

**Investigating nucleolar dynamics in karyotypically normal human cells by fluorescent protein tagging of genes encoding key nucleolar proteins.**



**NUI Galway**  
**O'É Gaillimh**

**A thesis submitted to the National University of Ireland as  
fulfilment of the requirement for the degree of**

**Doctor of Philosophy**

**by**

**Sebastian Johannes Fladerer**

**Date of Submission: December 2021**

**Centre for Chromosome Biology,  
School of Natural Sciences,  
National University of Ireland Galway**

**Thesis Supervisor:**

**Professor Brian McStay**



## Table of contents

Chapter I .....	1
1. Introduction.....	2
1.1 A historical overview about nucleolar structure and function.....	2
1.2 The architecture of ribosomal DNA (rDNA) .....	7
1.2.1 rDNA in prokaryotes.....	7
1.2.2 rDNA in yeast a simple eukaryote.....	7
1.2.3 rDNA in mammals.....	9
1.3 rDNA transcription mediated by RNA Polymerase I .....	14
1.3.1 RNA Polymerase I.....	14
1.3.3 Pol-1 $\beta$ Transcription Initiation.....	17
1.3.4 Elongation.....	17
1.3.5 Transcription Termination .....	17
1.3.6 Re-initiation.....	18
1.3.7 The role of UBF in rDNA transcription.....	18
1.3.8 Processing and ribosome biogenesis .....	21
1.4 NORs and the Nucleolus .....	23
1.4.1 The nucleolar structure.....	23
1.4.2 Nucleoli in different species.....	26
1.4.3 The Nucleolus during the Cell Cycle .....	27
1.4.4 Liquid-Like phase separation as a model for the dynamic behaviour of nucleoli?.....	30
1.5 Aims of this thesis. ....	36
Chapter II .....	37
2. Materials and Methods.....	38
2.1 DNA techniques .....	38
2.1.1 Bacterial transformation .....	38
2.1.2 Plasmid Purification .....	38
2.1.3 Restriction Digest.....	39
2.1.4 DNA Ligation.....	39
2.1.5 Gibson-Assembly .....	39
2.1.6 Agarose gel electrophoresis .....	39
2.1.7 DNA gel extraction and PCR purification .....	40
2.1.8 Gateway Cloning .....	40
2.1.9 Preparation of Genomic DNA from human cells.....	40
2.1.10 Polymerase Chain Reaction .....	41

2.1.11 Primers .....	41
2.1.12 Vectors .....	43
2.1.13 DNA sequencing .....	44
2.2. Protein techniques .....	44
2.2.1 Western Blot .....	44
2.2.2 Antibodies for Western-Blot's .....	45
2.3 Cell culture .....	45
2.3.1 Growth conditions and splitting .....	45
2.3.2 Liquid nitrogen stocks: freezing down and thawing out .....	46
2.4 CRISPR/Cas9 genome editing .....	46
2.4.1 RNP preparation .....	46
2.4.2 Transfection .....	46
2.5 Imaging .....	47
2.5.1 Imaging/live cell imaging .....	47
2.5.2 Fixed cells .....	47
2.5.3 Live cell imaging and Halo staining .....	48
<b>CHAPTER III .....</b>	<b>49</b>
<b>3. Establishing a strategy for C-terminal tagging of hUBF by CRISPR/Cas9 ....</b>	<b>50</b>
3.1. Crispr/Cas9 as a tool for FP tagging relevant nucleolar key proteins in hTERT-RPE1 cells .....	50
3.2 Aims of this chapter .....	55
3.3 Results .....	55
3.3.1 Developing a Crispr/Cas9 Knock-In technique for hTERT-RPE1 cells. ....	55
3.3.2 Developing a flexible tagging system .....	61
3.4 FP tagging of tagging of Nucleolar bookmarker proteins by CRISPR/Cas9 .....	64
3.5 Endogenous hUBF tagging with monomeric Azami Green (mAG) .....	64
<b>Chapter IV .....</b>	<b>70</b>
<b>4. Tagging of nucleolar markers for the DFC and GC compartment .....</b>	<b>71</b>
4.1 A wider perspective for tagging other marker proteins .....	71
4.2 Tagging of Fibrillarin as a bookmarker for the DFC compartment .....	71
4.2.1 A little background about the function of FIB .....	71
4.2.2 Results for tagging Fibrillarin with mCherry .....	72
4.3 Tagging of a suitable GC bookmarker protein .....	77
4.4. Tagging of Nop52 with HALO .....	77

4.4.1 Some background information of Nop52 .....	77
4.4.2 Results for tagging Nop52 with HALO.....	77
4.3 Discussion and future perspective.....	86
<b>Chapter V.....</b>	<b>88</b>
<b>5. The dynamic behaviour of the nucleolus .....</b>	<b>89</b>
5.1 Different aspects about the nucleolus .....	89
5.2 Results .....	94
5.2.1 The <i>in vivo</i> perspective on the nucleolus and its sub-compartments	94
5.2.2 Nucleolar formation from the FC/DFC perspective by live-cell imaging. .....	97
5.2.3 Nucleolar formation from the FC/GC perspective. ....	102
5.2.4 Nop52 and its dynamic behaviour during the cell cycle. ....	107
5.3 Nucleolar stress response of the FC/DFC and FC/GC compartments ...	111
5.4 Discussion and future perspective.....	117
<b>Chapter VI.....</b>	<b>121</b>
<b>6. Final Discussion .....</b>	<b>122</b>
6.1 The development and application of the CRISPR/Cas9 system in hTERT- RPE1 cells, a non-transformed cell line .....	122
6.2 Live Cell imaging short-term priorities .....	129
6.3 Measuring the biophysical properties of nucleolar compartments during stress .....	133
6.4 Final sum up.....	135
<b>References: .....</b>	<b>137</b>

**Declaration:**

I, Sebastian Johannes Fladerer, certify that the Thesis is all my own work and that I have not obtained a degree in this University or elsewhere on the basis of any of this work.

## **Abstract:**

The nucleolus, the site of ribosome biogenesis, is the largest membrane-less structure or organelle in the eukaryotic nucleus. It forms around arrays of ribosomal gene (rDNA) repeats termed nucleolar organizer regions (NORs). In humans, NORs are located on short (p) arms of the five acrocentric chromosomes, and are transcribed by RNA polymerase I to generate pre-ribosomal RNA (pre-rRNA) that are subsequently processed and assembled into ribosomes.

Nucleoli comprise a fibrillar component (FC), a dense fibrillar component (DFC) and a granular component (GC) as compartments that reflect the sequential steps of ribosome biogenesis, rDNA transcription, pre-rRNA processing and pre-ribosome assembly respectively. RNA Polymerase I (Pol I) rDNA transcription, is the first step for ribosomal assembly, and it is evolutionarily conserved and is generally considered to be one of the most energy consuming process required for cell growth, proliferation, and homeostasis. Sequences surrounding NORs on acrocentric p-arms contribute to the formation of peri-nucleolar heterochromatin, observed in most human cells.

In recent years, liquid like behaviour and liquid-liquid phase separation (LLPS) have been proposed to explain both nucleolar formation and internal organisation. A further refined model of nucleoli as comprising multiphase condensates has been derived from *in vitro* experiments with only a few selected recombinant fusion proteins and *in vivo* experiments with over expressed fusion proteins in transformed cell lines.

In this thesis I develop a strategy for efficient fluorescent protein and Halo-tagging of multiple genes encoding key nucleolar components. The cell model I use, hTert-RPE1, is karyotypically normal, has well characterised NORs and intact cell-cycle check points. Like other non-transformed human cell lines, hTert-RPE1 were previously shown to be highly refractile to CRISPR gene tagging. This is due to low rates of homology directed repair of CAS9 induced DNA double strand breaks<sup>1,2</sup>. The key advance is to link targeted integration with selection through the use of viral P2A sequences.

The cells lines I have developed provide a platform for future exploration of nucleolar dynamics in live 'normal' human cells. Exploitation of these cell models during the later stages of my thesis work was profoundly compromised by the emerging COVID-19 pandemic. Nevertheless, preliminary experiments confirm their utility for investigating

nucleolar breakdown and reformation before and after cell division respectively. Additionally, they will provide powerful tools for studying the dramatic reorganisation associated with nucleolar stress. Finally, the strategies and arrays of targeting cassettes I have developed will power future research into other aspects of this critical and most intriguing of nuclear organelles.

## **List of Figures.**

Fig. 1.1 First historical illustration of the nucleolus

Fig. 1.2 Miller-Spreads.

Fig. 1.3 Structural features of the human nucleolus on microscopy images

Fig. 1.4 The nucleolar proteome

Fig. 1.5 Schematic architecture of a prokaryotic rDNA repeat

Fig. 1.6 Structure and genomic organization of rDNA repeats in *Saccharomyces cerevisiae*

Fig. 1.7 A closer look on Yeast rDNA units.

Fig. 1.8 The schematic architecture of a single rDNA repeat.

Fig. 1.9 Localisation of the rDNA arrays on human acrocentric chromosomes

Fig. 1.10: A closer illustration of the human rDNA repeat unit, including an IGS.

Fig. 1.11: Comparison of the distribution of regulatory elements in the rDNA repeat of human and mouse.

Fig. 1.12 Schematic diagram illustrating the arrangement of the PIC in an rDNA array context

Fig. 1.13 Illustration of sequence conservation of UBF variants.

Fig. 1.14 Alternate Processing pathways of pre-rRNA transcript in human

Fig. 1.15 Organization of Nucleolar Organizer Regions (NORs) in Nucleoli

Fig. 1.16 Different perspectives on nucleoli and their internal organization

Fig. 1.17 The illustration of the nucleolar cell cycle by a drawing

Fig. 1.18 A role for Ki67 in Maintaining the chromosomal periphery compartment

Fig. 1.19 Liquid-like Properties of nucleoli during nucleolar fusion

Fig. 1.20 Internal liquid-like organisation of Nucleoli

Fig. 1.21 Nucleolar cap formation as a stress response

Fig. 1.22 location of Pseudo-NORs and their droplet-like structure in the context of the LLPS model.

Fig. 3.1 An illustration of the CRISPR/Cas9 complex and its interaction with substrate DNA

Fig. 3.2 Illustration of the traffic light reporter assay.

Fig. 3.3 Schematic illustration of Crispr/Cas9 genome editing, based on plasmid encoded constructs

Fig. 3.4 The IDT based 2-part gRNA format.

Fig. 3.5 IRES and Ribosomal Skipping as genetic tools

Fig. 3.6 Genome editing principle of N-terminal and C-terminal tagging based on developed Knock-In constructs,

Fig. 3.7 Schematic cloning strategy for creating different Knock-In cassettes

Fig. 3.8 Illustration of possible cassettes which can be used for CRISPR/Cas9 mediated Knock-In's

Fig. 3.9 Schematic illustration of the gene tagging procedure by single steps

Fig. 3.10 DNA sequence at the C-terminus of hUBF and targeting positions of two gRNAs (9 and 22)

Fig. 3.11 Illustration of the cloning strategy to create a UBF repair template, based on the original plasmid created by Dr. Chelly van Vuuren.

Fig. 3.12 gRNA and sequence design for the UBF tagging strategy

Fig. 3.13 A Different verification steps for the UBF-mAG Knock-In.

Fig 4.1 Cloning strategy for the FIB repair template

Fig. 4.2 gRNA and sequence design for the FIB tagging strategy

Fig. 4.3 Different verification steps for the UBF-mAG/FIB-mCherry double tagged cell line

Fig. 4.5 Map of the C-terminal Halo tagged Nop52 expression construct.

Fig. 4.6 The transient expression of a Nop52-Halo construct

Fig. 4.7 Cloning strategy for the Nop52 repair templates

Fig. 4.8 gRNA and sequence design for the Nop52 tagging strategy.

Fig. 4.9 Different verification steps for the Nop52-Halo tagged cell line

Fig. 4.10 Different verification steps for the UBF-mAG/Nop52-Halo double tagged cell line

Fig. 5.1 A closer look on the nucleolus and its sub compartments.

Fig. 5.2 Nucleolar cap formation as part of the stress response.

Fig. 5.3 A simple illustration of the cell cycle.

Fig. 5.4 UBF behaviour during meta- and anaphase

Fig. 5.5 Image of fixed and DAPI stained UBFmAG-FIBmCh-hTERT-RPE1 knock in cells.

Fig. 5.6 Illustration of the formation event of the FC and DFC nucleolar sub-compartments in hTERT-Rpe1-UBFmAG-FibmCherry Knock-In cell line.

Fig. 5.7 Enlargement of the formation event (from Fig. 5.6).

Fig. 5.8 Image of fixed and DAPI stained UBFmAG-Nop52Halo-hTERT-RPE1 knock in cells.

Fig. 5.9 Illustration of the formation event of the FC and GC nucleolar sub-compartments in hTERT-Rpe1-UBFmAG-Nop52Halo Knock-In cells.

Fig. 5.10 Closer look on the formation event shown in the figure above (Fig. 5.9).

Fig. 5.11 A closer look on nucleolar disruption

Fig. 5.12 Homozygous Nop52Halo-hTERT-RPE1HKO cell Image.

Fig. 5.13 A closer look on Nop52 and nucleolar formation.

Fig. 5.14 A closer look on Nop52 and nucleolar dynamics of a single interphase cell.

Fig. 5.15 Model for nucleolar organization in normal growth conditions and under nucleolar stress reorganization.

Fig. 5.16 Cap formation in of four hTERT-Rpe1-UBFmAG-FIBmCherry cells.

Fig. 5.17 Enlargement of the nucleolar stress response in a single cell under AMD treatment for UBFmAG-FIBmCherry tagged cells, marking the FC/DFC compartments.

Fig. 5.18 Cap formation in three different hTERT-Rpe1-UBFmAG-Nop52Halo cells. Time points for every image are indicated in white.

Fig 5.19 Enlargement of the nucleolar stress response in a single cell under AMD treatment for UBFmAG-Nop52Halo tagged cells, marking the FC/GC compartments

Fig. 6.1 Developed cloning strategy to generate repair constructs for C-terminal CRISPR/Cas9 tagging.

Fig. 6.2 The principal workflow of generating different hTERT-Rpe1 tagged cell lines, based on the developed method.

Fig. 6.3 General overview about the generated cell lines.

Fig. 6.4 Schematic procedure for in vitro experiments based on isolated nucleoli.

Fig. 6.5 The BODIPY-dye for the Halo-tag system.

Fig.6.6 Design and concept of the crowding sensor

## **List of Tables**

Table 2.1 Primers for FP-P2A cassettes

Table 2.2 Primers for FP-P2A-selection markers

Table 2.3 Primers and gRNA's for UBF Knock-In

Table 2.4 Primers and gRNA for FIB Knock-In

Table 2.5 Primers and gRNA Nop52 Knock-In

Table 2.6 Cloning vectors

Table 2.7 Antibody list for Western-Blot's

## Abbreviations:

AMD	Actinomycin D
BLAST	Basic Local Alignment Search Tool
bp	base pair
BODIPY	boron-diphyrrin
BSA	Bovine Serum Albumin
CF	Core Factor
ChIP	Chromatin Immunoprecipitation
CMV	Cytomegalovirus
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Da	Dalton
DAPI	4'-6-diamidino-2-phenylindole
DFC	dense fibrillar component
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
ds	double-stranded
FBS	foetal bovine serum
FC	fibrillar centre
FISH	fluorescence in situ hybridisation
GC	granular component
GFP	green fluorescent protein
HAT	histone acetyltransferase
HMG	high mobility group
H	hour
HALO	haloalkane dehalogenase
HRP	horseradish peroxidase
IgG	Immunoglobulin G
IGS	intergenic spacer
kb	kilo-base
LB	Luria broth

MEM	minimum essential medium
min	minute
ml	milliliter
nm	nanometer
NOR	nucleolar organizer region
O/N	overnight
orf	open reading frame
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
PIC	Pre-Initiation Complex
RNA	ribonucleic acid
RP	ribosomal protein
RPE	retinal pigment epithelial
rpm	revolutions per minute
RT	room temperature
SDS	sodium dodecyl sulfate
sec	second
shRNA	short hairpin RNA
SL1	selectivity factor 1
SLD	SANT-like domain
SNAP	SILAC nucleosome affinity purification
snoRNA	small nucleolar RNA
snoRNP	small nucleolar ribonucleoprotein
snRNA	small nuclear RNA
SSU	small sub-unit
TALEN	Transcription activator-like effector nuclease
TBP	TATA-binding protein
TEM	transmission electron microscopy
TERT	telomerase reverse transcriptase
TTF-1	transcriptional termination factor 1
UAF	upstream activating factor
UBF	upstream binding factor

UCE

upstream control element

UV

Ultraviolet

V

volts

v/v

volume to volume

w/v

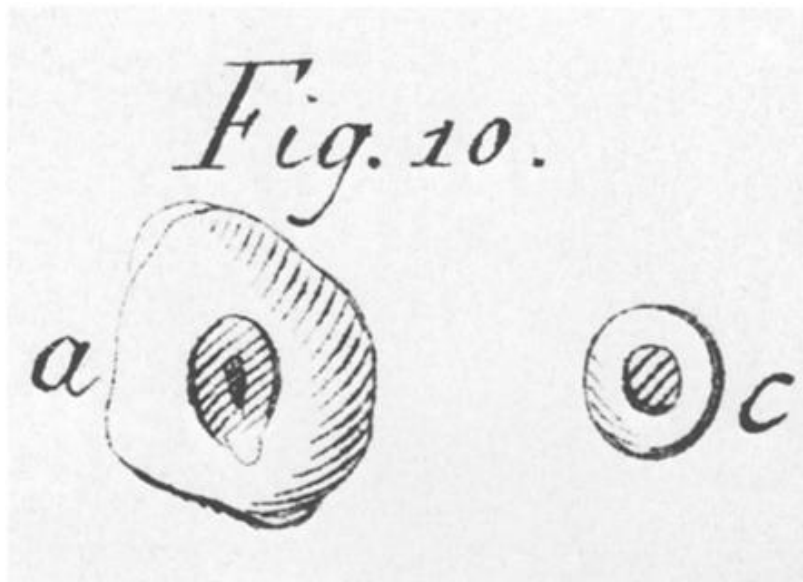
weight to volume

# Chapter I

## 1. Introduction

### 1.1 A historical overview about nucleolar structure and function

The nucleolus is the most prominent nuclear (sub-)structure in all eukaryotic organisms and is readily observed using phase contrast microscopy. The main function of the nucleolus is the synthesis of ribosomal RNA, however it also plays a role in other cellular processes, such as cell cycle control, stress response and cellular ageing<sup>3,4</sup>. The nucleolus was first observed by Fontana (1781), but it took another 40 years until Wagner (1835) and Valentin (1836) discovered and originally described the nucleolus independently.



