M., Houserova, D., Barnhill, E. C., Crucello, A., Polska, C. J., Brantley, L. W., Kaufman, G. C., Nguyen, M., Santana, M. W., Schiller, I. A., Spicciani, J. S., Zapata, A. K., Miller, M. M., Sherman, T. D., Ma, R., Zhao, H., Arora, R., … Borchert, G. M. (2017). Human snoRNA-93 is processed into a microRNA-like RNA that promotes breast cancer cell invasion. *NPJ breast cancer*, *3*, 25. https://doi.org/10.1038/s41523-017-0032-8

Pawlicki, J. M., & Steitz, J. A. (2008). Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *The Journal of Cell Biology*, *182*(1), 61–76. https://doi.org/10.1083/jcb.200803111

Pei, Y., Zhu, P., Dang, Y., Wu, J., Yang, X., Wan, B., Liu, J. O., Yi, Q., & Yu, L. (2008). Nuclear export of NF90 to stabilize IL-2 mRNA is mediated by AKT-dependent phosphorylation at Ser647 in response to CD28 costimulation. *Journal of Immunology (Baltimore, Md.: 1950)*, *180*(1), 222–229. https://doi.org/10.4049/jimmunol.180.1.222

Peng, Y., & Croce, C. M. (2016). The role of MicroRNAs in human cancer. *Signal Transduction and Targeted Therapy*, *1*, 15004. https://doi.org/10.1038/sigtrans.2015.4

Pfeifer, I., Elsby, R., Fernandez, M., Faria, P. A., Nussenzveig, D. R., Lossos, I. S., Fontoura, B. M. A., Martin, W. D., & Barber, G. N. (2008). NFAR-1 and -2 modulate translation and are required for efficient host defense. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(11), 4173–4178. https://doi.org/10.1073/pnas.0711222105

Pianigiani, G., Licastro, D., Fortugno, P., Castiglia, D., Petrovic, I., & Pagani, F. (2018). Microprocessor-dependent processing of splice site overlapping microRNA

exons does not result in changes in alternative splicing. *RNA (New York, N.Y.)*, *24*(9), 1158–1171. https://doi.org/10.1261/rna.063438.117

Piriyapongsa, J., Mariño-Ramírez, L., & Jordan, I. K. (2007). Origin and evolution of human microRNAs from transposable elements. *Genetics*, *176*(2), 1323–1337. https://doi.org/10.1534/genetics.107.072553

Pitchiaya, S., Heinicke, L. A., Park, J. I., Cameron, E. J., & Walter, N. G. (2017). Resolving sub-cellular miRNA trafficking and turnover at single-molecule resolution. *Cell reports*, *19*(3), 630–642. https://doi.org/10.1016/j.celrep.2017.03.075

Place, R. F., Li, L.-C., Pookot, D., Noonan, E. J., & Dahiya, R. (2008). MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(5), 1608–1613. https://doi.org/10.1073/pnas.0707594105

Pokornowska, M., Milewski, M. C., Ciechanowska, K., Szczepańska, A., Wojnicka, M., Radogostowicz, Z., Figlerowicz, M., & Kurzynska-Kokorniak, A. (2020). The RNA-RNA base pairing potential of human Dicer and Ago2 proteins. *Cellular and molecular life sciences: CMLS*, *77*(16), 3231–3244. https://doi.org/10.1007/s00018-019-03344-6

Pushpavalli, S. N. C. V. L., Sarkar, A., Bag, I., Hunt, C. R., Ramaiah, M. J., Pandita, T. K., Bhadra, U., & Pal-Bhadra, M. (2014). Argonaute-1 functions as a mitotic regulator by controlling Cyclin B during Drosophila early embryogenesis. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, *28*(2), 655–666. https://doi.org/10.1096/fj.13-231167

Qi, H. H., Ongusaha, P. P., Myllyharju, J., Cheng, D., Pakkanen, O., Shi, Y., Lee, S. W., Peng, J., & Shi, Y. (2008). Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature*, *455*(7211), 421–424. https://doi.org/10.1038/nature07186

Qin, S., Jin, P., Zhou, X., Chen, L., & Ma, F. (2015). The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human. *PloS One*, *10*(6), e0131365. https://doi.org/10.1371/journal.pone.0131365

Quévillon Huberdeau, M., Zeitler, D. M., Hauptmann, J., Bruckmann, A., Fressigné, L., Danner, J., Piquet, S., Strieder, N., Engelmann, J. C., Jannot, G., Deutzmann, R., Simard, M. J., & Meister, G. (2017). Phosphorylation of Argonaute proteins affects mRNA binding and is essential for microRNA-guided gene silencing in vivo. *The EMBO journal*, *36*(14), 2088–2106. https://doi.org/10.15252/embj.201696386

Quick-Cleveland, J., Jacob, J. P., Weitz, S. H., Shoffner, G., Senturia, R., & Guo, F. (2014). The DGCR8 RNA-binding heme domain recognizes primary microRNAs by clamping the hairpin. *Cell Reports*, *7*(6), 1994–2005. https://doi.org/10.1016/j.celrep.2014.05.013

Rajyaguru, P., & Parker, R. (2012). RGG motif proteins. *Cell cycle (Georgetown, Tex.)*, *11*(14). https://doi.org/10.4161/cc.20716

Ramalingam, P., Palanichamy, J. K., Singh, A., Das, P., Bhagat, M., Kassab, M. A., Sinha, S., & Chattopadhyay, P. (2014). Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. *RNA (New York, N.Y.)*, *20*(1), 76–87. https://doi.org/10.1261/rna.041814.113