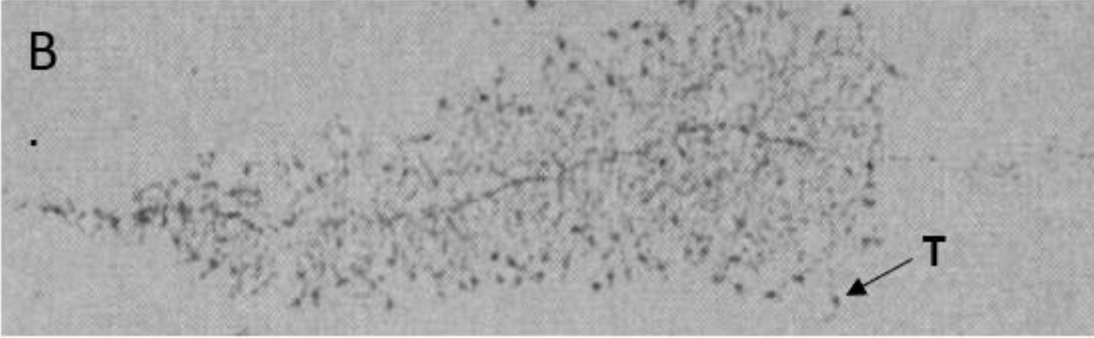
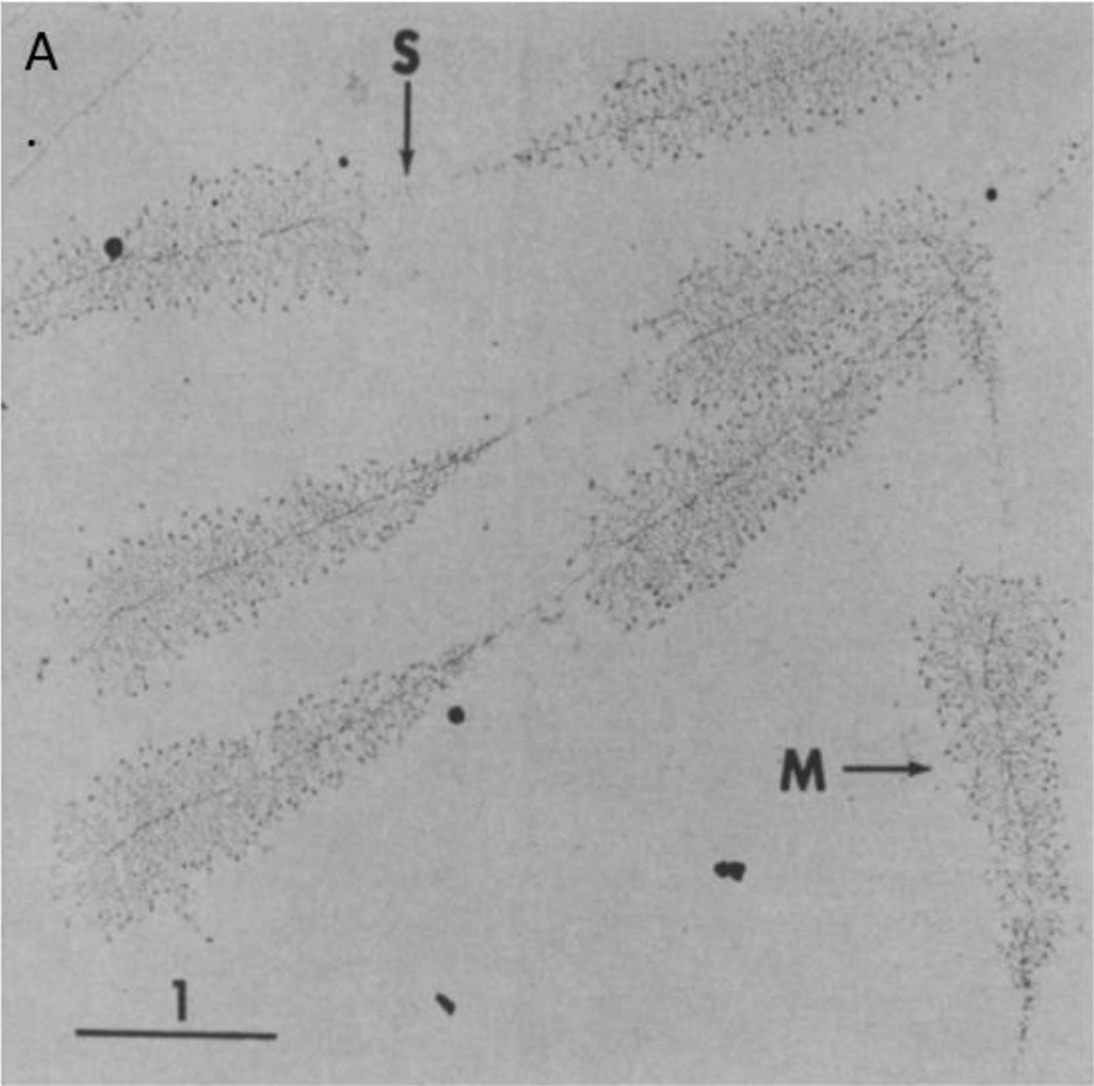
**Fig. 1.1:** First historical illustration of the nucleolus. a.) Epidermal cell of an eel with nucleus and nucleolus. c.) Red blood cell of the same species as scale. Figure adapted from<sup>5</sup> and originally published by Fontana (1781)

The first documented reports by Rudolph Wager in 1835, initially described them as a “fleck” (or “spot” in English), and later as “the germinative spot”. The term “nucleolus” was first used by Gustav Gabriel Valentin in 1836. In 1898 the first comprehensive review about the discoveries concerning nucleoli in the 18<sup>th</sup> and 19<sup>th</sup> century was published and it included 346 hand-drawn nuclei and nucleoli from a collection of

biological material.<sup>6</sup> Thirty years later Heitz (1931) and McClintock (1934) independently observed that nucleoli form at specific chromosomal loci, in *Zea mays*. Heitz further noted, that nucleoli form at the secondary constriction, in contrast to primary constrictions which are the centromeres, and that the size of this secondary constrictions is reflected by the size of the respective nucleolus. In addition, he also observed that nucleoli tend to fuse during the cell cycle, as nuclei contain only one nucleolus before cells entered mitosis. Furthermore, McClintock also initially described the nucleolar-organizing body, today known as nucleolar organiser region (NOR). This was the first time when the nucleolus was connected to a cytogenic entity. McClintock also hypothesised that the information held in NORs was redundant in nature as it could be split by chromosomal translocation resulting in the appearance of two nucleoli.

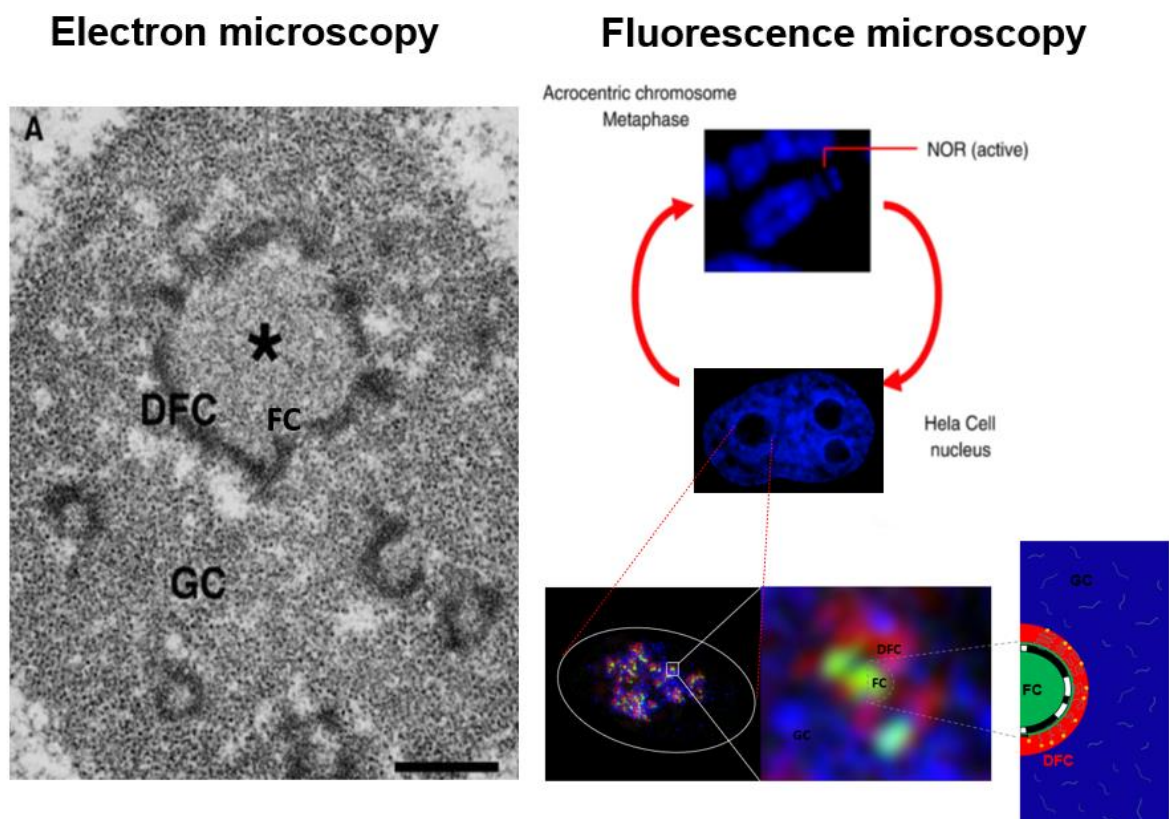
In the 50s and 60s of the last century, protocols were developed to isolate nucleoli and analyse their molecular composition.<sup>7-9</sup> In these approaches nucleoli were purified by sequential cell fractions and centrifugation. These protocols were then adapted in recent years to get deeper insights concerning molecular constitution and dynamics of nucleoli. Later, various studies led to the conclusion that a major component of the nucleolus is ribosomal RNA. Initial evidence came from studies of a *Xenopus* mutant lacking nucleoli due to a deletion of the NORs. In these mutants the tadpoles had a deficiency in ribosomal RNA synthesis (rRNA).<sup>10,11</sup> Conclusive evidence came then from experiments with radioactively labelled rRNA. When the rRNA was hybridised to metaphase chromosomes it localized to the secondary constrictions on acrocentric chromosomes, and when it was hybridized in interphase cells it preferentially localized to nucleoli.<sup>12-15</sup> The experiments demonstrated that rRNA is synthesized in nucleoli and encoded by ribosomal genes (rDNA array units), which are located at the NORs. Next, the synthesis of this RNA by enzymatic transcription was shown.<sup>16,17</sup> In 1969 O.L. Miller captured rDNA during transcription and showed heavy loading of the rDNA with Pol I, which was the first in situ visualization of a transcribed gene. The electron micrographs became then famous under the term “Miller spreads” or “Christmas trees.”<sup>18</sup> Multiple transcription units are lined up one after the other on a single rDNA unit and synthesize rRNA. The single nascent transcripts of the transcription units, which are active on a single rDNA array have different length, like branches of a fir tree (“Christmas tree”) (see Fig. 2A).

At the 5' end of every nascent transcript, so-called terminal ball structure can be observed (Fig. 2 B.). These structures are the ribosomal small subunit processome (SSU-processome) and will be discussed in more detail later. Between the single active transcribed rDNA units, the intergenic spacer (IGS) is localized.



**Fig. 1.2.** Miller-Spreads. **A.** Original Miller spread from nucleolar core, isolated from *Triturus viridescens* oocytes, the spread can be subdivided in two segments, a Matrix segment M and a matrix-free segment S. Scalebar 1  $\mu\text{m}$ . **B.** Enlarged segment of a Miller spread, which shows the typical “Christmas tree” form, terminal balls (T) are at the beginning of every new transcribed rRNA molecule. Figure adapted from <sup>18</sup>

However, electron microscopic experiments also opened a new perspective on the substructure of nucleoli. The first electron microscopic experiments shed light on the tripartite structure of the nucleolus in 1952 <sup>19</sup>, revealing that the nucleolus consists of the fibrillar centre (FC), the dense fibrillar component (DFC) and the granular component (GC).

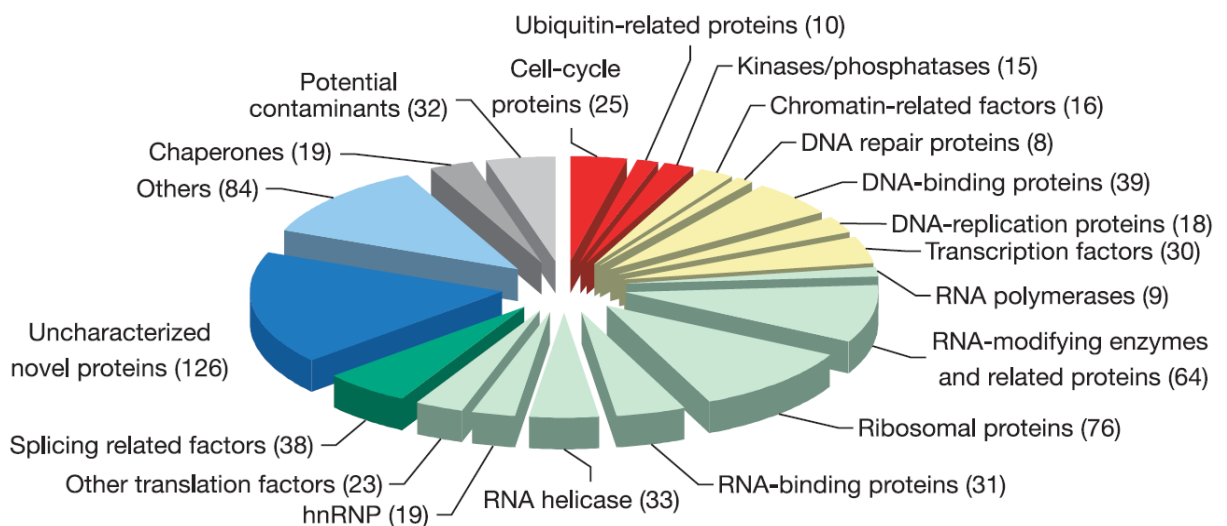


**Fig. 1.3.** Structural features of the human nucleolus on microscopy images. Electron microscopy image illustrating the three main parts of the nucleolus of HeLa cell (on the left). The asterisk marks the fibrillar centre, which is surrounded by the dense fibrillar component, which is enclosed by the granular component. Figure adapted from <sup>20</sup>. Fluorescence microscope images are shown on the right. Upper panel shows a DAPI stained human acrocentric chromosome with a NOR visible as secondary constriction (no DAPI signal). The middle panel depicts a DAPI stained HeLa cell nucleus. The density of DNA in the nucleolus is lower, therefore it appears as a black hole. Lower panes illustrate a nucleolus during the interphase with the three sub compartments FC, DFC and GC. Adapted from <sup>21</sup>. Scale bar 0.5  $\mu\text{m}$ .

After many years of contention, a consensus on the functions carried out in each compartment has been established.<sup>22</sup> The FC contains the ribosomal DNA and RNA polymerase I (Pol I), the transcription takes place at the boundary between the FC and DFC, the nascent RNA grows into the DFC where also the early processing of rRNA takes place, while later rRNA processing and ribosome assembly steps occur in the GC (See Fig. 1.3).<sup>23 24</sup>

Several factors, discovered in the 1980s, are essential for transcription initiation on the rDNA in human cells. A pre-initiation complex (PIC) formed on rDNA facilitates recruitment of RNA polymerase I. The PIC is comprised of Upstream Binding Factor (UBF) and Selectivity Factor 1 (SL-1).<sup>25,26</sup> When this complex is formed, a subfraction of Pol I that is associated with RRN3/TIF-1A is recruited to the promoter leading to transcription initiation.<sup>27,28</sup> Mechanistic details will be discussed later in this introduction.

The idea of a multifunctional nucleolus initially arose from quantitative analyses of the proteome of human nucleoli. In this study 692 proteins co purify with nucleoli have been identified. While most are involved directly in ribosome biogenesis a proportion are involved in cell cycle regulation, DNA damage repair and a significant fraction remain uncharacterized (see Fig. 1.4.).<sup>29</sup>



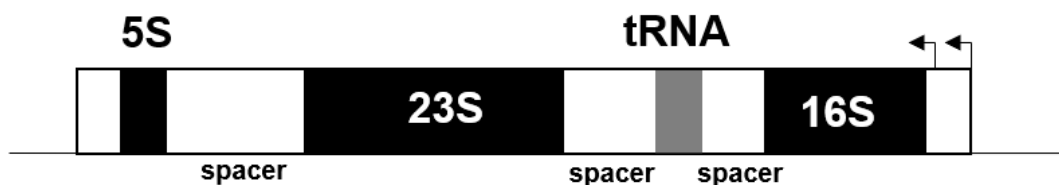
**Fig. 1.4:** The nucleolar proteome: Mass spectrometry revealed the presence of 692 proteins in the human nucleus. The identified proteins are clustered according to their molecular functions in a pie chart. Adapted from <sup>29</sup>.

## **1.2 The architecture of ribosomal DNA (rDNA)**

### **1.2.1 rDNA in prokaryotes**

The genome of a prokaryote like *E. coli* contains seven rRNA transcription units, these units are located at different positions. Every transcription unit has 2 promoters.<sup>30</sup> During transcription, these rRNA units gather together, forming what may be considered as a primitive nucleolus.<sup>31</sup>

The whole transcription unit, which is located downstream of the two promoters, is encoding genes for 16S, 23S and 5S ribosomal (RNA) subunits. The single rRNA subunits are separated by spacer regions. In the spacer region between the 16S and 23S subunits, a tRNA gene sequence is located (see Fig. 1.5).<sup>30,32</sup> However, it should be noted that this arrangement is a generalized composition for rRNA subunits. The number and location of the single 16S and 23S rRNA genes can vary.<sup>33</sup>

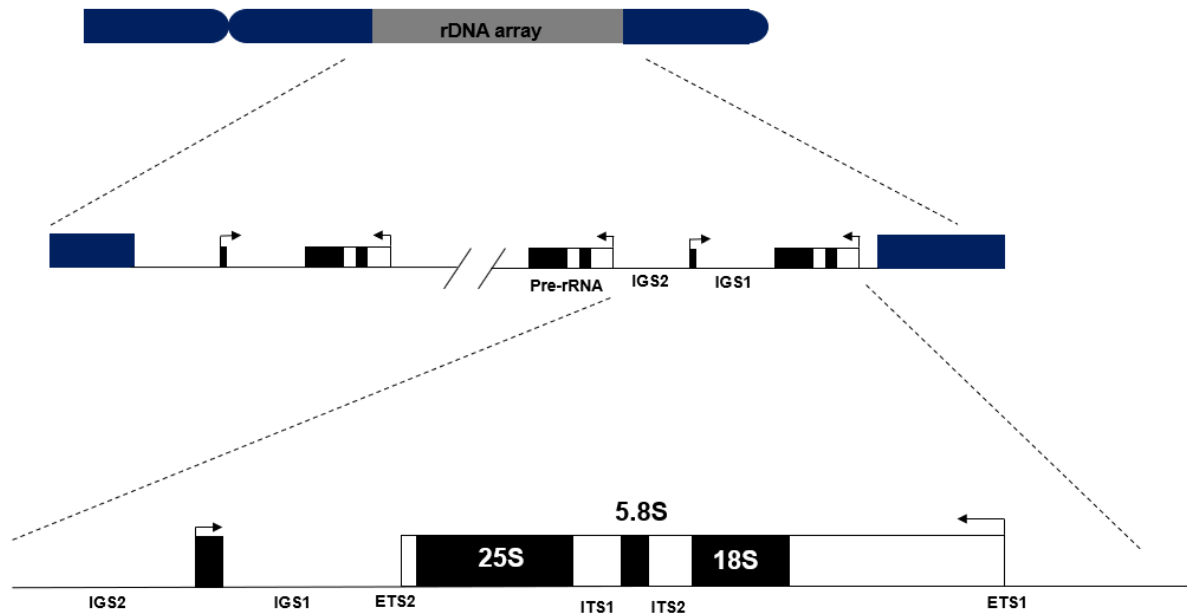


**Fig. 1.5:** Schematic architecture of a prokaryotic rDNA repeat

In many eubacteria, the genomic localisation of the rRNA transcriptional units is similar, but there is a difference in the genome of *B. subtilis*, in this organism the rRNA units are arranged in a cluster.<sup>34</sup>

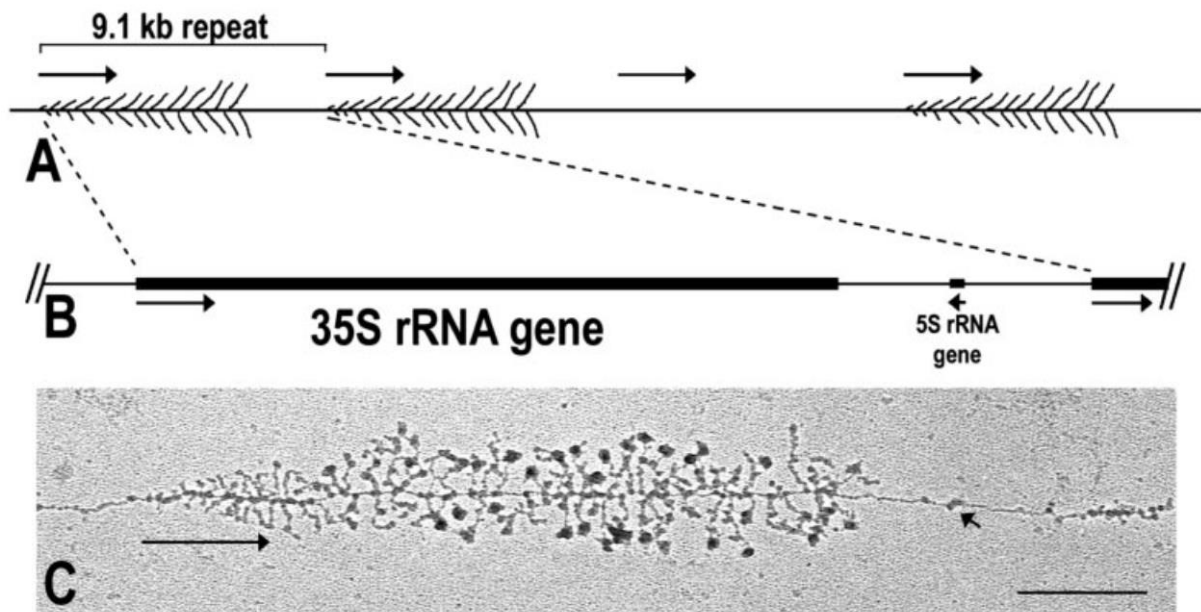
### **1.2.2 rDNA in yeast a simple eukaryote**

In haploid *Saccharomyces cerevisiae* cells only one nucleolus is formed, which contains 150 to 200 rRNA transcription units. Interestingly these genes are located in a single cluster of repeats organised in a head to tail tandem array on the right arm of chromosome 12.<sup>35,36</sup>



**Fig. 1.6:** Structure and genomic organization of rDNA repeats in *Saccharomyces cerevisiae*

Each rRNA transcription unit (size 35S), contains a 25S, a 5.8S and a 18S gene separated by internal transcription spacers (ITS) and flanked by external transcription spacers (ETS). Each 9.1 kb rDNA unit (enlarged section in Fig. 1.6.) has a Pol I enhancer element located 2.0 kb upstream of the transcription initiation site.<sup>37</sup> Each rDNA unit also contains a 5S gene, which is transcribed by RNA polymerase III from the opposite strand (Pol III) (see Fig. 1.6 and Fig. 1.7).<sup>38</sup> Miller spreads in budding yeast revealed, there are transcriptionally active and inactive/silent rDNA units present.