Reichholf, B., Herzog, V. A., Fasching, N., Manzenreither, R. A., Sowemimo, I., & Ameres, S. L. (2019). Time-Resolved Small RNA Sequencing Unravels the Molecular Principles of MicroRNA Homeostasis. *Molecular Cell*, *75*(4), 756–768.e7. https://doi.org/10.1016/j.molcel.2019.06.018

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, *403*(6772), 901–906. https://doi.org/10.1038/35002607

Remenyi, J., Bajan, S., Fuller-Pace, F. V., Arthur, J. S. C., & Hutvagner, G. (2016). The loop structure and the RNA helicase p72/DDX17 influence the processing efficiency of the mice miR-132. *Scientific Reports*, *6*, 22848. https://doi.org/10.1038/srep22848

Rolando, C., Erni, A., Grison, A., Beattie, R., Engler, A., Gokhale, P. J., Milo, M., Wegleiter, T., Jessberger, S., & Taylor, V. (2016). Multipotency of Adult Hippocampal NSCs In Vivo Is Restricted by Drosha/NFIB. *Cell Stem Cell*, *19*(5), 653–662. https://doi.org/10.1016/j.stem.2016.07.003

Romano, G., Veneziano, D., Acunzo, M., & Croce, C. M. (2017). Small non-coding RNA and cancer. *Carcinogenesis*, *38*(5), 485–491. https://doi.org/10.1093/carcin/bgx026

Rotival, M., Siddle, K. J., Silvert, M., Pothlichet, J., Quach, H., & Quintana-Murci, L. (2020). Population variation in miRNAs and isomiRs and their impact on human immunity to infection. *Genome Biology*, *21*(1), 187. https://doi.org/10.1186/s13059-020-02098-w

Ruby, J. G., Jan, C. H., & Bartel, D. P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature*, *448*(7149), 83–86. https://doi.org/10.1038/nature05983

Rüdel, S., Wang, Y., Lenobel, R., Körner, R., Hsiao, H.-H., Urlaub, H., Patel, D., & Meister, G. (2011). Phosphorylation of human Argonaute proteins affects small RNA binding. *Nucleic Acids Research*, *39*(6), 2330–2343. https://doi.org/10.1093/nar/gkq1032

Rüegger, S., & Großhans, H. (2012). MicroRNA turnover: When, how, and why. *Trends in Biochemical Sciences*, *37*(10), 436–446. https://doi.org/10.1016/j.tibs.2012.07.002

Rupaimoole, R., & Slack, F. J. (2017). MicroRNA therapeutics: Towards a new era for the management of cancer and other diseases [Number: 3 Publisher: Nature Publishing Group]. *Nature Reviews Drug Discovery*, *16*(3), 203–222. https://doi.org/10.1038/nrd.2016.246

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E. E., Nitsch, R., & Wulczyn, F. G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nature Cell Biology*, *10*(8), 987–993. https://doi.org/10.1038/ncb1759

Sahoo, M. R., Gaikwad, S., Khuperkar, D., Ashok, M., Helen, M., Yadav, S. K., Singh, A., Magre, I., Deshmukh, P., Dhanvijay, S., Sahoo, P. K., Ramtirtha, Y., Madhusudhan, M. S., Gayathri, P., Seshadri, V., & Joseph, J. (2017). Nup358 binds to AGO proteins through its SUMO-interacting motifs and promotes the association of target mRNA with miRISC. *EMBO reports*, *18*(2), 241–263. https://doi.org/10.15252/embr.201642386

Sakamoto, S., Aoki, K., Higuchi, T., Todaka, H., Morisawa, K., Tamaki, N., Hatano, E., Fukushima, A., Taniguchi, T., & Agata, Y. (2009). The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. *Molecular and Cellular Biology*, *29*(13), 3754–3769. https://doi.org/10.1128/MCB.01836-08

Sakamoto, S., Morisawa, K., Ota, K., Nie, J., & Taniguchi, T. (1999). A Binding Protein to the DNase I Hypersensitive Site II in HLA-DR Gene Was Identified as NF90 [Publisher: American Chemical Society]. *Biochemistry*, *38*(11), 3355–3361. https://doi.org/10.1021/bi982099g

Sala, L., Chandrasekhar, S., & Vidigal, J. A. (2020). AGO unchained: Canonical and non-canonical roles of Argonaute proteins in mammals. *Frontiers in Bioscience (Landmark Edition)*, *25*, 1–42.

Sambandan, S., Akbalik, G., Kochen, L., Rinne, J., Kahlstatt, J., Glock, C., Tushev, G., Alvarez-Castelao, B., Heckel, A., & Schuman, E. M. (2017). Activity-dependent spatially localized miRNA maturation in neuronal dendrites. *Science (New York, N.Y.)*, *355*(6325), 634–637. https://doi.org/10.1126/science.aaf8995

Sand, M., Skrygan, M., Georgas, D., Arenz, C., Gambichler, T., Sand, D., Altmeyer, P., & Bechara, F. G. (2012). Expression levels of the microRNA maturing microprocessor complex component DGCR8 and the RNA-induced silencing complex (RISC) components argonaute-1, argonaute-2, PACT, TARBP1, and TARBP2 in epithelial skin cancer. *Molecular Carcinogenesis*, *51*(11), 916–922. https://doi.org/10.1002/mc.20861

Saw, P. E., & Song, E.-W. (2020). siRNA therapeutics: A clinical reality. *Science China Life Sciences*, *63*(4), 485–500. https://doi.org/10.1007/s11427-018-9438-y

Schirle, N. T., & MacRae, I. J. (2012). The crystal structure of human Argonaute2. *Science (New York, N.Y.)*, *336*(6084), 1037–1040. https://doi.org/10.1126/science.1221551