**Fig. 1.7** A closer look on Yeast rDNA units. **A.** schematic cartoon of four tandem rRNA gene repeats showing transcriptionally active and inactive/silent units. **B.** schematic figure of one gene spacer unit, including the Pol I transcribed 35S rRNA gene (long thick line) the 5S rRNA gene, transcribed by Pol III (short thick line) and non-transcribed spacers (thin lines) **C.** A representative electron micrograph of an active rRNA gene, also called “Miller Spread”, arrows are indicating direction of transcription, the SSU processome can be seen as little “Knobs” at the start of the fresh transcribed rRNA molecules. Figure adapted from <sup>39</sup>

Replication commences, while rDNA transcription is still ongoing. The replication can move into two directions whereas the transcription machineries can just move in one direction. To avoid collisions between the transcription and replication machineries, a replication fork barrier (RFB), is located in the IGS1 near the 3' end of the 35S rRNA gene.<sup>40</sup> The increased levels of unequal recombination between the sister chromatids, is also affecting the genomic stability.<sup>41</sup>

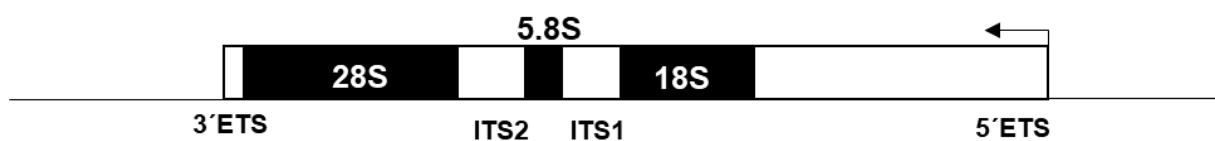
### **1.2.3 rDNA in mammals**

In mammals and in all other higher eukaryotes, the ribosomal genes for the 18S-28S and the 5S ribosomal subunits are getting transcribed by two different RNA Polymerases (I and III) and are located at different positions in the genome. In humans for example, the 5S rRNA gene arrays are located on chromosome 1 as a cluster of around 90 copies. These gene arrays are getting transcribed by RNA Pol III.<sup>42,43</sup>

In species where it has been determined, the 18S-28S rDNA genes are arranged as tandem arrays, in a head to tail orientation towards the centromere.<sup>44,45,46</sup> These rDNA clusters are the genetic foundation of the nucleolus and are called nucleolar organizer regions (NORs) for this reason.

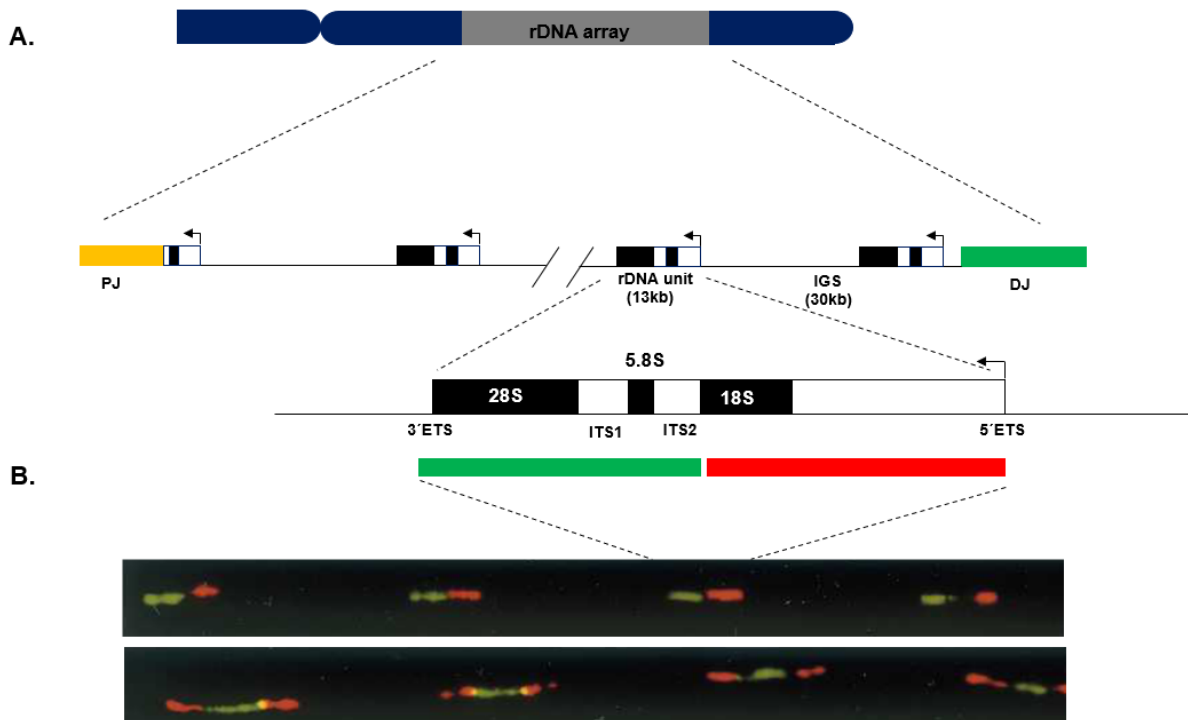
In mammalian cells the nucleolus forms around transcriptionally active NORs, which are located at the p-arms, amid sequences that form heterochromatin in interphase. In situ hybridization experiments unveiled, that those regions are located on the short arms of the human acrocentric chromosomes 13, 14, 15, 21 and 22.<sup>12</sup> Compared to yeast, mammalian rDNA repeat unit are large with roughly 43 kb in humans and about 45 kb in mice and consists of the rRNA gene and an intergenic spacer (IGS) of ~30k.<sup>47</sup>

The rRNA coding region is transcribed by RNA polymerase I, resulting in a 47S pre-rRNA precursor transcript, which is then processed into the 18S, 5.8S and 28S rRNA subunits needed for ribosome synthesis. The single ribosomal subunits are separated by internal transcript spacers (ITS1/2) and flanked by external transcript spacers at both ends (5' and 3'ETS).



**Fig. 1.8:** The schematic architecture of a single rDNA repeat.

Molecular combing experiments suggest that the human rDNA repeats are predominantly arranged in a head to tail fashion in humans, but a subset of repeats shows non-canonical arrangement (see Fig. 1.9).<sup>48</sup> The existence of non-canonical arranged rDNA repeats is still an open question. In the aspect of rDNA sequencing, there is no evidence that those arrangements do exist in karyotypic normal cells.<sup>49</sup> Thus, raising the possibility that non-canonical rDNA repeats are a recurring “DNA combing” artifact. Multiple rDNA strands can clump together and overlap because of the combing and deliver signals with a non-canonical orientation.



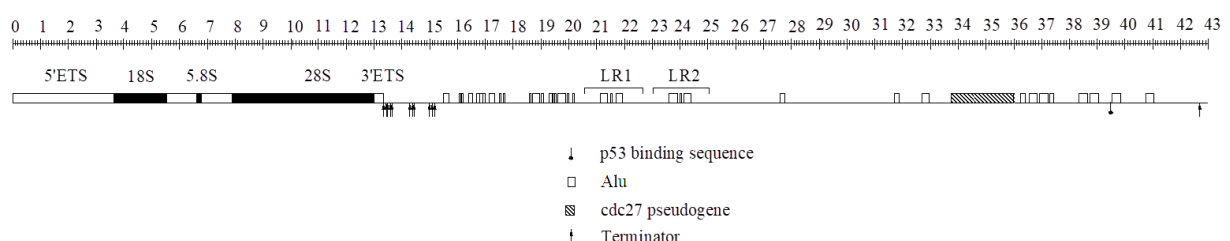
**Fig. 1.9:** Localisation of the rDNA arrays on human acrocentric chromosomes. **A.** The rDNA units are localised between the Proximal Junction and the Distal Junction **B.** Visualized rDNA arrays by DNA combing in head to tail and in a non-canonical orientation. Adapted from <sup>48</sup>

The presence of 300-400 copies of the rDNA repeat unit per haploid genome was revealed by pulse-field gel electrophoresis (PFGE). This was established by the release of intact NORs from chromosomes using restriction endonucleases, which do not cleave the rDNA gene. Individual NORs were then resolved by PFGE and visualized using Southern blot analysis.<sup>50</sup> Also newer PCR technologies like droplet digital PCR (ddPCR) can be used to measure the copy numbers of rDNA arrays.<sup>51</sup> Furthermore, the size of each individual NOR ranges between 50 kb and up to over 6 mb.<sup>47,52</sup> FISH experiments with different karyotypical normal cell lines, and human donors showed remarkable differences in the distribution of rDNA between the single acrocentric chromosomes.<sup>51</sup> Satellite repeat regions create a heterochromatic epigenetic context at the centromeres and telomeres, and it has been proposed that they isolate the rDNA cluster from genes that are transcribed by Pol II and III.<sup>53</sup> The flanking region, on the centromeric side includes a proximal junction (PJ) consist mainly of sequences that are segmentally duplicated elsewhere in the genome, mostly around centromeres of metacentric chromosomes, whereas the distal junction (DJ)

located as flanking region in telomere direction contains a high portion of unique sequences (albeit shared among the acrocentrics) (see Fig. 1.9). Analysis of these identified sequences showed similarities to euchromatic regions. In addition, the DJ was shown to localize to the peri nucleolar heterochromatin, providing an anchor function for active rDNA.<sup>54</sup> Further binding sites for a multivalent DNA binding protein CTCF were identified across the DJ region. CTCF has been previously shown to function in the organisation of human nucleoli.<sup>55</sup> Therefore, although the DJ is located in a heterochromatic region, its complex chromatin environment is indicative of an actively transcribed genomic locus. ENCODE datasets such as RNA-seq. and RNA Pol II and TAFI CHIP seq. datasets and GenBank datasets including mRNA and expressed sequence tags (EST) were mapped to the customised reference genome including the DJ region. This analysis confirmed the actively transcribed nature of the DJ. Two RNAPII promoters were confirmed, one on each arm of the inverted repeats. They are transcribed to generate transcripts that are both spliced, and poly adenylated. Both transcripts were exhibited low to medium expression levels in a numerous cell types and have been proposed to function as long non coding RNAs with potential functions at their site of synthesis<sup>54</sup>. These non-protein coding sequences or non-coding RNA's (ncRNA) are involved in many important biological processes with an impact on gene expression, protein complex formation and recruitment<sup>56</sup>.

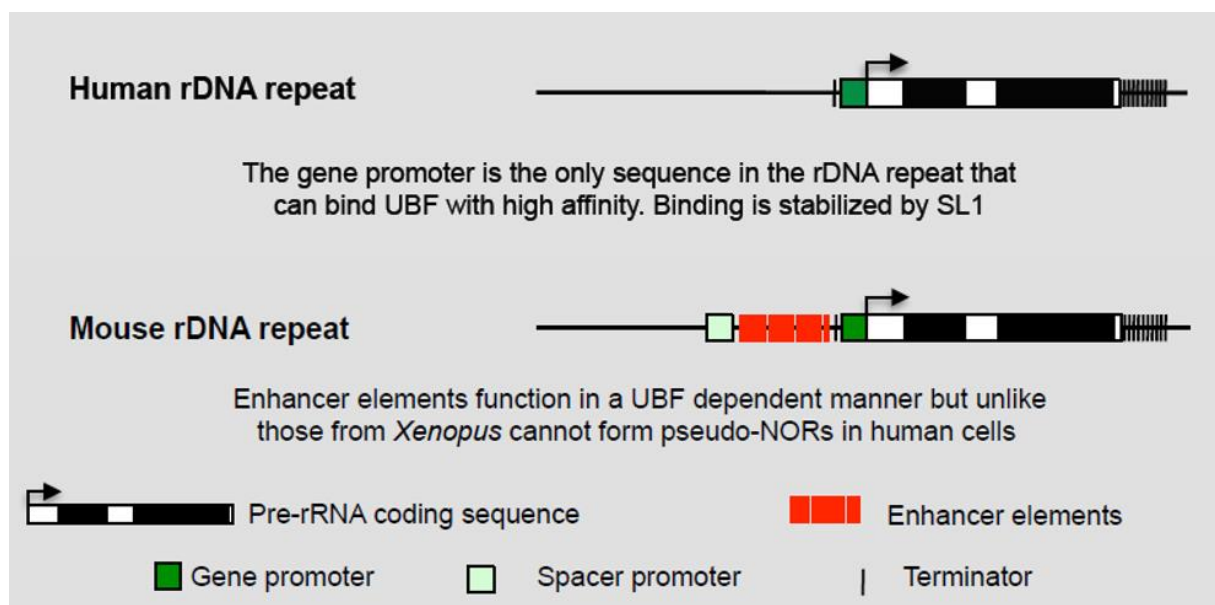
The approximately 30 kb long Inter-genic spacer region (IGS) of the rDNA repeat is housing different elements: such as short interspersed elements (SINEs), like Alu elements for humans (Deininger, 2006; Gonzalez et al., 1992), repetitive enhancer elements, and transcription terminators (see Fig. 1.10) (Goodfellow and Zomerdijk; 2013, McStay & Grummt, 2008). The gene promoter consists of two functional parts, a core promoter in close proximity to the transcription start site, and an upstream control element approximately 100 nucleotides further upstream.<sup>57</sup>

#### Human Ribosomal Repeat



**Fig. 1.10:** A closer illustration of the human rDNA repeat unit, including an IGS. The structure of the murine rRNA gene is depicted. The 5'ETS site indicate the starting site of transcription, initiation of the 47S pre-rRNA and the intergenic spacer promoter. Locations of mature rRNA encoding regions are marked like, terminator regions, repetitive enhancer repeats, Alu elements, and the *cdc27* pseudogene and the large repeats LR1 and LR2 are indicated. A scale is shown below: numbers indicate kilobase pairs and the zero point represents the transcription start site of the rRNA gene. Adapted from <sup>53</sup>

The mouse IGS contains a spacer promoter and repetitive enhancer elements <sup>53</sup>. Interestingly, neither spacer promoters nor enhancer elements have yet been identified in the human IGS (see Fig. 1.11).



**Fig. 1.11:** Comparison of the distribution of regulatory elements in the rDNA repeat of human and mouse. The relative positioning of the gene promoters, spacer promoters, terminator and enhancer elements is illustrated for humans and mouse rDNA repeat. Image was taken from the book “The Nucleus”.

The gene is flanked by transcription terminators at its 5'- and 3'-ends, which are binding elements for the transcription terminator factor I (TTF-I), that stops the elongation of RNA Pol I <sup>58</sup>.

The IGS also includes replication origins and replication fork barriers (RFBs) that prevent collisions between the replication and transcription machineries <sup>40</sup>. As previously described, Miller spreads in budding yeast revealed that there are inactive or silent rDNA units. The evidence is indirect, and it is still an open question if these

inactive rDNA units also exist in the human genome. Recently published Nanopore sequencing data suggest that these inactive rDNA units exist. In terms of rDNA methylation, they found, that there is an obvious difference between methylated and unmethylated 45S rDNAs. In unmethylated 45S rDNAs, the methylation rate was close to 0, the genes are likely to be transcribed. The methylation level or status was also similar in contiguous copies. This may be because heterochromatin forms around these regions and affects rDNA silencing. However, in contrast the noncoding IGS regions were always methylated and the level of the contiguous 45S rDNA. These observations suggest that the IGS is always in a similar heterochromatin structure and that 45S rDNA activation affects the region.<sup>49</sup>

The activity of transcription and replication contributes to genomic instability. Those silent rDNA units, in yeast, contribute a “working space” or a template for the recombinational repair of two sister chromatids.<sup>59</sup>

### **1.3 rDNA transcription mediated by RNA Polymerase I**

#### **1.3.1 RNA Polymerase I**

As evidenced by the density of branches on Miller spreads, rDNA repeats are heavily transcribed. Indeed transcription units appear fully packed by RNA polymerase I. It can be also seen in the Miller spreads, that at the end of the pre-rRNA molecules are large particles attached, the so-called terminal balls.<sup>60</sup> Pol I driven transcription is also very close coupled with processing of pre-rRNA, which is encoding the 18S, 5.8S and 28S rRNA subunits.<sup>61</sup> How this is achieved is uncertain, Pol I does not have a C-terminal domain (CTD) in contrast to Pol II.<sup>62</sup> For Pol II it has been shown that the CTD is involved in transcription initiation, transcript capping and splicing.<sup>63</sup> The Pol I machinery is composed of 14 different polypeptide subunits, 12 of those subunits are similar or have close related counterparts in Pol II and Pol III. The other two subunits are related to factors that activate RNA polymerase II and have structural similarities with Pol III. All these subunits form the catalytic core of the RNA Polymerase I machinery.<sup>64,65</sup> There are two different configurations of Pol I complexes, the most

abundant version is the Pol I alpha, predominantly involved in transcription elongation. A less common configuration is the Pol I  $\beta$  which is required for effective transcription initiation.<sup>66,67</sup> The different RNA polymerases have little promoter specificity and require accessory factors that form a PIC and recruit RNA Pol.<sup>64</sup>

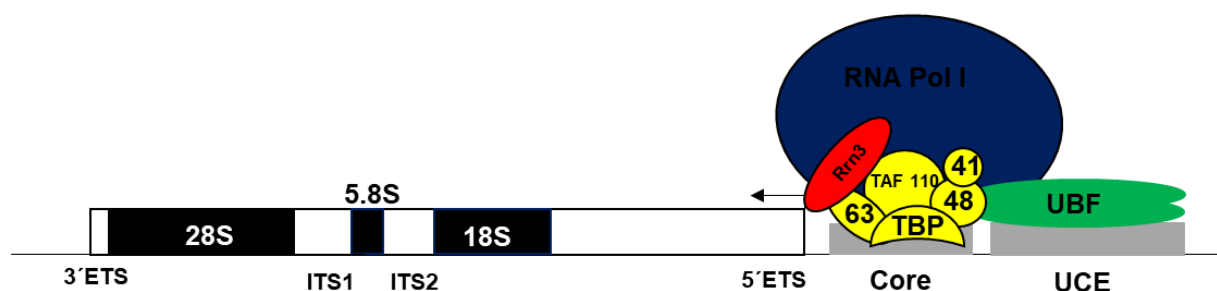
### **1.3.2 Polymerase initiation complex (PIC) Formation**

The different RNA polymerase enzymes themselves have no intrinsic ability to recognize and bind specifically to promoter sequences. The three different polymerases can bind to their promoters based on different transcription factors. The forming of specific PIC is the fundamental mechanism to create a functional transcription machinery. However, there are differences between the single polymerases. Pol II and III can form their specific PIC's at multiple promoter types, whereas Pol I mediated PIC formation just occurs at a single repeated promoter element. Because of this difference between the promoter types for the corresponding RNAP's, evolution of rDNA promoters is free from constraints. Recombination, which is the driving force for (genomic) evolution, has led to a "concerted divergence" of the promoter sequences of all the rDNA genes in a (given) genome. This explains why few sequence differences could be observed within a (given) genome, but they are highly divergent between taxa.<sup>68</sup> In all Eukaryotic cells, rDNA promoters are composed of a core promoter and an upstream control element (UCE).<sup>69</sup> However, there is little evolutionary sequence conservation.<sup>70</sup> Every evolutionary sequence change within the promoter correlates with the (specific) interacting factors, and can be described as molecular coevolution hypothesis.<sup>71</sup> This also explains why the Pol I transcription machinery has a higher rate of evolutionary change than the Pol II and III counterparts and Pol I machinery is incompatible between different species.

As previously described, RNA Polymerases (RNAP) have little intrinsic affinity for the promoter elements. A selectivity factor 1 (SL1) complex is essential for the formation of the Pol I pre-initiation complex (PIC) at the rDNA promoter region<sup>72</sup>.

The formation of a Pol I specific PIC and the transcription afterwards is a complex process, the selectivity factor (SL1) is composed of TATA-box binding protein (TBP) and its associated factors, like TAF<sub>I</sub>110, TAF<sub>I</sub>63, TAF<sub>I</sub>48, TAF<sub>I</sub>41 (now more usually referred to as TAF<sub>I</sub> A-D) <sup>73</sup>.