Schirle, N. T., Sheu-Gruttadauria, J., Chandradoss, S. D., Joo, C., & MacRae, I. J. (2015). Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. *eLife*, *4*. https://doi.org/10.7554/eLife.07646

Schmidt, T., Knick, P., Lilie, H., Friedrich, S., Golbik, R. P., & Behrens, S.-E. (2016). Coordinated Action of Two Double-Stranded RNA Binding Motifs and an RGG Motif Enables Nuclear Factor 90 To Flexibly Target Different RNA Substrates. *Biochemistry*, *55*(6), 948–959. https://doi.org/10.1021/acs.biochem.5b01072

Schmidt, T., Knick, P., Lilie, H., Friedrich, S., Golbik, R. P., & Behrens, S.-E. (2017). The properties of the RNA-binding protein NF90 are considerably modulated by complex formation with NF45. *The Biochemical Journal*, *474*(2), 259–280. https://doi.org/10.1042/BCJ20160790

Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., & Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, *115*(2), 199–208. https://doi.org/10.1016/s0092-8674(03)00759-1

Seco-Cervera, M., González-Rodríguez, D., Ibáñez-Cabellos, J. S., Peiró-Chova, L., Pallardó, F. V., & García-Giménez, J. L. (2018). Small RNA-seq analysis of circulating miRNAs to identify phenotypic variability in Friedreich's ataxia patients. *Scientific Data*, *5*, 180021. https://doi.org/10.1038/sdata.2018.21

Shabalina, S. A., & Koonin, E. V. (2008). Origins and evolution of eukaryotic RNA interference. *Trends in Ecology & Evolution*, *23*(10), 578–587. https://doi.org/10.1016/j.tree.2008.06.005

Shabman, R. S., Leung, D. W., Johnson, J., Glennon, N., Gulcicek, E. E., Stone, K. L., Leung, L., Hensley, L., Amarasinghe, G. K., & Basler, C. F. (2011). DRBP76 associates with Ebola virus VP35 and suppresses viral polymerase function. *The Journal of Infectious Diseases*, *204 Suppl 3*, S911–918. https://doi.org/10.1093/infdis/jir343

Shamanna, R. A., Hoque, M., Pe'ery, T., & Mathews, M. B. (2013). Induction of p53, p21 and apoptosis by silencing the NF90/NF45 complex in human papilloma virus-transformed cervical carcinoma cells. *Oncogene*, *32*(43), 5176–5185. https://doi.org/10.1038/onc.2012.533

Shapiro, J. S., Schmid, S., Aguado, L. C., Sabin, L. R., Yasunaga, A., Shim, J. V., Sachs, D., Cherry, S., & tenOever, B. R. (2014). Drosha as an interferon-independent antiviral factor. *Proceedings of the National Academy of*

Sciences of the United States of America, *111*(19), 7108–7113. https://doi.org/10.1073/pnas.1319635111

Shen, E., Wang, X., Liu, X., Lv, M., Zhang, L., Zhu, G., & Sun, Z. (2020). MicroRNA-93-5p promotes epithelial-mesenchymal transition in gastric cancer by repressing tumor suppressor AHNAK expression. *Cancer Cell International*, *20*, 76. https://doi.org/10.1186/s12935-019-1092-7

Shen, J., Xia, W., Khotskaya, Y. B., Huo, L., Nakanishi, K., Lim, S.-O., Du, Y., Wang, Y., Chang, W.-C., Chen, C.-H., Hsu, J. L., Wu, Y., Lam, Y. C., James, B. P., Liu, X., Liu, C.-G., Patel, D. J., & Hung, M.-C. (2013). EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. *Nature*, *497*(7449), 383–387. https://doi.org/10.1038/nature12080

Shenoy, A., & Blelloch, R. H. (2014). Regulation of microRNA function in somatic stem cell proliferation and differentiation. *Nature Reviews. Molecular Cell Biology*, *15*(9), 565–576. https://doi.org/10.1038/nrm3854

Sheu-Gruttadauria, J., & MacRae, I. J. (2017). Structural Foundations of RNA Silencing by Argonaute. *Journal of Molecular Biology*, *429*(17), 2619–2639. https://doi.org/10.1016/j.jmb.2017.07.018

Sheu-Gruttadauria, J., Xiao, Y., Gebert, L. F., & MacRae, I. J. (2019). Beyond the seed: Structural basis for supplementary microRNA targeting by human Argonaute2. *The EMBO journal*, *38*(13), e101153. https://doi.org/10.15252/embj.2018101153

Shi, L., Qiu, D., Zhao, G., Corthesy, B., Lees-Miller, S., Reeves, W. H., & Kao, P. N. (2007). Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells. *Nucleic Acids Research*, *35*(7), 2302–2310. https://doi.org/10.1093/nar/gkm117

Shin, C., Nam, J.-W., Farh, K. K.-H., Chiang, H. R., Shkumatava, A., & Bartel, D. P. (2010). Expanding the microRNA targeting code: Functional sites with centered pairing. *Molecular Cell*, *38*(6), 789–802. https://doi.org/10.1016/j.molcel.2010.06.005

Shrestha, A., Mukhametshina, R. T., Taghizadeh, S., Vásquez-Pacheco, E., Cabrera-Fuentes, H., Rizvanov, A., Mari, B., Carraro, G., & Bellusci, S.

(2017). MicroRNA-142 is a multifaceted regulator in organogenesis, homeostasis, and disease [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/dvdy.24477]. *Developmental Dynamics*, *246*(4), 285–290. https://doi.org/https://doi.org/10.1002/dvdy.24477

Shuman, S. (2015). RNA capping: Progress and prospects. *RNA (New York, N.Y.)*, *21*(4), 735–737. https://doi.org/10.1261/rna.049973.115

Smith, K. N., Starmer, J., Miller, S. C., Sethupathy, P., & Magnuson, T. (2017). Long Noncoding RNA Moderates MicroRNA Activity to Maintain Self-Renewal in Embryonic Stem Cells. *Stem Cell Reports*, *9*(1), 108–121. https://doi.org/10.1016/j.stemcr.2017.05.005

Smith, N. L., & Miskimins, W. K. (2011). Phosphorylation at serine 482 affects stability of NF90 and its functional role in mitosis. *Cell Proliferation*, *44*(2), 147–155. https://doi.org/10.1111/j.1365-2184.2011.00742.x

Soifer, H. S., Sano, M., Sakurai, K., Chomchan, P., Saetrom, P., Sherman, M. A., Collingwood, M. A., Behlke, M. A., & Rossi, J. J. (2008). A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Research*, *36*(20), 6511–6522. https://doi.org/10.1093/nar/gkn687

Song, D., Huang, H., Wang, J., Zhao, Y., Hu, X., He, F., Yu, L., & Wu, J. (2017). NF90 regulates PARP1 mRNA stability in hepatocellular carcinoma. *Biochemical and Biophysical Research Communications*, *488*(1), 211–217. https://doi.org/10.1016/j.bbrc.2017.05.037

Song, J., Zhuang, Y., Zhu, C., Meng, H., Lu, B., Xie, B., Peng, J., Li, M., & Yi, C. (2020). Differential roles of human PUS10 in miRNA processing and tRNA pseudouridylation. *Nature Chemical Biology*, *16*(2), 160–169. https://doi.org/10.1038/s41589-019-0420-5

Stark, A., Brennecke, J., Bushati, N., Russell, R. B., & Cohen, S. M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell*, *123*(6), 1133–1146. https://doi.org/10.1016/j.cell.2005.11.023

Stavast, C. J., & Erkeland, S. J. (2019). The Non-Canonical Aspects of MicroRNAs: Many Roads to Gene Regulation. *Cells*, *8*(11). https://doi.org/10.3390/cells8111465

Steiman-Shimony, A., Shtrikman, O., & Margalit, H. (2018). Assessing the functional association of intronic miRNAs with their host genes. *RNA (New York, N.Y.)*, *24*(8), 991–1004. https://doi.org/10.1261/rna.064386.117

Stein, P., Zeng, F., Pan, H., & Schultz, R. M. (2005). Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes. *Developmental Biology*, *286*(2), 464–471. https://doi.org/10.1016/j.ydbio.2005.08.015

Su, W., Hopkins, S., Nesser, N. K., Sopher, B., Silvestroni, A., Ammanuel, S., Jayadev, S., Möller, T., Weinstein, J., & Garden, G. A. (2014). The p53 transcription factor modulates microglia behavior through microRNA-dependent regulation of c-Maf. *Journal of Immunology (Baltimore, Md.: 1950)*, *192*(1), 358–366. https://doi.org/10.4049/jimmunol.1301397

Subtelny, A. O., Eichhorn, S. W., Chen, G. R., Sive, H., & Bartel, D. P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature*, *508*(7494), 66–71. https://doi.org/10.1038/nature13007

Sun, C., Liu, W., Lu, Z., Li, Y., Liu, S., Tang, Z., Yan, Y., Li, Z., Feng, H., Zhang, D., Liu, Y., Fang, Z.-Z., Jiang, C., Ding, Q., Jiang, J., & Ying, H. (2021). Hepatic miR-378 modulates serum cholesterol levels by regulating hepatic bile acid synthesis. *Theranostics*, *11*(9), 4363–4380. https://doi.org/10.7150/thno.53624

Sun, H.-L., Cui, R., Zhou, J., Teng, K.-Y., Hsiao, Y.-H., Nakanishi, K., Fassan, M., Luo, Z., Shi, G., Tili, E., Kutay, H., Lovat, F., Vicentini, C., Huang, H.-L., Wang, S.-W., Kim, T., Zanesi, N., Jeon, Y.-J., Lee, T. J., … Croce, C. M. (2016). ERK Activation Globally Downregulates miRNAs through Phosphorylating Exportin-5. *Cancer Cell*, *30*(5), 723–736. https://doi.org/10.1016/j.ccell.2016.10.001

Sun, Y., Ji, F., Kumar, M. R., Zheng, X., Xiao, Y., Liu, N., Shi, J., Wong, L., Forgues, M., Qin, L.-X., Tang, Z.-Y., Zhao, X., Wang, X. W., & Ji, J. (2017). Transcriptome integration analysis in hepatocellular carcinoma reveals

discordant intronic miRNA-host gene pairs in expression. *International Journal of Biological Sciences*, *13*(11), 1438–1449. https://doi.org/10.7150/ijbs.20836

Svoboda, P. (2014). Renaissance of mammalian endogenous RNAi. *FEBS letters*, *588*(15), 2550–2556. https://doi.org/10.1016/j.febslet.2014.05.030

Svobodova, E., Kubikova, J., & Svoboda, P. (2016). Production of small RNAs by mammalian Dicer. *Pflugers Archiv: European Journal of Physiology*, *468*(6), 1089–1102. https://doi.org/10.1007/s00424-016-1817-6