The SL1 complex binds with sequence specificity to the core promoter element and recruits Pol I, TAF<sub>110</sub> and TAF<sub>163</sub> are specific for this association and the contact to the promoter<sup>74</sup>. Species-specific Pol I transcription is controlled by the SL1 complexes. In mouse-human hybrid somatic cells it can be observed that human NORs are able to associate with mouse nucleoli<sup>75</sup>, and recruit UBF, to the human acrocentric chromosomes but the human promoter sequence is too divergent for recruiting mouse SL1. Human NORs, are transcriptionally silent in this untypical cellular environment<sup>74</sup> and can be reactivated by human TAFs. SL1 also plays a role in stabilizing the association of UBF, with the rDNA promoter region<sup>76</sup>. The only presence of SL1 and Pol I $\beta$  as a complex is essential and sufficient for basal levels of transcription, integration of UBF into the PIC results in an activation of high rates of rDNA transcription<sup>76</sup>. TAF<sub>48</sub> and TBP are subunits of the SL1 complex, which binds to the acidic C terminal domain of UBF<sup>74,77</sup>. The Pol I $\beta$  associated kinase CK2 converts stability to the PIC complex by binding to UBF and TAF<sub>110</sub><sup>78</sup>. Rrn3 (also referred as TIF1A) is essential for recruiting POL I to the PIC, therefore formation of the Pol I/Rrn3 complex is a key step in Pol I transcription initiation<sup>78</sup>, furthermore it should be noted that Rrn3 is functionally conserved in yeast and mammals<sup>79</sup>.



**Fig. 1.12:** Schematic diagram illustrating the arrangement of the PIC in an rDNA array context. A dimer of UBF, in green, is associated with the UCE of the promoter. SL1 is depicted in yellow and includes five TAFs and TBP, which interacts with sequence specificity with the Core promoter element. Rrn3, in red, and the RNA Pol I $\beta$ , in blue. The interactions, which are stabilizing the PIC structure are as follows: The C-terminus of UBF interacts with TAF<sub>48</sub>. UBF also interacts with Pol I $\beta$  subunits PAF<sub>49</sub> and PAF<sub>53</sub>, which are not illustrated. Rrn3 interacts with both the Pol I $\beta$  subunit RPA<sub>43</sub>, not illustrated, and the SL1 factors TAF<sub>110</sub> and TAF<sub>163</sub>.

### **1.3.3 Pol-I $\beta$ Transcription Initiation**

To initiate Pol I $\beta$  mediated transcription after the PIC formation, a promoter escape is necessary<sup>67,78,80</sup>. It has been shown that a promoter escape is rate limiting for rRNA production<sup>81</sup>. Until now it is not fully clear how the mechanism is working. In mice, the interruption of the complex binding between Rrn3/Pol I $\beta$  is essential for a promoter escape and catalysed by a phosphorylation of Rrn3 by kinase CK2<sup>82</sup>. In yeast, the untethering between Rrn3/Pol I $\beta$  is not essential for rRNA synthesis<sup>83</sup>. Probably the importance of the untethering of Rrn3/Pol I $\beta$  is species dependent but it also has been shown that UBF is vital for stimulating a promoter escape<sup>84</sup>.

### **1.3.4 Elongation**

Dividing cells require an enormous amount of ribosomes, and therefore also ribosomal RNA. The high processivity capabilities of Pol I are in part responsible for this elevated rate, each Pol I complex has a transcription rate of around 95 nucleotides per second. In mammalian cells, each active rDNA gene is loaded with around 100 Pol I complexes<sup>85</sup>. In yeast, a rDNA gene contains between 50 to 60 Pol I complexes, with a transcription rate, per complex, of around 60 nucleotides per second<sup>39</sup>.

A number of other factors play an essential role for the high and characteristic Pol I elongation rate. As well as being involved in PIC formation, UBF which binds across the entire rDNA gene aids elongation. The mechanism of UBF interaction and function in that context is not fully clear. It is speculated that UBF contributes towards Pol I elongation by maintaining the chromatin in a transcriptionally permissive state<sup>86</sup>. It can be assumed that, chromatin remodelling is playing an important role in Pol I elongation. Remodelling factors like (WSTF)-SNF2h and G9a (HMT)<sup>71,87,88</sup> can be mentioned here, but also histone chaperones like Nucleolin, Nucleophosmin and FACT<sup>89-91</sup>

### **1.3.5 Transcription Termination**

Like the initiation of Pol I transcription, termination also requires multiple factors and DNA elements. The elongation of Pol I is first halted and then followed by the release of both, the Pol I complex and the pre-rRNA transcript. In contrast to Pol II termination, Pol I termination requires site specific DNA binding proteins, for example TTF1-1 in mice<sup>92,93</sup> or Reb1 in yeast<sup>94</sup>. In recent years evidence showed up that Nsi1 could be

the true termination factor<sup>95,96</sup>. The termination factors bind to terminator elements at the end of the rRNA gene (See Figure) Pol I stops with transcriptionally activity upon contact with this protein. This termination mechanism is highly conserved from yeast to humans. A single conserved termination site (T0) has been identified in *Xenopus*, murine, and human rDNA repeat immediately upstream of the transcription initiation site. TTF1 binds to this sequence element and stimulates and stabilizes transcriptionally activity of Pol I. Therefore it could be possible that TTF1 is playing a dual role<sup>97,98</sup>. The Endonuclease Rnt1, with the exonuclease Rnt1 and RNA helicase Sen1 have been shown to be vital for dissociation of Pol I, in yeast, at the TTS, and because of that for the destabilization of rRNA production<sup>99–101</sup>.

### **1.3.6 Re-initiation**

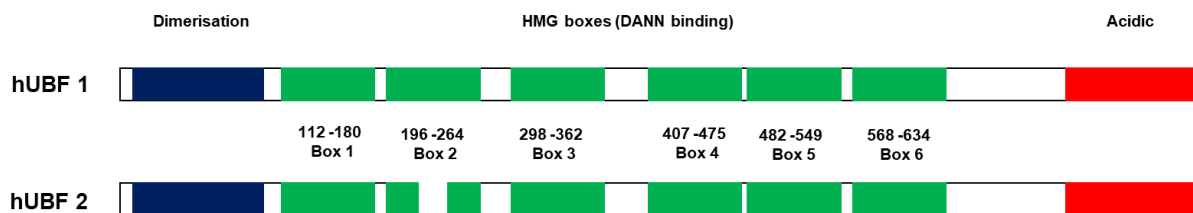
It is vital for every cell and organism to have the ability to re initiate rRNA production after termination. DNA looping formed by juxtaposition of the promoter and terminator elements is suggested to aid the optimized recycling of Pol I machinery. TTF-I binding to sites both downstream of the spacer promoter (Tsp) and the site immediately upstream of the promoter (T0) facilitates these loops. It is suggested that C-Myc is physically associated with the rDNA gene loops<sup>102–104</sup>. Several factors are involved in the formation of rDNA gene loops, which are thought to form a stable core for productive rRNA synthesis by facilitating recycling of Pol I machinery complexes. DNA binding proteins like UBF or complexes like SL1 are proposed to be bound at the promoter positions, after Pol I transcription is terminated. It is necessary after termination, to recruit Rrn3 and Pol I to regenerate the PIC for a re-initiation of transcription<sup>66,78</sup>. However, the assumption of PIC reformation should be viewed with caution. In yeast, the stable association of the core factor (CF) at the rDNA promotor is Pol I dependent<sup>105</sup>. So, the evidence is lacking, that the PIC remains intact at the promoter to support multiple rounds of transcription.

### **1.3.7 The role of UBF in rDNA transcription**

For a long time, it was thought, that UBF function is just limited to the formation of the PIC at the rDNA promoter. The exact mechanistical function of the PIC formation or even the role of UBF in it is still an open question, and less clear that previously believed. UBF is not essential for promoting *in vitro* transcription reactions, in the

human model system, however it is fundamental in *Xenopus* for transcription reactions. Furthermore it have been shown that SL1 can bind to the rDNA promoter in the absence of UBF<sup>106,107</sup>. The role of UBF is not just limited to transcriptional activation at the rDNA, it also includes the stimulation of the promoter clearance by Pol I<sup>107</sup>. UBF binds extensively across the rDNA repeat, in mouse, *Xenopus* and human<sup>86</sup>.

The transcription of ribosomal RNA is facilitated by Pol I in combination with the initiation complex (PIC) consisting of SL1 and UBF. The interaction of UBF and SL1 to form the PIC and activation of transcription was first shown by Tjian and colleagues<sup>72</sup>. UBF is a highly abundant protein, which contains a N-terminal dimerization domain, six high mobility group (HMG) boxes and at least four of them are known to mediate DNA binding<sup>72</sup> and an acidic tail at the C-terminus. UBF is a highly conserved protein, showing 73 % homology between mammals and *Xenopus*<sup>108</sup>. Furthermore, two different splice-variants of UBF, hUBF1 and hUBF2, were found in human with a size of 94 and 97 kDa (see Fig. 1.13)<sup>75</sup>.



**Fig. 1.13:** Illustration of sequence conservation of UBF variants. Schematic sequences of human UBF are illustrated. HMG-boxes 1-4, which mediate DNA binding (green), the N-terminal dimerization domain (blue) and the C-terminal acidic tail (red) are shown.

In vivo UBF dimerizes and the presence of multiple HMG boxes enable this dimer to bend the DNA to a 360° loop. This structure resembles the loop structure of a nucleosome and was called enhancesome reflecting its activating effect on transcription initiation<sup>109-112</sup>. In addition to this function, UBF plays a role also in defining the chromatin architecture of active rDNA repeat units, it binds all over the rRNA enhancer and coding regions and represents an essential structural component of the open rDNA chromatin conformation<sup>75,113,114</sup>. this architectural role of UBF is supported by observations showing that new secondary constrictions as well as

nucleoli form at sites where UBF binding sites of *Xenopus* are integrated into the human genome, in a non-NOR context (Mais 2005). Remarkably, factors of the Pol I transcription and ribosomal RNA processing machinery are recruited to these sites called pseudo-NORs, although the inserted sequences lack the promoter sequences and are therefore not capable of transcription <sup>21</sup>. A possible mechanism for creating and maintaining the open chromatin structure of rDNA is that UBF competes with histone H1, leading to decompaction of chromatin. Recently, evidence arose that besides playing a regulatory role in Pol I transcription UBF also regulates Pol II at highly transcribed histone genes, which is mediated by the shorter splice variant UBF2 <sup>115,116</sup>. In order to further investigate the function of UBF *in vivo* a conditional gene-deletion system was established in mouse, Hamdane et al. (2014) showed by using this system, that UBF defines the active chromatin domains of the rRNA genes and is essential for transcription of these genes. Furthermore, loss of UBF leads to the formation of a proteinaceous nucleolar precursor body containing rRNA synthesis and processing factors, which progressively separates from rDNA loci <sup>117</sup>.

During metaphase, active mammalian NORs can be visualized by the presence of secondary constrictions. Another and more reliable method is staining with Silver nitrate. This method is based on the acidic/argyophilic domains of many of the proteins of the RNA Pol I machinery such as Treacle and UBF, which remain bound to NORs during mitosis <sup>118–121</sup>. Another method for visualising these active NORs can be done by immunofluorescence using antibodies specific for the respective proteins <sup>122</sup>. These methods can just indicate on which chromosome, active NORs (and so active rDNA arrays) are located. Despite the use of these techniques it has been demonstrated that certain acrocentric chromosomes are more likely than others to lack NOR associated silver staining <sup>123</sup>. It is important to keep in mind that half of the rDNA genes in mammalian somatic cells are transcriptionally silent. <sup>124</sup>. This can be monitored by a technique based on the differential replication timing of the active and silent rDNA, with active genes being replicated early and silent genes in the late S-phase. Nascent DNA was pulse labelled with BrdU at various stages during S-phase. Anti-BrdU antibodies were used to isolate the newly replicated DNA. Real time PCR allowed a comparison of the relative fraction of rDNA replicated in early versus late S-phase <sup>125</sup>. This study showed that approximately half of the rDNA genes in mammalian somatic cells are transcriptionally silent. A model for the silencing mechanism is based on the NoRC complex, that consist of TIP5 (TTF-1-interacting prtotein-5) and the ATPase SNF2h.

TTF-1 binds to the promoter proximal terminator site ( $T_0$ ). TIP5 in NoRC interacts with the N-terminal part of full length TTF-I and unmasks its DNA binding site. This interaction is required for binding of TTF-I to its promoter-proximal target site and for the recruitment of NoRC to the promoter in chromatin. NoRC is then suggested to recruit Dnmt1 and Dnmt3b along with HDAC1. This results in DNA methylation and histone deacetylation in the corresponding region. The Methylation of a single CpG dinucleotide (at-133 upstream of the TSS), has been shown to disrupt UBF binding affinity to the promoter and because of this silencing of transcription follows <sup>126</sup>.

### **1.3.8 Processing and ribosome biogenesis**

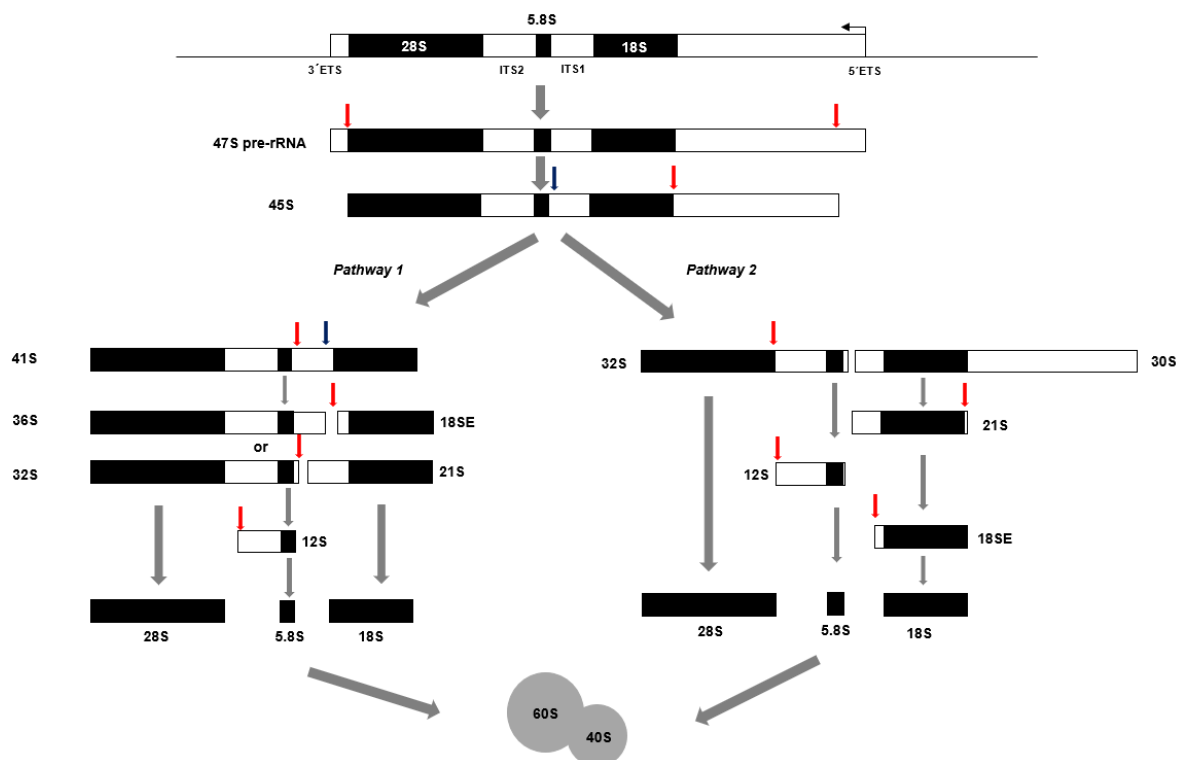
Early ribosomal proteins and pre-ribosomal factors assemble co-transcriptionally in the nucleolus with the newly-synthesized pre-ribosomal RNA (pre-rRNA), forming the so-called 90S pre-ribosomal particle, which can be seen as “terminal balls or terminal knobs” on the “Miller spreads” containing the SSU <sup>127,128</sup>.

In *Saccharomyces cerevisiae*, the majority of the 35S nascent transcripts are cleaved during transcription at specific sites, whereas higher eukaryotic rDNA transcripts are getting cleaved and processed after transcription <sup>129</sup>.

Pre-rRNA processing is a sequential enzymatic driven process which is including cleavage, methylation, pseudo uridylation and folding of the rRNA. The polycistronic pre-rRNA transcript is getting produced by translation in the FC, travers through the DFC and into the GC, the final assembly of ribosomes takes place in the cytoplasm. Multiple processing pathways have been identified and it is possible that these pathways can be employed in different fashion depending on the cell type <sup>130</sup>. For the 47S rRNA, initial processing takes place within the 5'-ETS and the 3'-ETS <sup>131</sup> of the transcript resulting in a 45S product, followed by a endonucleolytic cleavage within the spacer sequences <sup>129</sup>. The resultant 45S pre-rRNA can be processed in two different ways (pathway 1 and 2).

Cleavage of the 45S pre-rRNA can either start in the 5'-ETS or in the ITS1, which defines the two mentioned pathways. If cleavage of the 5'-ETS occurs first (41S pre-rRNA), subsequent cleavage in the ITS1 takes place either at two positions (indicated as red and blue arrows) and will result in two different pre rRNA species.

Initial cleavage at the first positions major pathway in HeLa cells, considering the abundance of the 30S pre-rRNA relative to the 41S. The endonuclease NOB1 is necessary for maturation of the 3' end of the 18S-E pre-rRNA intermediate that is exported to the cytoplasm resulting in 18S rRNA. Formation of the long and short 5' ends of the 5.8S rRNA is not fully documented in mammalian cells. The 5.8S rRNA 3'-end maturation pathway primarily involves exonucleases, but the 7S pre-rRNA was also proposed to result from endonucleolytic cleavage of the 12S pre-rRNA (not shown in the figure) <sup>129</sup>.



**Fig. 1.14:** Alternate Processing pathways of pre-rRNA transcript in human. Red arrows are indicating the single processing site.

Another example of an alternative pathway, starting at the 45S product would lead to additional intermediates, a 43S and subsequently 41S pre rRNA. Processing of this 41S intermediate can then proceed to be modified in a similar manner to the primary processing pathway <sup>131</sup>

In human ribosomes, the 18S rRNA assembles with 33 ribosomal proteins (RPSs) to form the 40S ribosomal subunit, or small subunit (SSU), while the 5S, 5.8S, and 28S rRNAs associate with 47S ribosomal proteins (RPLs) to assemble the 60S large subunit (LSU). Several nucleases, which were initially discovered in yeast, are involved

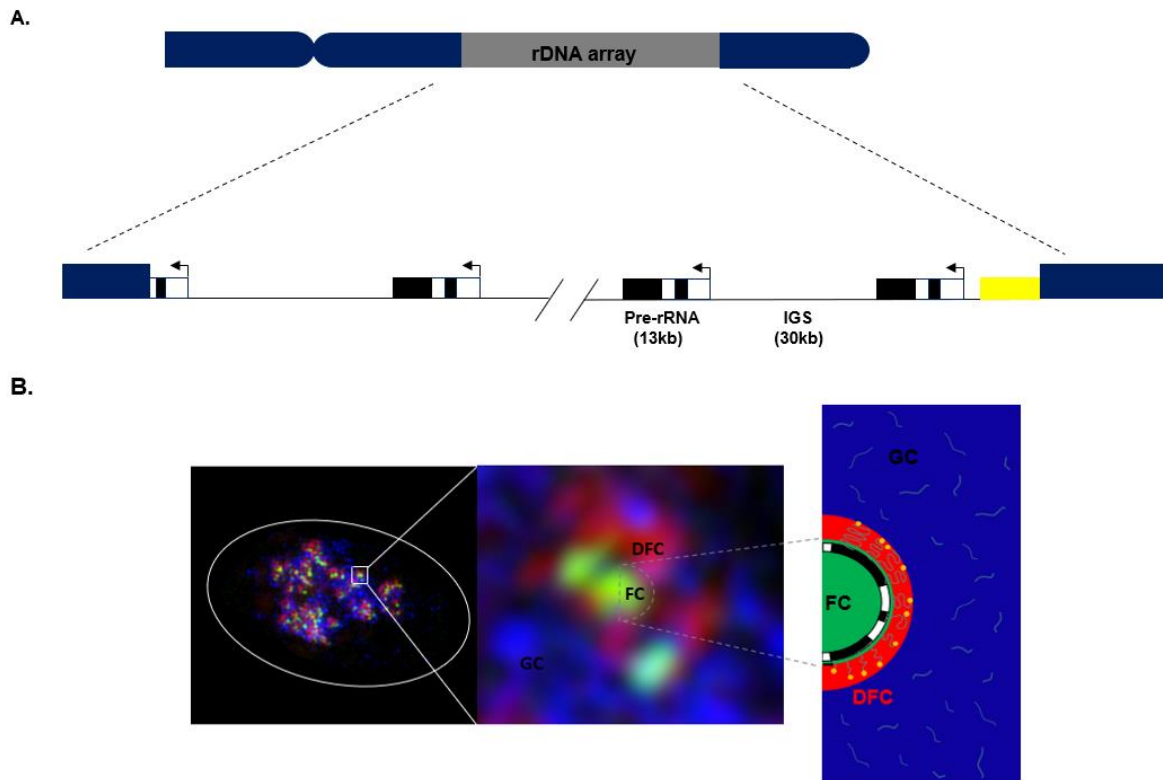
in pre-rRNA processing. Their human orthologs are also involved in rRNA maturation, so the overall processing scheme has been conserved through evolution <sup>131</sup>. However, the pre-rRNA processing in mammalian cells turns out to be more complex than in yeast.

The 5S rRNA is transcribed by RNA polymerase III from repeated gene copies located on human chromosome 1. There little is known about the maturation of 5S precursors in humans <sup>132</sup>.