Takeshita, D., Zenno, S., Lee, W. C., Nagata, K., Saigo, K., & Tanokura, M. (2007). Homodimeric structure and double-stranded RNA cleavage activity of the C-terminal RNase III domain of human dicer. *Journal of Molecular Biology*, *374*(1), 106–120. https://doi.org/10.1016/j.jmb.2007.08.069

Tan, G. C., Chan, E., Molnar, A., Sarkar, R., Alexieva, D., Isa, I. M., Robinson, S., Zhang, S., Ellis, P., Langford, C. F., Guillot, P. V., Chandrashekran, A., Fisk, N. M., Castellano, L., Meister, G., Winston, R. M., Cui, W., Baulcombe, D., & Dibb, N. J. (2014). 5′ isomiR variation is of functional and evolutionary importance. *Nucleic Acids Research*, *42*(14), 9424–9435. https://doi.org/10.1093/nar/gku656

Tang, X., Li, M., Tucker, L., & Ramratnam, B. (2011). Glycogen synthase kinase 3 beta (GSK3) phosphorylates the RNAase III enzyme Drosha at S300 and S302. *PloS One*, *6*(6), e20391. https://doi.org/10.1371/journal.pone.0020391

Taylor, D. W., Ma, E., Shigematsu, H., Cianfrocco, M. A., Noland, C. L., Nagayama, K., Nogales, E., Doudna, J. A., & Wang, H.-W. (2013). Substrate-specific structural rearrangements of human Dicer. *Nature Structural & Molecular Biology*, *20*(6), 662–670. https://doi.org/10.1038/nsmb.2564

Tian, Y., Simanshu, D. K., Ma, J.-B., Park, J.-E., Heo, I., Kim, V. N., & Patel, D. J. (2014). A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Molecular Cell*, *53*(4), 606–616. https://doi.org/10.1016/j.molcel.2014.01.003

Ting, N. S., Kao, P. N., Chan, D. W., Lintott, L. G., & Lees-Miller, S. P. (1998). DNA-dependent protein kinase interacts with antigen receptor response element

binding proteins NF90 and NF45. *The Journal of Biological Chemistry*, *273*(4), 2136–2145. https://doi.org/10.1074/jbc.273.4.2136

Todaka, H., Higuchi, T., Yagyu, K.-i., Sugiyama, Y., Yamaguchi, F., Morisawa, K., Ono, M., Fukushima, A., Tsuda, M., Taniguchi, T., & Sakamoto, S. (2015). Overexpression of NF90-NF45 Represses Myogenic MicroRNA Biogenesis, Resulting in Development of Skeletal Muscle Atrophy and Centronuclear Muscle Fibers. *Molecular and Cellular Biology*, *35*(13), 2295–2308. https://doi.org/10.1128/MCB.01297-14

Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A. D., Filipowicz, W., Ramos, A., Gherzi, R., & Rosenfeld, M. G. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature*, *459*(7249), 1010–1014. https://doi.org/10.1038/nature08025

Treiber, T., Treiber, N., Plessmann, U., Harlander, S., Daiß, J.-L., Eichner, N., Lehmann, G., Schall, K., Urlaub, H., & Meister, G. (2017). A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Molecular Cell*, *66*(2), 270–284.e13. https://doi.org/10.1016/j.molcel.2017.03.014

Triboulet, R., Chang, H.-M., Lapierre, R. J., & Gregory, R. I. (2009). Post-transcriptional control of DGCR8 expression by the Microprocessor. *RNA (New York, N.Y.)*, *15*(6), 1005–1011. https://doi.org/10.1261/rna.1591709

Tsai, V. W. W., Husaini, Y., Sainsbury, A., Brown, D. A., & Breit, S. N. (2018). The MIC-1/GDF15-GFRAL Pathway in Energy Homeostasis: Implications for Obesity, Cachexia, and Other Associated Diseases. *Cell Metabolism*, *28*(3), 353–368. https://doi.org/10.1016/j.cmet.2018.07.018

Tu, C.-C., Zhong, Y., Nguyen, L., Tsai, A., Sridevi, P., Tarn, W.-Y., & Wang, J. Y. J. (2015). The kinase ABL phosphorylates the microprocessor subunit DGCR8 to stimulate primary microRNA processing in response to DNA damage. *Science Signaling*, *8*(383), ra64. https://doi.org/10.1126/scisignal.aaa4468

Urbanek-Trzeciak, M. O., Jaworska, E., & Krzyzosiak, W. J. (2018). miRNAmotif-A Tool for the Prediction of Pre-miRNAProtein Interactions. *International Journal of Molecular Sciences*, *19*(12). https://doi.org/10.3390/ijms19124075

Vance, V., & Vaucheret, H. (2001). RNA silencing in plants–defense and counterdefense. *Science (New York, N.Y.)*, *292*(5525), 2277–2280. https://doi.org/10.1126/science.1061334

van Solingen, C., Bijkerk, R., de Boer, H. C., Rabelink, T. J., & van Zonneveld, A. J. (2015). The Role of microRNA-126 in Vascular Homeostasis. *Current Vascular Pharmacology*, *13*(3), 341–351. https://doi.org/10.2174/15701611113119990017

Viranaicken, W., Gasmi, L., Chauvin, C., Denoulet, P., & Larcher, J.-C. (2006). Identification of a newly spliced exon in the mouse Ilf3 gene generating two long and short isoforms of Ilf3 and NF90. *Genomics*, *88*(5), 622–632. https://doi.org/10.1016/j.ygeno.2006.08.006

Viswanathan, S. R., Daley, G. Q., & Gregory, R. I. (2008). Selective blockade of microRNA processing by Lin28. *Science (New York, N.Y.)*, *320*(5872), 97–100. https://doi.org/10.1126/science.1154040

Vumbaca, F., Phoenix, K. N., Rodriguez-Pinto, D., Han, D. K., & Claffey, K. P. (2008). Double-stranded RNA-binding protein regulates vascular endothelial growth factor mRNA stability, translation, and breast cancer angiogenesis. *Molecular and Cellular Biology*, *28*(2), 772–783. https://doi.org/10.1128/MCB.02078-06

Wada, T., Kikuchi, J., & Furukawa, Y. (2012). Histone deacetylase 1 enhances microRNA processing via deacetylation of DGCR8. *EMBO reports*, *13*(2), 142–149. https://doi.org/10.1038/embor.2011.247

Wahle, E., & Winkler, G. S. (2013). RNA decay machines: Deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochimica Et Biophysica Acta*, *1829*(6-7), 561–570. https://doi.org/10.1016/j.bbagrm.2013.01.003

Wahlquist, C., Jeong, D., Rojas-Muñoz, A., Kho, C., Lee, A., Mitsuyama, S., van Mil, A., Park, W. J., Sluijter, J. P. G., Doevendans, P. A. F., Hajjar, R. J., & Mercola, M. (2014). Inhibition of miR-25 improves cardiac contractility in the failing heart. *Nature*, *508*(7497), 531–535. https://doi.org/10.1038/nature13073

Wang, B., Love, T. M., Call, M. E., Doench, J. G., & Novina, C. D. (2006). Recapitulation of short RNA-directed translational gene silencing in vitro. *Molecular Cell*, *22*(4), 553–560. https://doi.org/10.1016/j.molcel.2006.03.034

Wang, H.-W., Noland, C., Siridechadilok, B., Taylor, D. W., Ma, E., Felderer, K., Doudna, J. A., & Nogales, E. (2009). Structural insights into RNA processing by the human RISC-loading complex. *Nature Structural & Molecular Biology*, *16*(11), 1148–1153. https://doi.org/10.1038/nsmb.1673

Wang, J., Lee, J. E., Riemondy, K., Yu, Y., Marquez, S. M., Lai, E. C., & Yi, R. (2020). XPO5 promotes primary miRNA processing independently of RanGTP. *Nature Communications*, *11*(1), 1845. https://doi.org/10.1038/s41467-020-15598-x

Wang, X.-H., Aliyari, R., Li, W.-X., Li, H.-W., Kim, K., Carthew, R., Atkinson, P., & Ding, S.-W. (2006). RNA interference directs innate immunity against viruses in adult Drosophila. *Science (New York, N.Y.)*, *312*(5772), 452–454. https://doi.org/10.1126/science.1125694

Wei, X., Liu, H., Li, X., & Liu, X. (2019). Over-expression of MiR-122 promotes apoptosis of hepatocellular carcinoma via targeting TLR4. *Annals of Hepatology*, *18*(6), 869–878. https://doi.org/10.1016/j.aohep.2019.07.005

Wen, X., Liu, X., Mao, Y.-P., Yang, X.-J., Wang, Y.-Q., Zhang, P.-P., Lei, Y., Hong, X.-H., He, Q.-M., Ma, J., Liu, N., & Li, Y.-Q. (2018). Long non-coding RNA DANCR stabilizes HIF-1 and promotes metastasis by interacting with NF90/NF45 complex in nasopharyngeal carcinoma. *Theranostics*, *8*(20), 5676–5689. https://doi.org/10.7150/thno.28538

Westholm, J. O., & Lai, E. C. (2011). Mirtrons: microRNA biogenesis via splicing. *Biochimie*, *93*(11), 1897–1904. https://doi.org/10.1016/j.biochi.2011.06.017

Wilkins, C., Dishongh, R., Moore, S. C., Whitt, M. A., Chow, M., & Machaca, K. (2005). RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. *Nature*, *436*(7053), 1044–1047. https://doi.org/10.1038/nature03957

Willimott, S., & Wagner, S. D. (2010). Post-transcriptional and post-translational regulation of Bcl2. *Biochemical Society Transactions*, *38*(6), 1571–1575. https://doi.org/10.1042/BST0381571

Wilson, R. C., Tambe, A., Kidwell, M. A., Noland, C. L., Schneider, C. P., & Doudna, J. A. (2015). Dicer-TRBP complex formation ensures accurate

mammalian microRNA biogenesis. *Molecular Cell*, *57*(3), 397–407. https://doi.org/10.1016/j.molcel.2014.11.030

Winter, J., & Diederichs, S. (2011). Argonaute proteins regulate microRNA stability: Increased microRNA abundance by Argonaute proteins is due to microRNA stabilization. *RNA biology*, *8*(6), 1149–1157. https://doi.org/10.4161/rna.8.6.17665

Winter, J., Jung, S., Keller, S., Gregory, R. I., & Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nature Cell Biology*, *11*(3), 228–234. https://doi.org/10.1038/ncb0309-228

Wolkowicz, U. M., & Cook, A. G. (2012). NF45 dimerizes with NF90, Zfr and SPNR via a conserved domain that has a nucleotidyltransferase fold. *Nucleic Acids Research*, *40*(18), 9356–9368. https://doi.org/10.1093/nar/gks696

Worden, A. Z., Lee, J.-H., Mock, T., Rouzé, P., Simmons, M. P., Aerts, A. L., Allen, A. E., Cuvelier, M. L., Derelle, E., Everett, M. V., Foulon, E., Grimwood, J., Gundlach, H., Henrissat, B., Napoli, C., McDonald, S. M., Parker, M. S., Rombauts, S., Salamov, A., … Grigoriev, I. V. (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes Micromonas. *Science (New York, N.Y.)*, *324*(5924), 268–272. https://doi.org/10.1126/science.1167222