## **1.4 NORs and the Nucleolus**

### **1.4.1 The nucleolar structure**

The nucleolus is the most obvious structure of the interphase nucleus, it houses the rDNA repeats which are organized as NORs and it can be easily seen by a dark field phase contrast microscope. In the previous chapters, NORs were described in the aspect of their scientific history, genomic architecture (see Fig. 1.15 A) and their function as transcription templates to generate rRNA for ribosome biogenesis. Individual nucleoli form around NORs immediately after transcription re-initiates this takes place after M-phase. All involved protein complexes and components for transcription and rRNA processing take place in the new forming nucleoli. Self-interactions of these protein complexes with each other and with different polynucleotides, such as rDNA and rRNA leads to a self-assembly organization based on a mechanism of the nucleolus <sup>133,134</sup>. We will discuss this in more detail later.

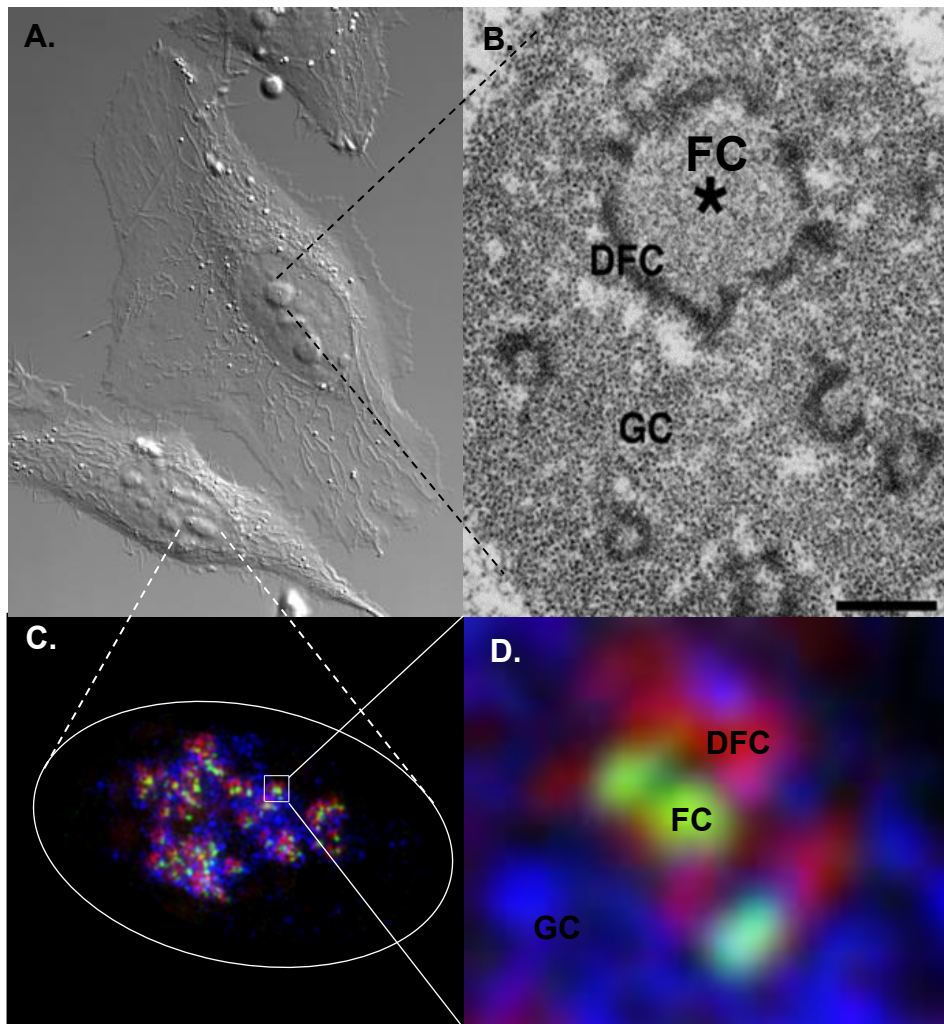


**Fig. 1.15:** Organization of Nucleolar Organizer Regions (NORs) in Nucleoli **A.** the genomic architecture of NORs, indicating their position on the short arms of the acrocentric chromosomes. Transcription of rDNA arrays occurs in a telomere to centromere direction. The distal junction (DJ) (yellow) is located adjacent to the ribosomal gene arrays. **B.** internal organization of the nucleolus, in a mature interphase nucleolus multiple NORs come together. The fibrillar centres (FCs) (green) are surrounded by the dense fibrillar component (DFC) (red). At the interphase between FC and DFC, RNA polymerase I (Pol I) transcription is initiated. Nascent pre-ribosomal RNA (pre-rRNA) are bound by the small subunit processome (grey lines with yellow balls). Late pre-rRNA processing intermediates released into the granular component (GC) (blue) for maturation and assembly into ribosomes. IF images taken from <sup>135</sup>.

The nucleolus of higher eukaryotes exhibits a tripartite structure intimately dependent on the ongoing transcription of the rDNA by the RNA Pol I complex. Three sub-compartments dominate the nucleolar interior when observed using electron microscopy <sup>24</sup>. The fibrillar centre (FC) hosts rDNA and a pool of transcription factors. The size and number of FC's can vary depending on cell type and ribosome biogenesis requirements <sup>20</sup>. The FC's are surrounded by the dense fibrillar components (DFC), rDNA transcription occurs between the boundary layer of the FC/DFC, whereas early rRNA processing takes place in the DFC. The rRNA transcript is subject to early

processing events within the DFC. The granular component (GC), surrounded by a peri nucleolar heterochromatin (PNH) shell, is hosting the FC/DFC and is the site of late processing of rRNA and initial ribosome assembly (see Fig. 1.15 B and Fig. 1.16)

20.



**Fig. 1.16:** Different perspectives on nucleoli and their internal organization. **A.** a phase contrast image of a cell. **B.** electron microscope image of a nucleolus. **C.** and **D.** SIM-IF image of a nucleolus IF images taken from <sup>135</sup>

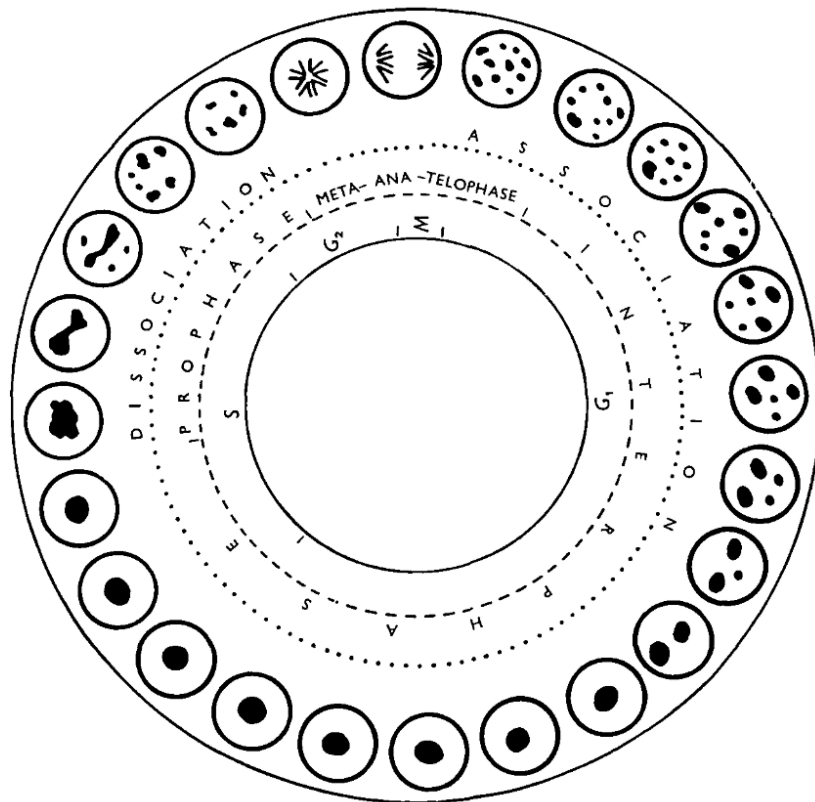
The processes which are necessary to start from an rDNA array to a fully assembled ribosome can be seen, as a sequential process. Especially rDNA transcription and rRNA processing allows us to identify the different compartments of nucleoli using immunofluorescence (IF). Good markers for the FC compartments are UBF and

Treacle, a protein based on the TCOF-1 gene <sup>136</sup>. Good markers for the DFC compartment are proteins, which are involved in early processing of rRNA, such as Fibrillarin <sup>137</sup>. Later processing events take place in the GC compartment, Nucleolin and Nop52 can be used as typical bookmarkers for this compartment, which is representing the nucleolus <sup>51</sup>. Based on these bookmarker proteins, for the different sub-compartments, we can take advantage and facilitate the study of nucleolar dynamic events and the mechanism and possible models of nucleolar formation and fusion.

#### **1.4.2 Nucleoli in different species**

The typical tri-partite structure of nucleoli in higher eukaryotes, does not extend to all organisms. Yeast, for example, just contain a single nucleolus that forms around a single NOR. The yeast nucleolus just consists two compartments, a fibrillar- strand network embedded in granules, that remains connected to the nuclear envelope. this kind of nucleolar architecture is not just unique to yeast, it is also existing in invertebrates and anamniotes <sup>138</sup>. *Drosophila melanogaster* also lack a fibrillar centre <sup>20</sup>. One explanation of the tri-partite structure of nucleoli, as we can find it in human and mice, could be linked to the increase in the length of the IGS. This could lead to a change in the fibrillar network, which is now enclosing the FC compartment as a own layer (DFC) <sup>138</sup>.

### 1.4.3 The Nucleolus during the Cell Cycle

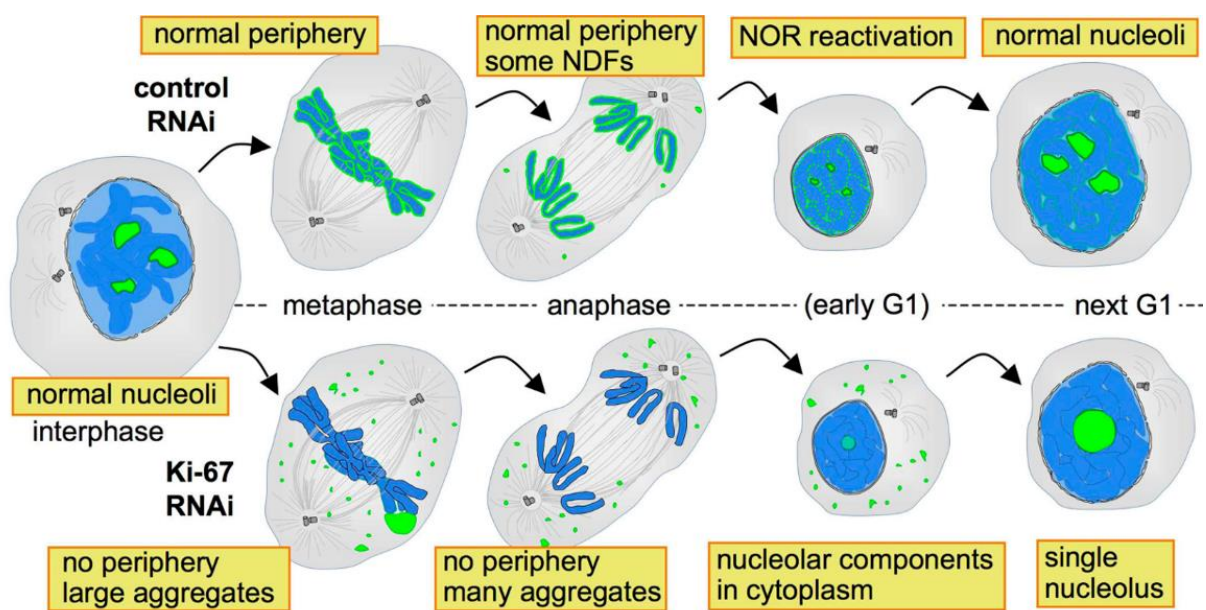


**Fig. 1.17:** The illustration of the nucleolar cell cycle by a drawing. Figure adapted from <sup>139</sup>

Somatic cells of higher Eukaryotes undergo “open” mitosis. Whereas yeast would exhibit “closed” mitosis, which would mean that nucleoli remain intact and transcription rate slows down by the action of Cdc14 during the mitotic event <sup>140</sup>. During this “open” mitosis, transcription of the rDNA stops and the, and the whole tripartite organization of the nucleolus is disrupted and breaks down <sup>141</sup>. This dismantling of the nucleolus is a sequential process, processing factors are released and followed by the cessation of Pol I transcription.

rDNA transcription rate is reduced by 30% in early prophase and will completely stop by the end of prophase <sup>140</sup>. By prometaphase nucleoli will be no longer visible and many of the DFC and GC components relocate to a peri chromosomal compartment, whereas many FC proteins localize at the secondary constriction. This kind of localization is observed in many vertebrates and plant cells. Typical DFC and GC components like, Fibrillarin, U3 snoRNAs, Nucleolin, Nop140, NPM/B23, Bop1, Nop52,

PM-ScI 100 and Ki67 can be found as a part of this peri chromosomal compartment<sup>142–144</sup>. Recent findings suggest that Ki67 could play a fundamental role. It is widely used as a proliferation marker in basic research and cancer prognosis. During interphase, the Ki67 protein can be exclusively detected within the cell nucleus, whereas in mitosis most of the protein is relocated to the surface of the chromosomes<sup>145</sup>. It is present during all active phases of the cell cycle (G1, S, G2), but is absent in quiescent cells (G0)<sup>146</sup>. Cellular content of Ki67 increases during cell progression through S phase<sup>147</sup>.



**Fig. 1.18:** A role for Ki67 in Maintaining the chromosomal periphery compartment. Blue: chromatin, green: perichromosomal proteins (Ki67), grey: microtubules. Figure obtained from<sup>148</sup>

Depleting Ki67 results in the failure of many proteins usually associated with the chromosome periphery to localize, and moreover revealed Ki67 to be required for generating this compartment<sup>148</sup>.

Mitotic “active” NORs form constrictions, or achromatic gaps on metaphase chromosomes and can be indirectly visualized with DNA stains such as DAPI. Certain FC factors remain associated with these “active” NORs ready for resumption of transcription at the exit from mitosis and initiating nucleolar biogenesis. PNB proteins

are recruited to these reforming nucleoli in a pre-rRNA dependent fashion.<sup>149,150</sup> Responsible for the formation of this secondary constriction and the high affinity for silver staining is the RNA Pol I machinery, which is associated with the mitotic NORs<sup>151–153</sup>. Because of that the secondary constrictions are also called AgNORs. FRAP analysis has illustrated the dynamic nature of the association of several FC relevant proteins with mitotic NORs. RNA Pol I subunits may transiently dissociate from NORs during mitosis, therefore it can be assumed that the FC like structure localizing at NORs keeps the dynamic nature and is not composed of locked down or “deactivated” RNA Pol I machinery<sup>154,155</sup>.

The reactivation of rDNA transcription in telophase results in assembly of nucleoli around the NORs. The entire process of nucleolar reassembly takes a specific amount of time. Experiments in HeLa cells showed that the full assembly can take around 1.5 hours<sup>156</sup>. In human cells, the pool of transcription machinery components at the FC-like structures is getting reactivated, transcription is starting and based on that also the rRNA processing. Relevant rRNA processing factors are getting recruited, which are part of the DFC and GC, the tri-partite nucleolar structure takes place again. However, it should be noted that, that activation of the Pol I machinery is the first step in a chain reaction of nucleolar formation, but it is neither necessary nor sufficient to drive nucleolar formation<sup>157</sup>. Mediator proteins for the disassembly process have been identified, these proteins which are capable of catalysing the reversibly disassembly of nucleoli in the presence of ongoing rRNA transcription. As a result, it is possible to see nucleolar formation as a process which is uncoupled from rDNA transcription<sup>158</sup>. Nucleolar assembly is dependent on the involved processing protein complexes, snoRNAs, the transcribed and processed rRNA (45S from prophase) and of the ribosomal proteins. All these complexes and components localize to the perichromosomal compartment during metaphase and must be translocated to the sites of active rDNA transcription. This process is accomplished in two distinct phases. When chromosomes unwind into chromatin (telophase), the perichromosomal compartment is dismantled to form pre-nucleolar bodies (PNBs), which have the function of a temporary nucleolar subdomain. It was assumed that they have simply the function as a pool of processing factors for the developing nucleolus. However, this view could be too simple. It was found that pre-rRNA maturation of the 45S rRNA, transcribed during prophase, occurs at these sites. Blocking of maturation steps, results in an

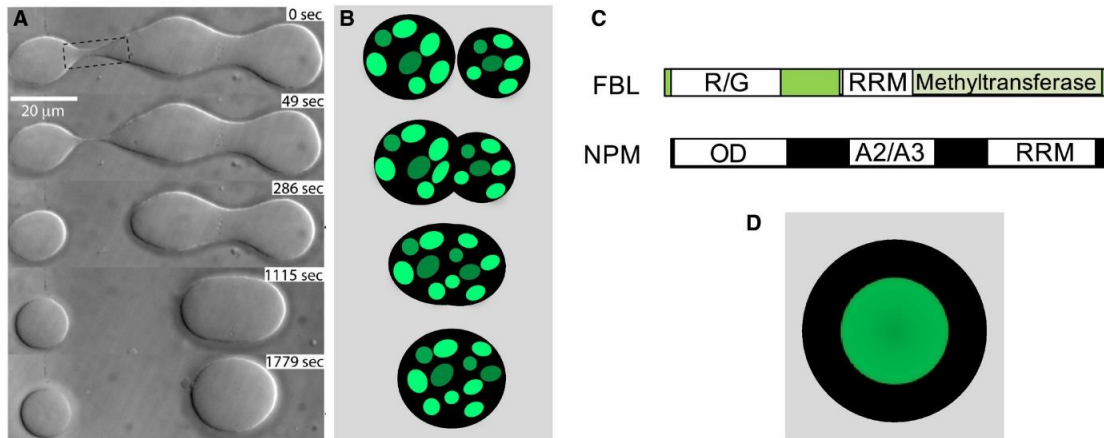
accumulation of stalled pre ribosomes in the PNBs delayed PNB dissolution in G1. So it can be stated, that PNBs fulfil the role of a extra-nucleolar ribosome maturation site<sup>159</sup>.

As previously described, the recruitment of rRNA processing factors to the actively transcribed NORs occurs in telophase, and single nucleoli form around them. Now the assembly and growing of the nucleoli take place. The concentration of those factors is decreasing in the PNBs and increasing in the nucleoli. This “concentration gradient” or flow of processing factors is complete within 15-20 minutes, for factors which are associated with the FCs. The complete transfer of the GC factors requires around 90 minutes. Interestingly, the traffic between PNBs and nucleoli was unidirectional for the DFC protein fibrillarin, and bidirectional for proteins such as B23/NPM and Nop52 as typical GC proteins. Another interesting fact is, that the transfer of those factors was not driven by close proximity and fusion of the PNBs with the (growing) nucleoli. It can be suggested that PNB formation is a way of controlling the formation of nucleoli by step wise release of DFC and then GC factors<sup>149</sup>. The entire process, of nucleolar biogenesis can take up to 1.5 hours in HeLa cells, which is also including the fusion event(s) of individual nucleoli into one large centrally located mature nucleolus<sup>160</sup>. It has been accepted for a very long time that nucleoli can fuse in somatic cells<sup>161</sup>. A recent study in *Xenopus laevis* oocytes offered one explanation for this phenomenon. Fusion events were induced by mechanical action or disruption of actin, observations based on these experiments brought up the idea of viscous liquid-like droplets<sup>162</sup>.

#### **1.4.4 Liquid-Like phase separation as a model for the dynamic behaviour of nucleoli?**

In 2011 Brangwynne used amplified nucleoli within the germinal vesicle (GV) of mature oocytes of *Xenopus laevis* to show the dynamic aspects of nucleoli in the context of a model following the behaviour of viscous liquids (liquid phases). But there is also a historical connection to the term, in 1946, Ehrenberg described the nucleolus as “A separated phase out of a saturated solution” and also as a spherical structure. He showed a connection between temperature changes and the size of the nucleolus, in the plant *Salix fragilix alba* (willow)<sup>163</sup>

By isolating GVs, we can physically manipulate GV to facilitate the merging of two unfused nucleoli, the merging nucleoli initially exhibit an hourglass conformation before fusing to form a larger spherical structure (Figure 1.19 A.)<sup>162</sup>. The bridge forming, disruption and the fusion event of nucleoli indicate a viscous liquid-like behaviour.