Wostenberg, C., Lary, J. W., Sahu, D., Acevedo, R., Quarles, K. A., Cole, J. L., & Showalter, S. A. (2012). The role of human Dicer-dsRBD in processing small regulatory RNAs. *PloS One*, *7*(12), e51829. https://doi.org/10.1371/journal.pone.0051829

Wu, C., So, J., Davis-Dusenbery, B. N., Qi, H. H., Bloch, D. B., Shi, Y., Lagna, G., & Hata, A. (2011). Hypoxia potentiates microRNA-mediated gene silencing through posttranslational modification of Argonaute2. *Molecular and Cellular Biology*, *31*(23), 4760–4774. https://doi.org/10.1128/MCB.05776-11

Wu, H., Ye, C., Ramirez, D., & Manjunath, N. (2009). Alternative processing of primary microRNA transcripts by Drosha generates 5′ end variation of mature microRNA. *PloS One*, *4*(10), e7566. https://doi.org/10.1371/journal.pone.0007566

Wu, T.-H., Shi, L., Adrian, J., Shi, M., Nair, R. V., Snyder, M. P., & Kao, P. N. (2018). NF90/ILF3 is a transcription factor that promotes proliferation over differentiation by hierarchical regulation in K562 erythroleukemia cells. *PloS One*, *13*(3), e0193126. https://doi.org/10.1371/journal.pone.0193126

Xie, M., Li, M., Vilborg, A., Lee, N., Shu, M.-D., Yartseva, V., Šestan, N., & Steitz, J. A. (2013). Mammalian 5′-capped microRNA precursors that generate a single microRNA. *Cell*, *155*(7), 1568–1580. https://doi.org/10.1016/j.cell.2013.11.027

Xu, Y., Xia, F., Ma, L., Shan, J., Shen, J., Yang, Z., Liu, J., Cui, Y., Bian, X., Bie, P., & Qian, C. (2011). MicroRNA-122 sensitizes HCC cancer cells to adriamycin and vincristine through modulating expression of MDR and inducing cell cycle arrest. *Cancer Letters*, *310*(2), 160–169. https://doi.org/10.1016/j.canlet.2011.06.027

Yan, C., Wan, R., & Shi, Y. (2019). Molecular Mechanisms of pre-mRNA Splicing through Structural Biology of the Spliceosome. *Cold Spring Harbor Perspectives in Biology*, *11*(1). https://doi.org/10.1101/cshperspect.a032409

Yan, K. S., Yan, S., Farooq, A., Han, A., Zeng, L., & Zhou, M.-M. (2003). Structure and conserved RNA binding of the PAZ domain. *Nature*, *426*(6965), 468–474. https://doi.org/10.1038/nature02129

Yan, S., & Jiao, K. (2016). Functions of miRNAs during Mammalian Heart Development. *International Journal of Molecular Sciences*, *17*(5). https://doi.org/10.3390/ijms17050789

Yang, F., Huo, X.-s., Yuan, S.-x., Zhang, L., Zhou, W.-p., Wang, F., & Sun, S.-h. (2013). Repression of the Long Noncoding RNA-LET by Histone Deacetylase 3 Contributes to Hypoxia-Mediated Metastasis. *Molecular Cell*, *49*(6), 1083–1096. https://doi.org/10.1016/j.molcel.2013.01.010

Yang, Q., Li, W., She, H., Dou, J., Duong, D. M., Du, Y., Yang, S.-H., Seyfried, N. T., Fu, H., Gao, G., & Mao, Z. (2015). Stress induces p38 MAPK-mediated phosphorylation and inhibition of Drosha-dependent cell survival. *Molecular Cell*, *57*(4), 721–734. https://doi.org/10.1016/j.molcel.2015.01.004

Yang, Q., & Jankowsky, E. (2006). The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nature Structural & Molecular Biology*, *13*(11), 981–986. https://doi.org/10.1038/nsmb1165

Yang, W., Chendrimada, T. P., Wang, Q., Higuchi, M., Seeburg, P. H., Shiekhattar, R., & Nishikura, K. (2006). Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nature Structural & Molecular Biology*, *13*(1), 13–21. https://doi.org/10.1038/nsmb1041

Yang, X.-Q., Zhang, C.-X., Wang, J.-K., Wang, L., Du, X., Song, Y.-F., & Liu, D. (2019). Transcriptional regulation of the porcine miR-17-92 cluster. *Molecular genetics and genomics: MGG*, *294*(4), 1023–1036. https://doi.org/10.1007/s00438-019-01560-0

Yao, B., La, L. B., Chen, Y.-C., Chang, L.-J., & Chan, E. K. L. (2012). Defining a new role of GW182 in maintaining miRNA stability. *EMBO reports*, *13*(12), 1102–1108. https://doi.org/10.1038/embor.2012.160

Ye, J., Jin, H., Pankov, A., Song, J. S., & Blelloch, R. (2017). NF45 and NF90/NF110 coordinately regulate ESC pluripotency and differentiation. *RNA*, *23*(8), 1270–1284. https://doi.org/10.1261/rna.061499.117

Yen, W.-H., Ke, W.-S., Hung, J.-J., Chen, T.-M., Chen, J.-S., & Sun, H. S. (2016). Sp1-mediated ectopic expression of T-cell lymphoma invasion and metastasis 2 in hepatocellular carcinoma. *Cancer Medicine*, *5*(3), 465–477. https://doi.org/10.1002/cam4.611

Yeom, K.-H., Lee, Y., Han, J., Suh, M. R., & Kim, V. N. (2006). Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Research*, *34*(16), 4622–4629. https://doi.org/10.1093/nar/gkl458