**Fig. 1.19:** Liquid-like Properties of nucleoli during nucleolar fusion. Figure from adapted from<sup>164</sup>

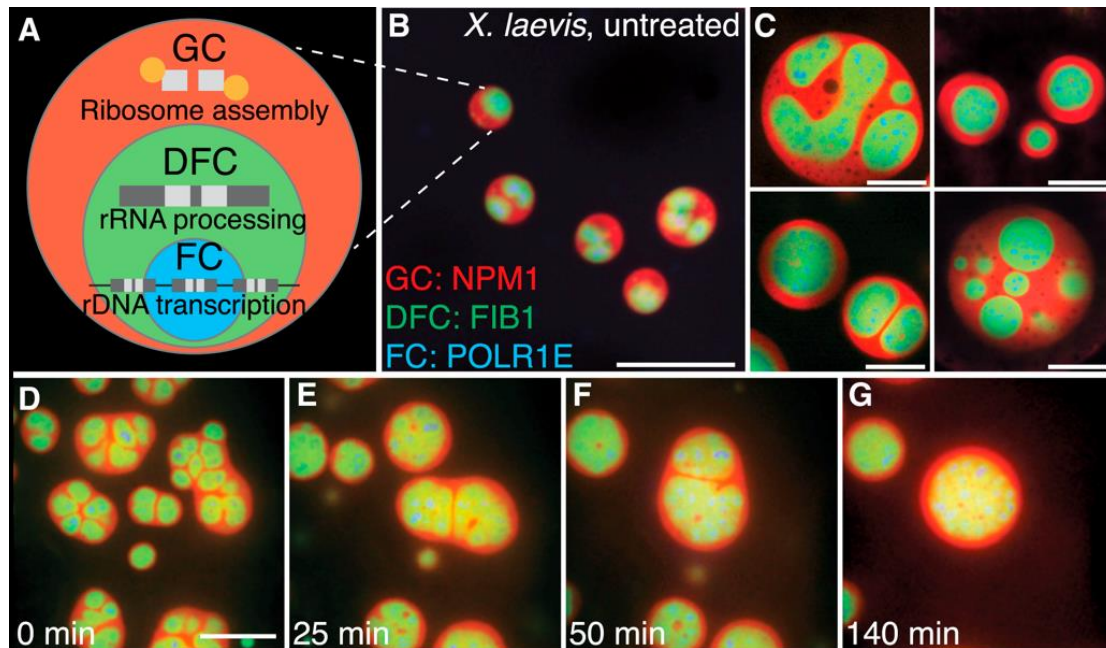
In addition to liquid-like behaviour, which is playing a role in nucleolar fusion events, the compartmentalization of the nucleolar interior also shows LLPS properties.

By injecting mRNAs into the ooplasm, which are encoding, relevant FP tagged key proteins for the different sub-compartments like the GC and DFC (Gfp:FBL for the DFC and RFP:NPM for the GC), the behaviour of those sub compartments could be tracked during fusion events. The GC compartments coalesce and an additive number of DFC's can be observed in the new emerged nucleolus (Fig. 1.19 B).

Recombinant NPM and FBL individually combined with purified rRNA condense, *in vitro*, into a liquid-like droplet interphase (Fig. 1.19 D.). It is the energetically preferred condition, because of the lowest state of surface tension.

In the GV oocytes of *Xenopus laevis*, the amplified nucleoli are immobilised in a 3D space by an F-actin scaffold. By fluorescently tagging of nucleolar key proteins for the single sub compartments (FC, DFC, GC) and disrupting the F-active scaffold restricting nucleolar fusion merging of the unfused nucleoli could be observed (Fig 1.19.)<sup>165</sup>. Furthermore, the internal organisation of the nucleolar follow a LLPS behaviour. The GC encompasses the DFC that appears to fuse into a spherical structure containing

dispersed FC compartments throughout. Interestingly the disruption of the dense nuclear actin network facilitates nucleolar fusion, in oocytes of *Xenopus laevis*. Conversely, there is no filamentous actin network present in human (interphase) cell nucleus.



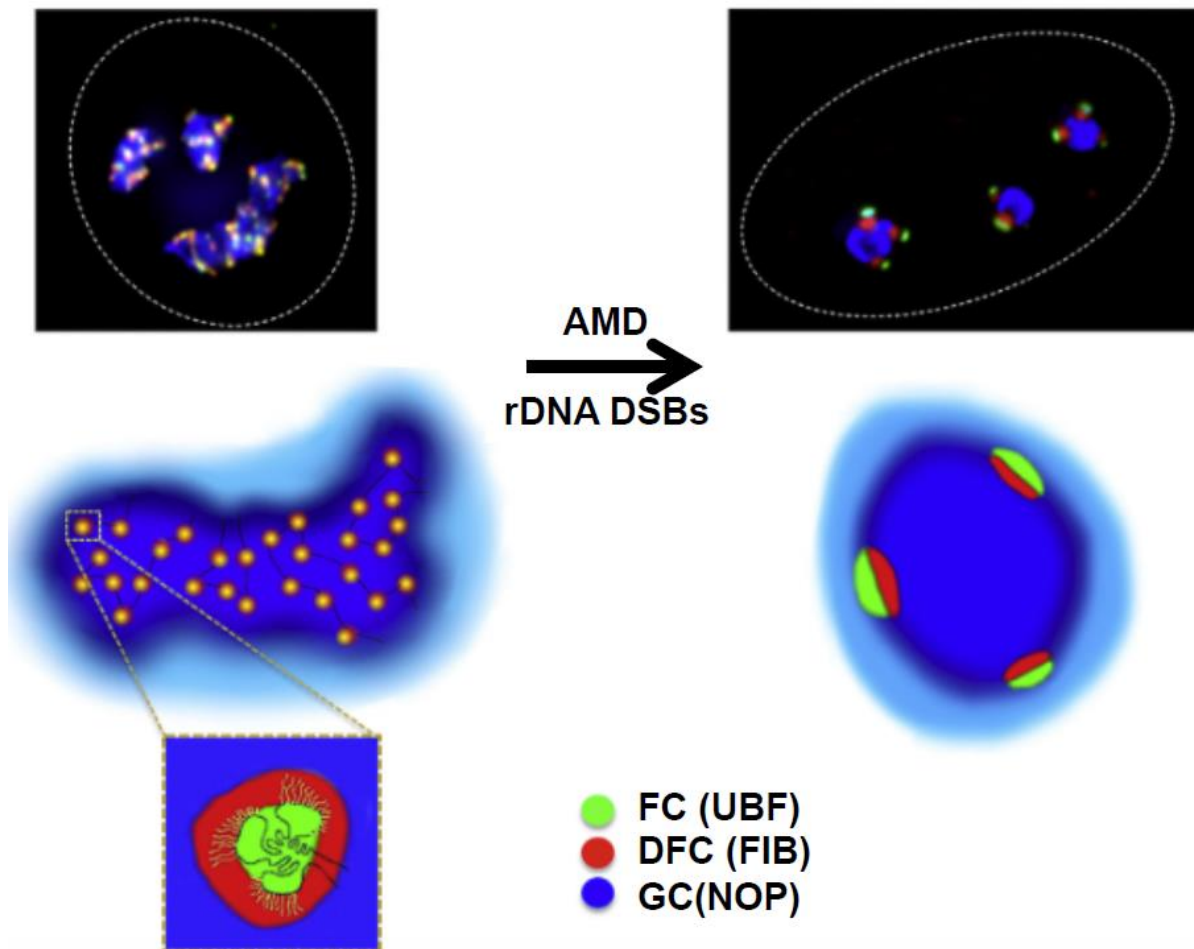
**Fig. 1.20:** Internal liquid-like organisation of Nucleoli. (Figure adapted from <sup>165</sup>)

Disordered domains within each protein are responsible for this droplet formation with RNA recognition motifs (RRMs) facilitating the maintenance of phase separation <sup>164,165</sup>.

Based on these *in vitro* and *in vivo* experiments we were able to see that the involved components for the single nucleolar sub-compartments follow the rules of viscosity and show a typical liquid-like behaviour. The driving physical force behind this behaviour is surface tension. However, this behaviour just relies on the interactions of the single biomolecules, which are part of the nucleolar sub-compartments. But it is also fundamental, to think about the starting process of nucleolar assembly and the resulting function, rRNA transcription. This process is one of the most energy consuming processes, and active over most stages of the cell cycle. Wieschaus et al, showed that nucleolar assembly is a thermodynamical process, in *Drosophila* embryos. Active assembly is faster at higher temperatures, and slower at lower temperatures, which can be expected for an enzymatic dependant process. But this

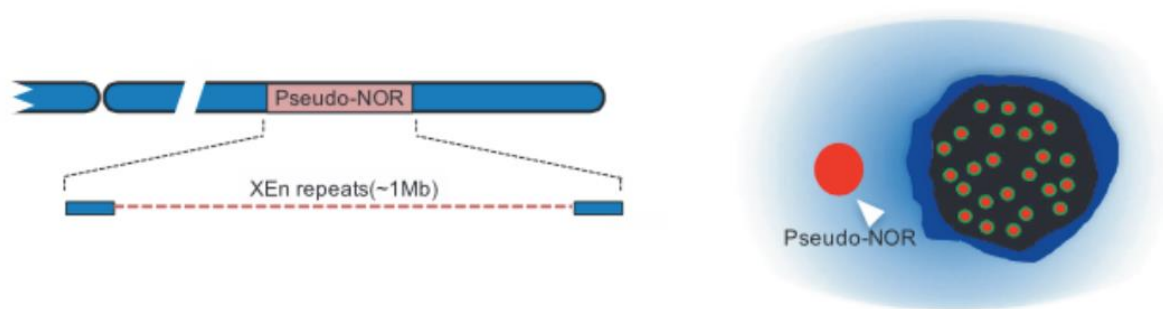
behaviour is the opposite of a thermodynamic LLPS, in which more condensation would occur at lower temperatures, seen from the perspective of thermodynamics. Physical quantities such as temperature, time and concentration have a massive influence on the behaviour of single substances, which are forming specific (liquid) phases. The formation of nucleoli in *Drosophila* is precisely timed and follows the activation of rDNA transcription <sup>150</sup>. This shows that LLPS is just a very vague description in the context of a ribosome biogenesis, these *in vivo* experiments are simply reporting the behaviour of overexpressed fusion proteins. Endogenous tagging of relevant (nucleolar) targets would offer a possibility to observe the natural behaviour of a fully functional protein based on the natural cell-controlled expression rate.

Another interesting perspective comes the disruption of rDNA transcription by inhibition of RNA Polymerase I using Actinomycin D (AMD) or due to the presence of DNA double strand breaks (DSB) within the rDNA. Both result in nucleolar segregation. The FC and DFC compartment depart from their typical internal localization to form dense bi-partite nucleolar caps at the nucleolar periphery of spherical nucleoli (GC). This would make the rDNA breaks accessible to the DNA repair machinery, which is located external to the nucleolus <sup>166</sup>.



**Fig. 1.21:** Nucleolar cap formation driven by nucleolar stress response, catalysed by RNA Pol I inhibitor AMD or by DNA double strand breaks (DSBs). Figure adapted from <sup>167</sup>

As previously described the recent findings presenting LLPS as a model for nucleolar formation in *Xenopus laevis*. This does not address endogenous NORs associated with chromosomes. Pseudo NORs, an ectopic array incorporated into non-acrocentric chromosomes highlighted the importance of UBF maintaining secondary constrictions at NORs . During interphase, pseudo NORs appear as a spherical droplet-like structure in the nucleoplasm. They recruit factors that are present within the FC compartments of nucleoli but are not transcribed <sup>168</sup>. Pseudo-NORs can be just viewed as a FC droplet that has not been enveloped by a DFC, because of the absence of transcription.



**Fig. 1.22:** location of Pseudo-NORs and their droplet-like structure in the context of the LLPS model. Figure adapted from <sup>164</sup>

Based on these observations the chromosomal location may be a factor, because the pseudo NOR is located on a non-acrocentric chromosome and thus it is in chromosomal context different from endogenous NORs <sup>169</sup>. Therefore, it is necessary to see the LLPS model also from the perspective of chromosomal location.

Based on those results we can conclude that the LLPS model could play a role in the behaviour of nucleoli and their sub compartments. Nevertheless, it is necessary to think about the general (physical) definition of phases and phase separation. Phase separation is the creation of two distinct phases from a single homogeneous mixture of two (pure) components. Physical quantities and properties like critical concentration, temperature, solubility, and surface tension play a key role for the mechanism of this event. LLPS in the context of cellular biology is more complicated and not clearly defined. We do not have single components, which would follow the described physical event, instead we have a very complex mixture of different biomolecules and biopolymers (DNA, RNA) in a specific cellular environment. LLPS criteria in cellular environment would be the concentration of the single involved components, temperature and ion strength, and of course also the chromosomal location of the starting point of those events (NOR forming). Therefore McSwiggen et al, concluded endogenous levels of involved proteins are the best way for further investigations on this topic of an LLPS mechanism in living cells <sup>170</sup>.

### **1.5 Aims of this thesis.**

The aims of this thesis are to develop a method to endogenously tag key essential nucleolar proteins in a suitable cell line, such as hTERT-Rpe1 cells. This would offer the possibility of visualising the dynamic behaviour of nucleoli in a non-transformed human cell line, certain in the knowledge that the FP tagged proteins are fully functional and expressed at normal levels. hTERT-Rpe1 cells are a karyotypically normal diploid cell line and are well characterised in the context of its NORs. Therefore hTERT-RPE1 cells are a suitable model system for future experiments. Live cell imaging based on tagged nucleolar key proteins for the single nucleolar sub-compartments would help to increase the perspective for nucleolar dynamics. This could be particularly helpful in view of the current LLPS discussion.

At the outset of this PhD, no method for generating CRISPR/Cas9 generated knock-In's in this normal/non-transformed cell line had been described. Older methods such as overexpression based on plasmid constructs do not necessarily show the natural behaviour of the tagged proteins. Therefore, the CRISPR/Cas9 nuclease system is the newest and most reliable approach to tag genes with fluorescent proteins (FP).

Regarding to the published literature there are some Knock-Out's in hTERT-Rpe1 cells based on CRISPR/Cas9 but no Knock-In's are known. Therefore, it was necessary to develop a method to achieve this goal. Chapter 3 and 4 will show the development of a method to tag genes at the C-terminus in hTERT-Rpe1 cells.

## **Chapter II**

## **2. Materials and Methods**

### **2.1 DNA techniques**

#### **2.1.1 Bacterial transformation**

An 50µL aliquot of chemically competent cells (stored at -80°C) was thawed on ice. 1-100 ng of plasmid DNA was added. The Eppendorf tube was mixed briefly and then incubated for 15 minutes on ice. The cells were heat shocked for 1 minute at 42°C and stored on ice for another 5 minutes. 1ml of LB broth was added to the cells, which were then incubated for an hour at 37°C. The cells were plated on two agar plates, supplemented with the appropriate antibiotic (1:1000 dilution for ampicillin (50 mg/mL) or kanamycin (25 mg/mL). Typically one tenth of the transfected cells were plated onto one plate with the remaining cells being pelleted and plated on the other. The agar plates were incubated overnight at 37°C.

#### **2.1.2 Plasmid Purification**

Small cultures of 10 ml LB broth were incubated overnight at 37°C at 250 RPM in the presence of the appropriate antibiotic. The cultures were then pelleted at 4000g for 15 minutes at 4°C. Plasmid DNA was isolated using the NucleoSpin® Plasmid kit (Machery-Nagel Cat No.740588.250).

Large scale plasmid preparations required 400ml broth cultures. The cultures were then pelleted at 4000g for 15 minutes at 4°C. Plasmid DNA was isolated using the NucleoBound® Xtra Maxi (Machery-Nagel Cat No. 740414.50) kit. DNA was eluted from the column and then precipitated with 10.5ml 100% isopropanol (Fisher). The DNA was pelleted at 4000 g for 30 minutes at 4°C. The pellet was then rinsed with 70% ethanol and centrifuged at 4000g for 30 minutes at 4°C before being re-suspended in 0.5-1ml 10mM Tris pH 8.0, 0.1mM EDTA (TE, Fisher), depending on the size of the DNA pellet.

### **2.1.3 Restriction Digest**

A restriction enzyme digest was typically performed on 0.5-1 µg of plasmid using 20 units (U) of a particular restriction enzyme from NEB or Roche. The reaction was incubated at 37°C for 30 min in a volume of 20 µl.

Cas9 restriction digest was performed with 2 µl of the RNP solution (see 2.4.1) and with 3-5 µg of plasmid at 37°C for 2h.

### **2.1.4 DNA Ligation**

Routine DNA ligations were performed with T4 DNA ligase (NEB) in accordance with the manufacturer's instructions. 10µL of the ligation reaction was transformed into competent bacterial cells.

### **2.1.5 Gibson-Assembly**

Gibson-Assembly reactions were performed with Gibson Assembly® Master Mix – Assembly (NEB E2611) according to the manufacturer's protocol with minor alterations. 5µL of the reaction was prepared rather than 10µL. This reaction was incubated at 50°C for 30 min and 3µL were transformed into competent cells. The two components for this reaction, vector and at least one insert were mixed in an amount ratio of at least 1:3. The inserts, usually PCR products, include a 25-30 bp overhang at both ends.

### **2.1.6 Agarose gel electrophoresis**

An agarose gel was prepared with 0.8-1.0 % w/v agarose based on 1xTAE or 1xTBE buffer, depending on the requirements. The agarose solution was heated in a microwave and allowed to cool before adding 0.5 µg/ml EtBr and pouring it into a cast to set. The DNA was loaded to the 1X TAE agarose gel wells with 10x DNA loading dye (10mM Tris pH8, 10mM EDTA pH8, 50% (w/v) sucrose (Fisher), 0.15% (w/v) Bromophenol Blue). The gel was run in a horizontal electrophoresis chamber with the appropriate running buffer. The DNA fragments were resolved using 80-100 volts (V). The appropriate HyperLadder™ (Bioline) was loaded alongside the DNA samples. The gel was then imaged using on a Gbox Imager (Syngene).

### **2.1.7 DNA gel extraction and PCR purification**

Appropriate restriction digests were performed and DNA bands visualised on bench top UV transilluminator (Benchtop 3UV™ UVP) at 302nm. The selected bands were extracted from the gel using a scalpel and transferred to an eppendorf. DNA was purified from extracted bands using NucleoBound® Extract II (Machery-Nagel Cat No. 740609.50) according to manufacturer's protocol.

PCR purifications were also performed using the NucleoBound® Extract II (Machery-Nagel Cat No. 740609.50) kit according to the manufacturers instructions. Elutions were performed with TE.

### **2.1.8 Gateway Cloning**

Sequences were shuttled from one Gateway® vector to another using LR Clonase II Enzyme according to the manufacturers protocol with some minor alterations. A 10µL reaction was prepared rather than 20µL. This reaction was incubated for 30 min at 25°C before being transformed into competent cells.

### **2.1.9 Preparation of Genomic DNA from human cells**

Cells were cultured in a 6 well plate, and harvested at log growth phase. Media was removed and cells were washed twice with PBS. 5ml of lysis buffer (20mM Tris pH8, 2mM EDTA, 0.5% SDS, 0.3mg/mL Proteinase K) was added to the flask. The flask was then incubated ON at 50°C with gentle agitation. The sample was then transferred to a 50mL tube, and DNA was purified using Phenol/Chloroform. 1/10 the volume of 3M Sodium Acetate was added to the sample followed by an equal volume of Phenol/Chloroform. The sample was incubated with gentle rotation at RT for 20 minutes, before being centrifuged at 4000g for 10min. The aqueous phase was transferred to a fresh tube and supplemented with 500µg RNase I followed by a 30 minute incubation at 37°C. The Phenol/Chloroform reaction was repeated and DNA was precipitated by addition of 2.5 volumes of 100% Ethanol and incubation at -20C for 20 minutes. The DNA was spooled onto a plastic Pasteur pipette and desalted in 70% Ethanol. The DNA was resuspended in TE (10mM Tris pH8, 0.1mM EDTA).

### **2.1.10 Polymerase Chain Reaction**

Q5 polymerase (NEB) was used for the amplification of PCR products. The Q5 polymerase was used according to manufacturers' instructions. All PCR reactions were run on Applied Biosystems ProFlex Thermal Cyclers (Life Technologies™).

### **2.1.11 Primers**

The primer design is based on primer blast (online tool: <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) and the SnapGene viewer software, depending on the function of the designed primer pairs. Primer pairs based on genomic DNA templates were usually designed by primer blast.