Yi, R., Doehle, B. P., Qin, Y., Macara, I. G., & Cullen, B. R. (2005). Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA (New York, N.Y.)*, *11*(2), 220–226. https://doi.org/10.1261/rna.7233305

Yi, R., Qin, Y., Macara, I. G., & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development*, *17*(24), 3011–3016. https://doi.org/10.1101/gad.1158803

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C.-C. G., Tolia, N. H., Joshua-Tor, L., Mitani, S., Simard, M. J., & Mello, C. C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell*, *127*(4), 747–757. https://doi.org/10.1016/j.cell.2006.09.033

Yin, S., Yu, Y., & Reed, R. (2015). Primary microRNA processing is functionally coupled to RNAP II transcription in vitro. *Scientific Reports*, *5*, 11992. https://doi.org/10.1038/srep11992

Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., & Tomari, Y. (2010). ATP-dependent human RISC assembly pathways. *Nature Structural & Molecular Biology*, *17*(1), 17–23. https://doi.org/10.1038/nsmb.1733

Yoon, J.-H., Jo, M. H., White, E. J. F., De, S., Hafner, M., Zucconi, B. E., Abdelmohsen, K., Martindale, J. L., Yang, X., Wood, W. H., Shin, Y. M., Song, J.-J., Tuschl, T., Becker, K. G., Wilson, G. M., Hohng, S., & Gorospe, M. (2015). AUF1 promotes let-7b loading on Argonaute 2. *Genes & Development*, *29*(15), 1599–1604. https://doi.org/10.1101/gad.263749.115

Yuan, Z., Ding, S., Yan, M., Zhu, X., Liu, L., Tan, S., Jin, Y., Sun, Y., Li, Y., & Huang, T. (2015). Variability of miRNA expression during the differentiation of human embryonic stem cells into retinal pigment epithelial cells. *Gene*, *569*(2), 239–249. https://doi.org/10.1016/j.gene.2015.05.060

Zdanowicz, A., Thermann, R., Kowalska, J., Jemielity, J., Duncan, K., Preiss, T., Darzynkiewicz, E., & Hentze, M. W. (2009). Drosophila miR2 primarily targets the m7GpppN cap structure for translational repression. *Molecular Cell*, *35*(6), 881–888. https://doi.org/10.1016/j.molcel.2009.09.009

Zeng, Y., & Cullen, B. R. (2004). Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Research*, *32*(16), 4776–4785. https://doi.org/10.1093/nar/gkh824

Zhang, C., Xie, C., Wang, X., Huang, Y., Gao, S., Lu, J., Lu, Y., & Zhang, S. (2020). Aberrant USP11 expression regulates NF90 to promote proliferation and metastasis in hepatocellular carcinoma. *American Journal of Cancer Research*, *10*(5), 1416–1428.

Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E., & Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell*, *118*(1), 57–68. https://doi.org/10.1016/j.cell.2004.06.017

Zhang, W., Xiong, Z., Wei, T., Li, Q., Tan, Y., Ling, L., & Feng, X. (2018). Nuclear factor 90 promotes angiogenesis by regulating HIF-1/VEGF-A expression through the PI3K/Akt signaling pathway in human cervical cancer. *Cell Death & Disease*, *9*(3), 276. https://doi.org/10.1038/s41419-018-0334-2

Zhang, X., Xu, Y., Qian, Z., Zheng, W., Wu, Q., Chen, Y., Zhu, G., Liu, Y., Bian, Z., Xu, W., Zhang, Y., Sun, F., Pan, Q., Wang, J., Du, L., & Yu, Y. (2018). circRNA_104075 stimulates YAP-dependent tumorigenesis through the regulation of HNF4a and may serve as a diagnostic marker in hepatocellular carcinoma. *Cell Death & Disease*, *9*(11). https://doi.org/10.1038/s41419-018-1132-6

Zhang, Y., Hu, Y., Fang, J.-Y., & Xu, J. (2016). Gain-of-function miRNA signature by mutant p53 associates with poor cancer outcome. *Oncotarget*, *7*(10), 11056–11066. https://doi.org/10.18632/oncotarget.7090

Zhang, Y., Sun, J., Qi, Y., Wang, Y., Ding, Y., Wang, K., Zhou, Q., Wang, J., Ma, F., Zhang, J., & Guo, B. (2020). Long non-coding RNA TPT1-AS1 promotes angiogenesis and metastasis of colorectal cancer through TPT1-AS1/NF90/VEGFA signaling pathway. *Aging*, *12*(7), 6191–6205. https://doi.org/10.18632/aging.103016

Zheng, B., Zhou, J., & Wang, H. (2020). Host microRNAs and exosomes that modulate influenza virus infection. *Virus Research*, *279*, 197885. https://doi.org/10.1016/j.virusres.2020.197885

Zhu, L., Kandasamy, S. K., & Fukunaga, R. (2018). Dicer partner protein tunes the length of miRNAs using base-mismatch in the pre-miRNA stem. *Nucleic Acids Research*, *46*(7), 3726–3741. https://doi.org/10.1093/nar/gky043

Zhuang, J., Shen, L., Yang, L., Huang, X., Lu, Q., Cui, Y., Zheng, X., Zhao, X., Zhang, D., Huang, R., Guo, H., & Yan, J. (2017). TGF1 Promotes Gemcitabine Resistance through Regulating the LncRNA-LET/NF90/miR-145 Signaling Axis in Bladder Cancer. *Theranostics*, *7*(12), 3053–3067. https://doi.org/10.7150/thno.19542