<b>mAGP2Af</b>	GCAAGCTTCCATGGTGAGCGTGATCAAGCC
<b>mAGP2Ar</b>	GGGTCCAGGGTTCTCCTCCACGTCTCCAGCCTGCTTCAGCAGG CTGAAGTTAGTAGCTCCGCTTCCCTTGGCCTGGCTGGGCAGC
<b>mChP2Af</b>	GCAAGCTTCCATGGTGAGCAAGGGCGAGGAG
<b>mChP2Ar</b>	GGTCCAGGGTTCTCCTCCACGTCTCCAGCCTGCTTCAGCAGGC TGAAGTTAGTAGCTCCGCTTCCCTTGTACAGCTCGTCCATGC
<b>HALOP2Af</b>	GCAAGCTTCCATGGCAGAAATCGGTAAGTGG
<b>HALOP2Ar</b>	GGGTCCAGGGTTCTCCTCCACGTCTCCAGCCTGCTTCAGCAGG CTGAAGTTAGTAGCTCCGCTTCCAGATCCTCAGTGGTTGGC

**Table 2.1** Primers for FP-P2A cassettes

<b>P2AHygf</b>	AGACGTGGAGGAGAACCCTGGACCCATGAAAAAGCCTGAACTC ACCGC
<b>P2AHygr</b>	CCGCTCTAGAACTAGTGGATCCCCCTATTCTTTGCCCTCGG ACGAG
<b>P2Absdf</b>	GCTGGAGACGTGGAGGAGAACCCTGGACCCATGGCCAAGCCT TTGTCTCA
<b>P2Absdr</b>	GCGGCCGCTCTAGAACTAGTGGATCCCCCTTAGCCCTCCCACA CATAACC
<b>P2ANeof</b>	GCTGGAGACGTGGAGGAGAACCCTGGACCCATGATTGAACAA GATGGATT
<b>P2ANeor</b>	GGCGGCCGCTCTAGAACTAGTGGATCCCCCTCAGAAGAACTC GTCAAGAA

<b>P2APurof</b>	GCTGGAGACGTGGAGGAGAACCCTGGACCCATGACCGAGTAC AAGCCCAC
<b>P2APuror</b>	GGCGGCCGCTCTAGAACTAGTGGATCCCCCTCAGGCACCGGG CTTGCGGG
<b>P2AMacsf</b>	GCTGGAGACGTGGAGGAGAACCCTGGACCCATGGCACCCCTGC ATGCTGCT
<b>P2AMacsr</b>	GGCGGCCGCTCTAGAACTAGTGGATCCCCCTACCCTCCTTTT CCACCTG
<b>P2AZeof</b>	GCTGGAGACGTGGAGGAGAACCCTGGACCCATGGCCAAGTTG ACCAGTGC
<b>P2AZeor</b>	GGCGGCCGCTCTAGAACTAGTGGATCCCCCTCAGTCCTGCTCC TCGGCCA

**Table 2.2** Primers for FP-P2A-selection markers

<b>gRNA9</b>	rGrGrGrUrGrGrGrGrCrUrGrArGrCrCrUrCrArGrU
<b>gRNA22</b>	rArGrGrGrCrArGrCrCrArGrGrGrArGrArGrCrCrC
<b>UBFmAGf</b>	GGGACTCCTCAGACTCTGACTCCAACGGATCCATGGTGAGCGT GATCAAGCC
<b>UBFHygr</b>	GAGAAACAAAGGTGGTTCAGTTGGGGGATCCCTATTCCTTTGCC CTCGGACG
<b>UBFKlf</b>	GGGACCTCTCCTAGACTCCTG
<b>UBFKlr</b>	GGAATGGGTCTTCCCCAAGC

**Table 2.3** Primers and gRNA's for UBF Knock-In

<b>gRNAFib</b>	rArCrUrUrCrArGrUrUrCrUrUrCrArCrCrUrUrGrG
<b>LFibPf</b>	TGAAGTTCAGCCCTGAGCGGATTGCGAG
<b>LFibPr</b>	GCCGTCCACGGCGCTGTTGG
<b>FibmChf</b>	CAAGCCCAACAGCGCCGTGGACGGCACC GCCGGCCCCGGCAT GGTGAGCAAGGGCGAGGA
<b>FibPBSRr</b>	GCAATCCGCTCAGGGCTGAACTTCATTAGCCCTCCCACACATAA C
<b>FibKlf</b>	GCCCAAAGCCCTGTTGTAGA
<b>FibKlr</b>	CTGTTGCTAGGCGACGAAGA

**Table 2.4** Primers and gRNA for FIB Knock-In

<b>gRNANop52</b>	rArArGrArArArCrGrCrArGrGrGrArGrUrGrArUrG
<b>Nop52f</b>	CTCGAGAAACAGCCAGAGCGTGACA
<b>Nop52r</b>	CCGCTGCTCCTGGGCATTAGAA
<b>HALONop52f</b>	CAGCTTTCTTGTACAAAGTGGTTGATCTAGAGGGCCCGATG GCAGAAATCGGTACTGG
<b>Nop52Hygr</b>	CTGCCTGTCCTTGGCCCGGCCACATCTATTCTTTGCCCTCG GAC
<b>HALONop52f</b>	CAGCTTTCTTGTACAAAGTGGTTGATCTAGAGGGCCCGATG GCAGAAATCGGTACTGG CTGCCTGTCCTTGGCCCGGCCACATTTAGCCCTCCCACA CATAAC
<b>Nop52Linkf</b>	GAGAAGAAGAAGAAACGCAGGGAGTATCTAGACCCAGCTTT CTTGTACAAAGTGGTTGATCTAGAGGG
<b>Nop52Linkr</b>	CTGCCTGTCCTTGGCCCGG
<b>Nop52Kif</b>	GGGATTCACCCCTTCCCTTG
<b>Nop52Klr</b>	CTCCTTGGCCCCTAACTGTG

**Table 2.5** Primers and gRNA for Nop52 Knock-In

### 2.1.12 Vectors

<b>pENTR4</b>	Gateway® Entry Vector, pUC ori, Kanr	Creating a Gateway® entry clone for recombination with a destination vector to generate an expression clone
<b>pDEST17</b>	Gateway® Destination Vector, T7 promoter, Ampr, pBR322 origin	Creating an Expression clone for the inducible expression of C-terminal tagged fusion protein (Nop52)

<b>pBluescript SK(-)</b>	Stratagene Bacterial Expression vector	Used for all “Knock-In cassettes” constructs, also the base for the UBF-mAG-P2A-Hyge repair construct
<b>pJET1.2</b>	Part of the CloneJET PCR Cloning Kit (cat. K1231; Fisher)	Used for the Nop52-Halo-P2A-Hyge/Blast. repair constructs.

**Table 2.6** Cloning vectors

### **2.1.13 DNA sequencing**

Plasmid DNA was sent to Sources BioScience for sequencing along with the appropriate primers. The resulting sequence was analysed using SnapGene viewer, Nucleotide Blast (<https://blast.ncbi.nlm.nih.gov>) as well as DNA strider 1.4.

## **2.2. Protein techniques**

### **2.2.1 Western Blot**

Cells were lysed in the appropriate volume of NuPAGE® LDS Sample Buffer to yield a concentration of  $1 \times 10^4$  cells/ $\mu$ L.  $10 \mu$ L of each lysate was loaded onto the NuPAGE® Novex® 4-12% Bis-Tris gel (Life Technologies™) for protein separation. The gel was run in NuPAGE® MES SDS Running Buffer using an XCell® SureLock Mini-cell (Invitrogen). This was conducted in accordance with the manufacturer’s instructions. Resolved proteins were transferred to Hybond™- ECL™ nitrocellulose membranes (Amersham) using XCell® II blot module (Invitrogen) and NuPAGE® transfer buffer (Invitrogen). In order to assess the efficiency of the protein transfer, membranes were stained with Ponceau S (0.1% (w/v) Ponceau S (Sigma Aldrich®), 5% acetic acid). After imaging destaining was conducted in 1xPBS. Membranes were subsequently blocked in 5% (w/v) non-fat Milk (Marvel) in PBS for 45 minutes at RT. The primary antibodies were then diluted in blocking buffer using the appropriate dilution stated in point 2.2.3. The membrane was incubated in the primary antibody solution ON at 4°C with gentle agitation. The membrane was then subjected to three 10 minute PBS washes before incubation with the appropriate secondary antibody. The horseradish

peroxidase (HRP)-conjugated antibody ( $\alpha$ -Sheep-SigmaAldrich®,  $\alpha$ -Mouse and  $\alpha$ -Rabbit both Jackson ImmunoResearch) was diluted in blocking buffer using the guidelines from Table 1.8. The membrane was subjected to a 2 hour incubation in this solution at RT. After 3 PBS washes, Western Lightning® Plus-ECL (PerkinElmer) was used to detect bound antibodies. Membranes were imaged using a CCD Syngene G-Box XT16 camera and GeneSnap software (Syngene). Quantification of signals was conducted using.

### **2.2.2 Antibodies for Western-Blot´s**

<b>Reactivity</b>	<b>Dilution</b>	<b>Source</b>	<b>Cat. number</b>
<b>Fibrillarin (rabbit)</b>	1/1000	Cell Signaling	Cat. C13C3
<b>Nop52 (sheep)</b>	1/1000	McStay lab	N/A
<b>UBF (sheep)</b>	1/1000	McStay lab	N/A
<b>Alpha-Sheep HRP (mouse)</b>	1/2000	SIGMA-Aldrich	A9452
<b>Alpha-Rabbit HRP (donkey)</b>	1/5000	Jackson ImmunoResearch	711-035-152

**Table 2.7** Antibody list for Western-Blot´s

## **2.3 Cell culture**

### **2.3.1 Growth conditions and splitting**

hTERT immortalized retinal pigment epithelial (RPE-1) cells, as well as hTERT-RPE1HKO cells were maintained in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12; Sigma Aldrich®) supplemented with 2.5mM L-glutamine (Sigma Aldrich®), 100 I.U./mL penicillin and 50µg/mL streptomycin (PS, Sigma Aldrich®), 10% (v/v) fetal bovine serum (FBS) (Serotec) and 0.25% (v/v) sodium bicarbonate (Sigma Aldrich®). During maintenance of the created cell lines, cells were passaged when 80-90% confluent. The cells were washed 2 times with PBS and detached using 0.25% Trypsin (Sigma Aldrich®) supplemented with 1mM EDTA, pH8.0

at 37°C until >95%. The trypsinisation reaction was quenched with four times the volume of growth medium. The detached cells were resuspended in media and diluted (1:5-1:10) appropriately before reseeding into new flasks.

### **2.3.2 Liquid nitrogen stocks: freezing down and thawing out**

Cells were trypsinised and spun down at 1200g for 5 minutes. Pellets were resuspended in an appropriate volume of 90% FBS, 10% Dimethyl sulfoxide (DMSO; cell culture grade- Sigma Aldrich®). Aliquots were frozen in Nunc™ cryovials in a Nalgene® Cryobox™ at -80°C overnight. The vials were transferred to liquid nitrogen tanks for long-term storage. To thaw, cells were quickly warmed and diluted in medium. Cells were centrifuged at 1200g for 5 minutes before seeding. Cells were rinsed with 1X phosphate buffered saline (PBS; Fisher) to remove any residual DMSO and fresh growth media was added.

## **2.4 CRISPR/Cas9 genome editing**

### **2.4.1 RNP preparation**

200 µM stock solutions were prepared for tracrRNA and the single crRNA's stocks, by dissolving them in Nuclease Free Duplex Buffer (Integrated DNA technologies, 11-01-03-01). Stocks can be stored at -20°C. Based on these stocks, a mixed 44 µM dilution (of each component) with Nuclease Free Duplex Buffer was prepared. A suitable amount of this dilution (at least 2 µL) was incubated at 95°C (PCR machine) for 5 minutes. After the incubation the dilution must cool down at RT for at least 15 min. A Cas9HiFi (IDT, 1081060) dilution will be prepared of 42 µM with Buffer R (Neon Transfection 10 µL Kit, Thermo Fischer Scientific MPK1096). For a single transfection 18 pmol of the gRNA dilution and 22 pmol of the Cas9 dilution are getting pre-complexed for 10 minutes at RT (0.5 µL of each dilution).

### **2.4.2 Transfection**

hTERT-RPE1 cells were transfected with CRISPR/Cas9 genome editing reagents via electroporation using the NEON Transfection System (Thermo Fisher Scientific

MPK5000). Manufacturer's instructions were followed. A single transfection solution (total volume 10  $\mu$ L) was prepared, based on BufferR, including 2.5  $\mu$ g of the repair template and 1  $\mu$ L of the RNP complex (see 2.3.1 RNP preparation). The hTERT-Rpe1 cells (10000 cells/ $\mu$ L) were electroporated with the Neon System, conditions of 1100 V, 20 ms, 2 Pulses. Resulting in a transfection efficiency of around 70-80%, however it should be noted that the survival rate of the cells depends on the target gene. After the electroporation the cells should rest for 2-5 minutes and can be seeded out in Penicillin-Streptomycin (PS) free media. After 24-48 hours selection can be performed. Used concentration for the Knock-In's were 25  $\mu$ g/ml for Hygromycin B (Invitrogen), and 5  $\mu$ g/ml Blasticidin (Melford). The selection runs for 2-3 weeks depending on the cell growth.

## **2.5 Imaging**

### **2.5.1 Imaging/live cell imaging**

Images were captured and merged using a Hamamatsu Orca Flash 4.0 V2 camera and LASX 2.0 software (Leica) with a 63x Plan Apo- chromat objective mounted on a Leica DMI8 imaging microscope illuminated by a Lumencor Spectra X Fluorescence Light Source. Typically, Z-stacks from the bottom to the top of a cell were taken. The number and distance between the stacks were system optimized. Image stacks were deconvolved by a blind method using auto-generated point spread functions (PSFs) with 10 iterations in LASX 3.3 software containing the 3D deconvolution module (Leica). For live cell imaging the chamber was preheated at 37°C for at least an hour before imaging, carbon dioxide (5-7%) atmosphere in the imaging chamber can be produced.

### **2.5.2 Fixed cells**

Cells were grown on glass cover slips. They were fixed by incubation for 10 min with 4% PFA in PBS. Cells were washed 3 times with PBS. Permeabilization buffer (0.5% Tritpon X-100 in PBS) was added and the cells were incubated for 5 minutes. Further washes with PBS (3 times) followed. The cover slips were mounted with VectorShield (Vector Laboratories), this included DAPI. The fixed DAPI stained cell lines were imaged to observe their behaviour during specific stages of the cell cycle (meta, anaphase).

### **2.5.3 Live cell imaging and Halo staining**

For live cell imaging cells were grown in complete F-12 media without phenol red in  $\mu$ -Dish 35-mm microscopy dishes (Ibidi, 81156). Halo tagged cell lines, were stained with TMR Halo dye (Promega, G8251) 0.5  $\mu$ M in phenol red Free F-12 media for 15 minutes at 37°C (Incubator). After the Incubation, the cells were washed 3 times with phenol red free F-12 media. The cells were incubated for another 45 minutes in phenol red free F-12 media to reduce background signal. After this final incubation, media was changed and (live cell) imaging can start.

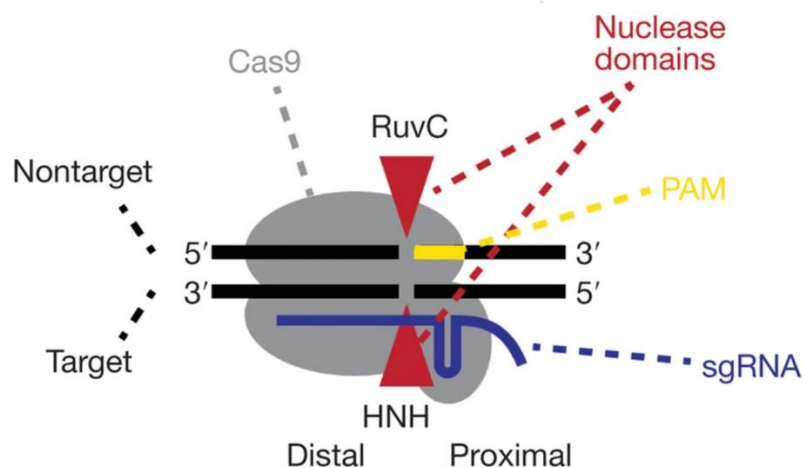
## **CHAPTER III**

### **3. Establishing a strategy for C-terminal tagging of hUBF by CRISPR/Cas9**

The targeted creation of DNA DSB at a desired site, in combination with DNA repair mechanism like NHEJ or HR, is the basis of any controlled genetic manipulation. The key step is to create a DNA DSB at the right position. Nuclease systems like ZFNs or TALENS were established systems for creating these DBS in the past. Their targeting principle is based on a protein-DNA interaction and thus they are less accurate in terms of the cutting position. The CRISPR/Cas9 system is, on the other hand, can cut with base precision because the targeting is based on RNA-DNA interactions. In general, it can be stated that the CRISPR/Cas9 system, compared to ZFNs and TALENS is undeniably simpler, cheaper, and more efficient. Therefore, it was the system of choice for my thesis.

#### **3.1. Crispr/Cas9 as a tool for FP tagging relevant nucleolar key proteins in hTERT-RPE1 cells.**

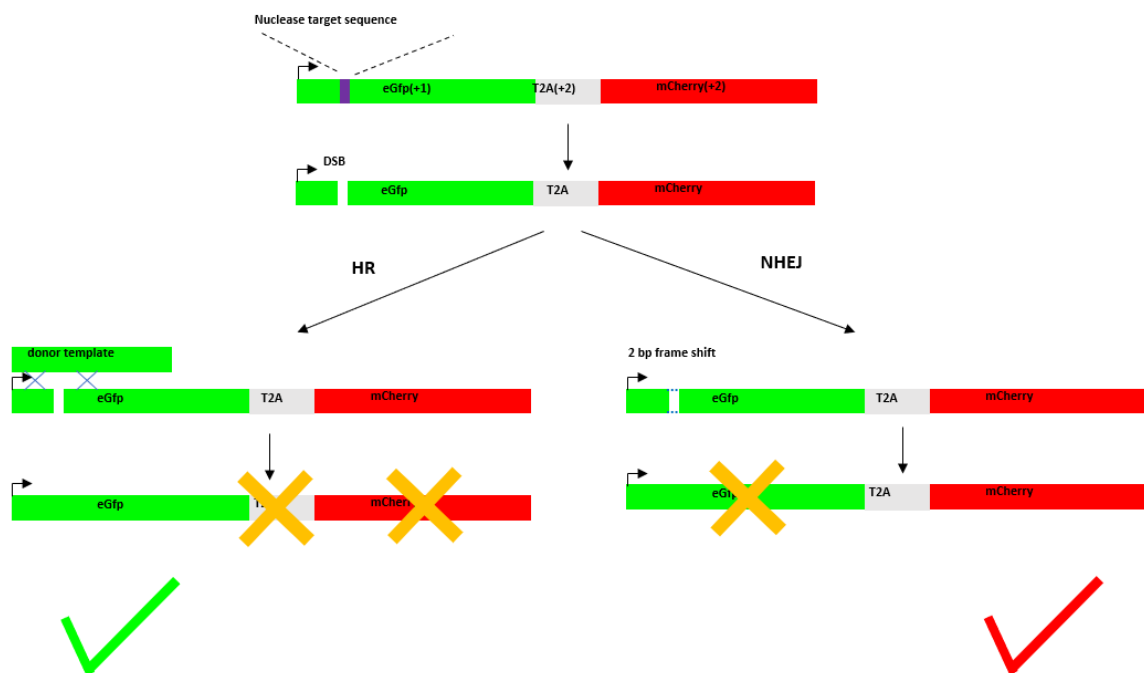
In February 2013 two ground-breaking papers were published in Science, the development of clustered regularly interspaced short palindromic repeats (Crispr)/Crispr associated 9(Cas9) technology has revolutionized genome editing in virtually all (model) organisms. The system was originally introduced as a tool for genome editing in mammalian cells <sup>171,172</sup>



**Fig.3.1** An illustration of the CRISPR/Cas9 complex and its interaction with substrate DNA. (a) Schematic of Cas9's interaction with target DNA. Cas9 (grey) complexed with sgRNA (dark blue) binds to DNA (black) comprising target and nontarget strands. Cas9- PAM interactions occur on the nontarget

strand; sgRNA-DNA annealing occurs on the target strand. RuvC (His840) and HNH (Asp10) nuclease domains cut the nontarget and target strands, respectively (red triangles) <sup>173</sup>.

These two studies demonstrated that Cas9 nucleases can be combined with short RNAs to induce precise cleavage at endogenous genomic loci in human and mouse cells. Mali et al., used a human codon optimised Cas9 with a C-terminal SV40 NLS and optimised the chimeric RNA approach by expressing crRNA-tracrRNA fusion transcripts for the human U6 polymerase III promoter following validation of this approach in vitro <sup>174</sup>. They tested the efficiency of the system in stimulating HR and compared it to that of TALENs. Both systems successfully stimulated HR as analysed via a traffic light reporter (TLR) assay (see Fig. 3.2).



**Fig. 3.2** Illustration of the traffic light reporter assay. The arrow represents the promoter and initial eGFP start codon. Reading frames relative to the initial eGFP start codon are indicated in parentheses. Schematic illustration of different engineered outcomes after the specific double-strand break (DSB). If the break is repaired through the HDR pathway, the full eGFP sequence will be reconstituted, the eGFP sequence is getting transcribed and translated in the right frame, so that a STOP codon will take place before the T2A sequence. The cells will show a green fluorescent signal. If the created break undergoes

NHEJ resulting in a frameshift, eGFP will be translated out of frame and the T2A and mCherry sequence is getting translated (in the right frame), the cells will show a red fluorescent signal.

The assay is based on two different fluorescent protein sequences, which are separated by a T2A element. The method itself was developed by Creto et al.. The first sequence, which is usually eGfp, is including an engineered DNA frameshift, which is part or close to the cutting site of the used nuclease system. The second fluorescent sequence is usually encoding a red FP signal, which can be Tomato or mCherry. The frameshift is designed in a way that it can be corrected in two different variants, depending on the repair mechanism. Homologous recombination (HR) would lead to a correction of the eGfp sequence (including a stop codon), based on a suitable repair template, so the cells would show a green fluorescent signal. However, it does not result in a translation of the T2A sequence, which is followed by mCherry. Non-homologous end joining (NHEJ) can correct the frameshift in a way, that just a correct translation of the T2A-mCherry sequence occurs. The cells would show a red fluorescent signal.<sup>175</sup> However, the CRISPR based system was more rapid with GFP positive cells present after twenty hours relative to forty hours for TALENs. They also demonstrated that CRISPR mediated HR stimulation can be used to integrate foreign DNA at endogenous loci<sup>171</sup>.

While cleavage at sites with perfect complementarity to the protospacer domain of the gRNA typically occurs with highest efficiency, cleavage at sites having one to several bases of mismatch with the gRNA can occur. These undesired cleavage events are called “off-target effects”. As a general rule it can be concluded, a high amount of Cas9 and gRNAs is resulting in a higher “off target” activity. Therefore, a delivery method based on a “fast on and fast off” principle is the best solution here. Pre-complexed Cas9 protein and gRNA’s, so called ribonucleoprotein complexes (RNPs) fulfil these requirements, quite in contrast to gRNA and Cas9 encoding DNA or RNA constructs. The RNP delivery is also a method to avoid intra-cellular innate immune response that is highly sensitive in sensing the delivery of foreign nucleic acids, that would be for example the case in human primary cells. A second also important point to reduce this side effect is to increase the Cas9 binding specificity. A non-catalytic domain, called REC3, within the Cas9 recognizes target complementarity and governs the HNH nuclease activity. A targeted induced mutation in this REC3 domain led to a substantial

improvement of this activity regulation <sup>176</sup>. The Cas9 proteins based on this mutation are commonly known as high fidelity Cas9 (HiFi Cas9) proteins. The use of this modified HiFiCas9 protein in CRISPR related in vivo experiments showed a dramatically reduced off target activity <sup>177</sup>.

The programmable RNA guided endonuclease *Streptococcus pyogenes* CAS9 induces DNA double strand breaks (DSBs) at target sites in the genome. These DSB then trigger DNA repair mechanisms, which can be classified into two categories: homology dependent and homology-independent repair <sup>178</sup> Homology-dependent DNA repair occurs mainly through homologous recombination (HR), whereas homology-independent DNA repair (HIDR) occurs through nonhomologous end joining (NHEJ). These DNA repair systems have both advantages and disadvantages. Although HR can achieve precise repair, the repair rate is relatively low because it occurs in limited periods during the cell cycle namely the late S to G2 phases, which is approximately one third of the period of the overall cell cycle. On the other hand, the rate of HIDR is high, although it is error prone. One of the reasons for this high rate is that NHEJ can be performed during all phases of the cell cycle. Therefore, DNA repair strategies should be selected in accordance with the actual purpose of genome editing, which could be a deletion or inactivation of a gene (Knock-Out) or the integration of a sequence at a specific position in the gene (Knock-In).

hTERT-RPE1, also often referred as RPE or RPE1, is a cell line that was derived from human normal retinal pigment epithelial (RPE) cells from a normal human female, which were immortalized by exogenous expression of human telomerase reverse transcriptase (hTERT) <sup>179,180</sup>. The hTERT-RPE1 cell line was originally established by immortalizing the RPE-340 cell line by transfection of the pGRN145 plasmid (ATCC MBA-141), which contains not only the hTERT gene but also the Hygromycin B and Puromycin resistance genes, followed by selection in presence of Hygromycin B. Therefore, conventional transfection protocols, for CRISPR/Cas9 or other purposes for involving transient puromycin selection is not applicable to hTERT-RPE1 cells. In the context of this work, I will use the term hTERT-RPE1 for the used cell line.

Distinctive features of hTERT-RPE1 cells are their high potential for ciliogenesis and normal karyotype. Because hTERT-RPE1 cells efficiently form primary cilia under serum starved conditions, these cells have been widely used for studies of ciliogenesis

and ciliary function. A useful characteristic of this cell line are the intact checkpoints and the normal karyotype. Furthermore, they are also well characterised in the aspect of their NORs <sup>51</sup>. Many cultured cell lines, such as HeLa and HEK293T, are abnormal with respect to the number of chromosomes. Sequencing of HeLa genome demonstrated a high level of aneuploidy <sup>181</sup> and this can cause difficulties in genome editing because the copy number of the target genes can vary. This discontinuity in the number of chromosomes should also lead to doubts about suitability as a model system. However, HeLa cells are still a very popular and established model system for cell culture-based experiments. Therefore, immortalized normal diploid cell lines, such as hTERT-RPE1 provide to be a better and all above a more reliable model system for normal cellular functions, such as those being investigated through the course of my thesis work.

HR activity varies among different cell lines, and in general it is lower in normal cell lines than in cancer cell lines, such as HeLa or HEK293T. Because hTERT-RPE1 cell retain characteristics of normal primary cells, HR activity is believed to be low in these cells. In fact, Miyamoto et al, reported that the establishment of biallelic KIF2A knockout (KO) lines from hTERT-RPE1 cell was difficult, but not impossible <sup>2</sup>.

Another popular Knock-In method is the use of single stranded ultra-long oligomers as DNA (ssODN) repair templates. A mechanism proposed by Richardson et al, assumes that the single stranded DNA can already anneal to the genomic DNA during the cutting process by the CRISPR/Cas9 complex. It is assumed that HR activity is increasing because of this mechanism <sup>173</sup>. Haapaniemi et al, showed that this method is working in hTERT-RPE1 cells, by the negative regulation of the p53 response <sup>182</sup>. They targeted a stable integrated plasmid encoding a non-fluorescent GFP mutation and co-transfected MDM2 (a negative regulator of p53) together with an ssODN, as a repair template. A small proportion of the cells showed a green fluorescent signal, because the ssODN got integrated at the right position and corrected the GFP sequence <sup>182</sup>. The disadvantage or the limiting factor of this method is currently the short length of the ssODN, which can be synthesized. Longer ssODNs (around 1000nt) can be synthesised, but they are very expensive and the resulting secondary structure of those ssODN could influence the HR mechanism.

### **3.2 Aims of this chapter.**

The aim of the work described in this chapter is to develop a reliable and efficient strategy, based on CRISPR/Cas9, to FP-tag endogenous genes encoding key nucleolar proteins in hTERT-RPE1 cells. This would offer the possibility of visualising the dynamic behaviour of nucleoli in a non-transformed human cell line, certain in the knowledge that the FP-tagged proteins are fully functional and expressed at normal levels.

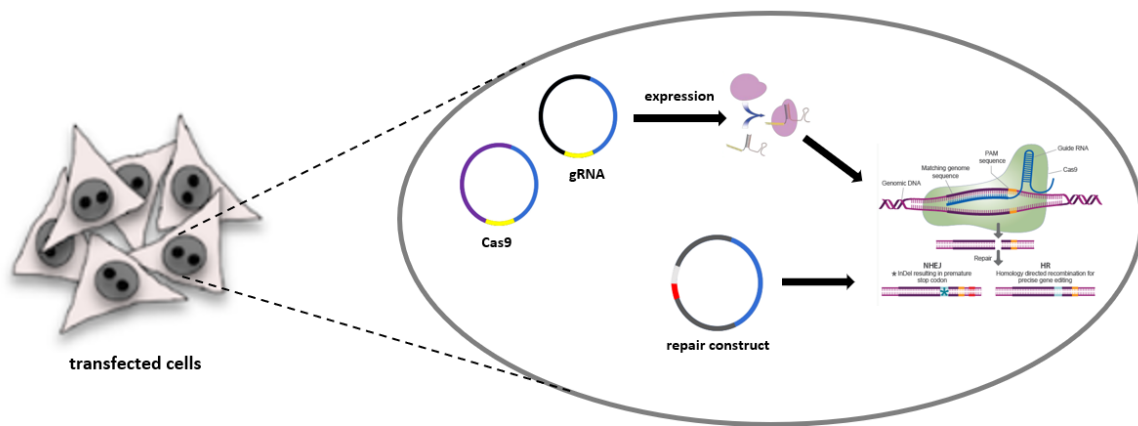
### **3.3 Results**

#### **3.3.1 Developing a Crispr/Cas9 Knock-In technique for hTERT-RPE1 cells.**

As previously described, the HR activity in hTERT-RPE1 cells and primary cells is low compared to cancer cells. Also, the transfection efficiency with conventional DNA transfection methods for plasmid encoded constructs such as Calcium- phosphate or Lipofectamine is very inefficient in this cell line. In addition, an enrichment of positive transfected cells based on puromycin selection is also not possible because of the immortalization construct (pGRN145)<sup>179</sup>.

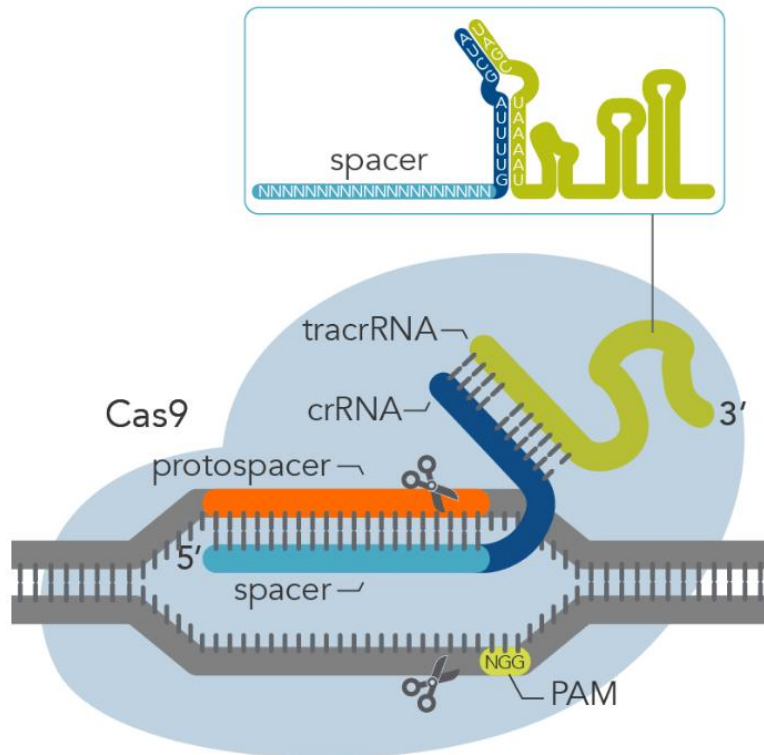
Therefore, it was necessary to find a way to increase the transfection efficiency, and a way to distinguish appropriately targeted cells from those in which the targeting construct was randomly integrated.

Most of the Crispr/Cas9 based Knock-In in the past were based on DNA constructs, encoding the Cas9 protein and a specific gRNA sequence, which is targeting a gene of interest. The third construct has the function of a DNA repair template, so it is including homology arms as a recognition sequence for HR and the sequence, which should be added to the gene locus. The positive co-transfected cells need time until the expression of Crispr/Cas9 construct is starting, during this time the concentration of DNA repair template, is decreasing. Furthermore, it must be also considered that transiently expression of Crispr/Cas9 constructs is inducing DNA damage by unspecific binding (off targeting effect), as long the constructs are transiently active in the cells (see Fig. 3.3).



**Fig. 3.3.** Schematic illustration of Crispr/Cas9 genome editing, based on plasmid encoded constructs. After a successful transfection, the expression constructs are getting transiently expressed, the CRISPR/Cas9 complex is forming and getting active. The repair construct is necessary for the HR pathway of repair at the designed genomic position, whereas NHEJ is independent from the presence of a repair construct.

A solution for this problem, is the transfection of fully assembled ribonucleoprotein complexes (RNPs). This complex can be divided into three parts: A crRNA which binds to a specific target sequence (protospacer, see Fig. 3.4.), but also anneals to a tracrRNA (green, see Fig. 3.4.). These two components form a gRNA, which can be assembled with a recombinant Cas9 (NLS) to an active CRISPR/Cas9 complex.



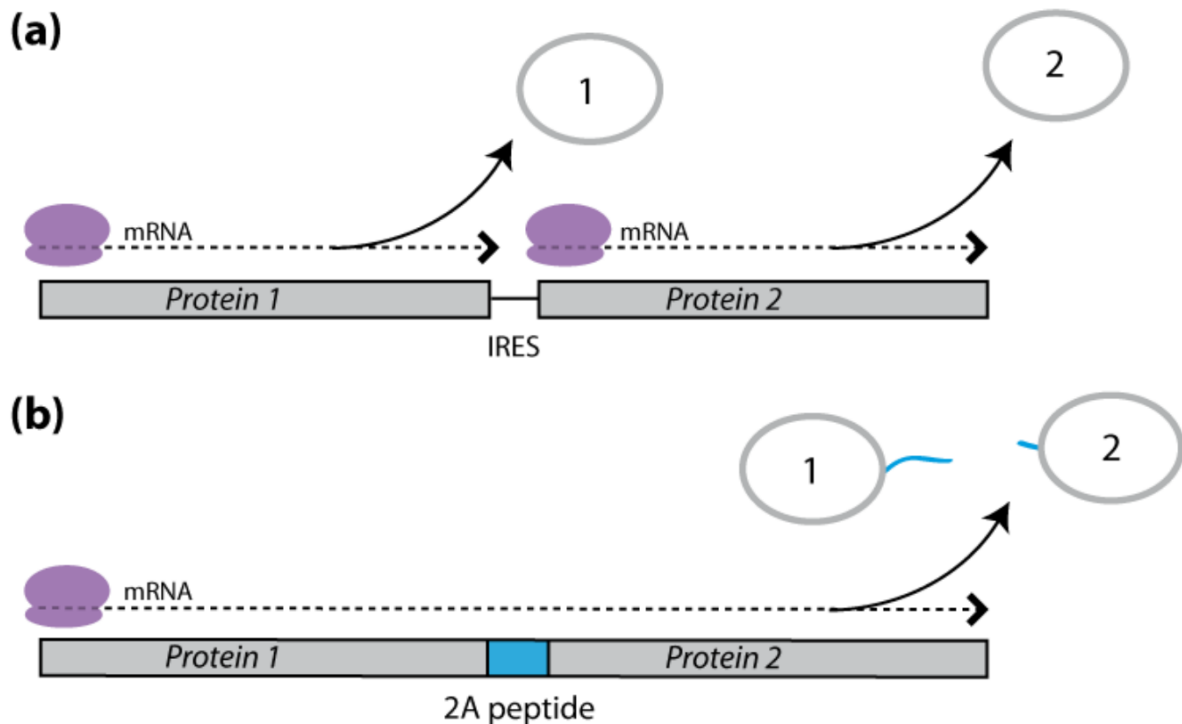
**Fig. 3.4:** The IDT based 2-part gRNA format. crRNA:tracrRNA complex uses a optimized crRNA/tracrRNA sequences that hybridize and then form a gRNA. Together with Cas9 endonuclease a fully functional RNP complex is formed. The cleavage site is specified by the spacer element of the crRNA (light blue). The crRNA spacer element recognizes 19 or 20 nt on the strand opposite from the PAM (NGG) site. The PAM site must be presented immediately downstream of the protospacer element (orange) for cleavage to occur. tracrRNA is illustrated in green. (Figure adapted from IDT: <https://eu.idtdna.com/pages/products/crispr-genome-editing/alt-r-crispr-cas9-system>.)

These components can be produced in a lab by protein and RNA expression of suitable constructs. But they are also commercially available, Integrated-DNA-Technologies (IDT) can be mentioned here as an example. The complexes are already active as soon they are assembled and so they can cut DNA in the cells (*in vivo*) immediately after the transfection but they can be also used for DNA *in vitro* digests. They are also short living (*in vivo*) compared to plasmids encoding Crispr/Cas9 complexes. This reduces the problem of DNA damage and off-targeting effects. There is also an optimized Cas9 nuclease available, engineered for higher fidelity (HiFi)<sup>183</sup>. Throughout the course of this work, I have used the CRISPR/Cas9 (HiFi) system from IDT

Electroporation is the optimum method to co-transfect DNA and RNPs into hTERT-RPE1 cells <sup>135</sup>. This method allows to co-transfect RNP complexes together with DNA (repair) constructs. It also shows a higher transfection efficiency and a better cell survival compared to well-practiced methods like Calcium-precipitation or Lipofectamine.

We looked at the problem of low transfection efficiency, but we still have the problem of the low HR activity in hTERT-RPE1 cells. This activity is fundamental for a successful Knock-In, and just active during the late S to G2 phase of the cell cycle. Selection markers are offering the possibility to kill cells, in which the DNA DSB has been repaired by NHEJ, but also in situations where the targeting construct has been randomly integrated. Creating a situation where expression of the selection marker is contingent on accurate integration by HR offers a solution. This would lead to a combined expression of our tagged (target) gene combined with a selection marker system. With an active selection it should be possible to enrich positive tagged cells and kill all untagged cells.

Two genetic tools would allow us to create a connection between a FP and a selection marker sequence, the internal ribosomal entry site (IRES) and the 2A self-cleaving peptides (ribosomal skipping) (see Fig. 3.5).



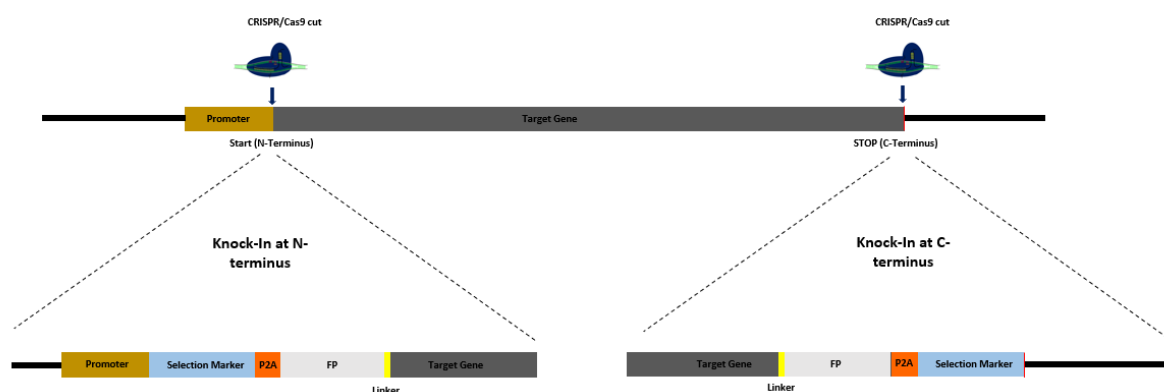
**Fig. 3.5** IRES and Ribosomal Skipping as genetic tools. **A.** IRES-mediated translation: An internal ribosome entry site (IRES) sequence is included between the coding sequences of Protein 1 and Protein 2 to drive a cap-independent, internal initiation of Protein 2 translation, in parallel to Protein 1 cap-dependent translation initiated at the polycistron transcript 5' end. **B.** 2A peptide-mediated translation: A viral 2A peptide sequence is included between the coding sequences of Protein 1 and Protein 2 to induce 'ribosomal skipping' during translation leading to the co-translational release of the two proteins. Translating ribosomes are indicated in purple. <sup>184</sup>

The IRES sequence was originally discovered in 1988 in the RNA genomes of poliovirus and encephalomyocarditis viruses by Sonenberg et al, this sequence allows IRES-mediated translation. Two coding sequences for proteins can be separated by an IRES sequence. The first protein coding sequence is a cap dependent translation, RNA Pol II, initiated at the polycistronic transcript 5' end. The IRES sequence, which is separating the two coding regions for the single genes, drives a cap independent internal initiation of RNA Pol II translation of the protein which is located downstream <sup>185</sup>. IRES separated genes can be co-expressed under the control of the same promoter, by mimicking a polycistronic mRNA. However, the expression rate of the second gene, downstream of the IRES is usually lower.

The second genetic tool would be a 2A self-cleaving peptide sequence, this sequence can be also used, like the IRES sequence, as a tool to separate and co-express two protein coding sequences. Originally the 2A self-cleaving peptides were discovered in

foot-and-mouth disease virus, and named after a gene of a peptidase (2Apro) of the virus<sup>184</sup>, whereas the P2A peptide was originally discovered in the Porcine teschovirus 1 (PTV1). The 2A sequence led to the co-translational division of two proteins within a single open reading frame. The exact mechanism is still unknown, but it is assumed that the 2A sequence inhibits the formation of a regular peptide bond on the ribosome and as a consequence the peptides separate during translation at the glycyl-prolyl peptide bond formation, rather than true proteolytic cleavage<sup>186,187</sup>. The cleavage always takes place before the C-terminal proline of the 2A peptide, therefore the second (C-terminal) peptide has a proline as the N-terminus and the first (N-terminal) peptide has the remaining 2A sequence as the C-terminus<sup>188</sup>. The most outstanding advantage of ribosomal skipping is the combined transcription and translation of the two (2A) connected genes. This can be just done the correct way if the two sequences are combined in correct way (open reading frame (ORF) is correct). In case a frameshift would occur for the upstream gene, a separation would not be possible. Therefore, ribosomal skipping is also an *in vitro* control mechanism, because selection is entirely dependent on appropriate translation of the targeted fusion protein

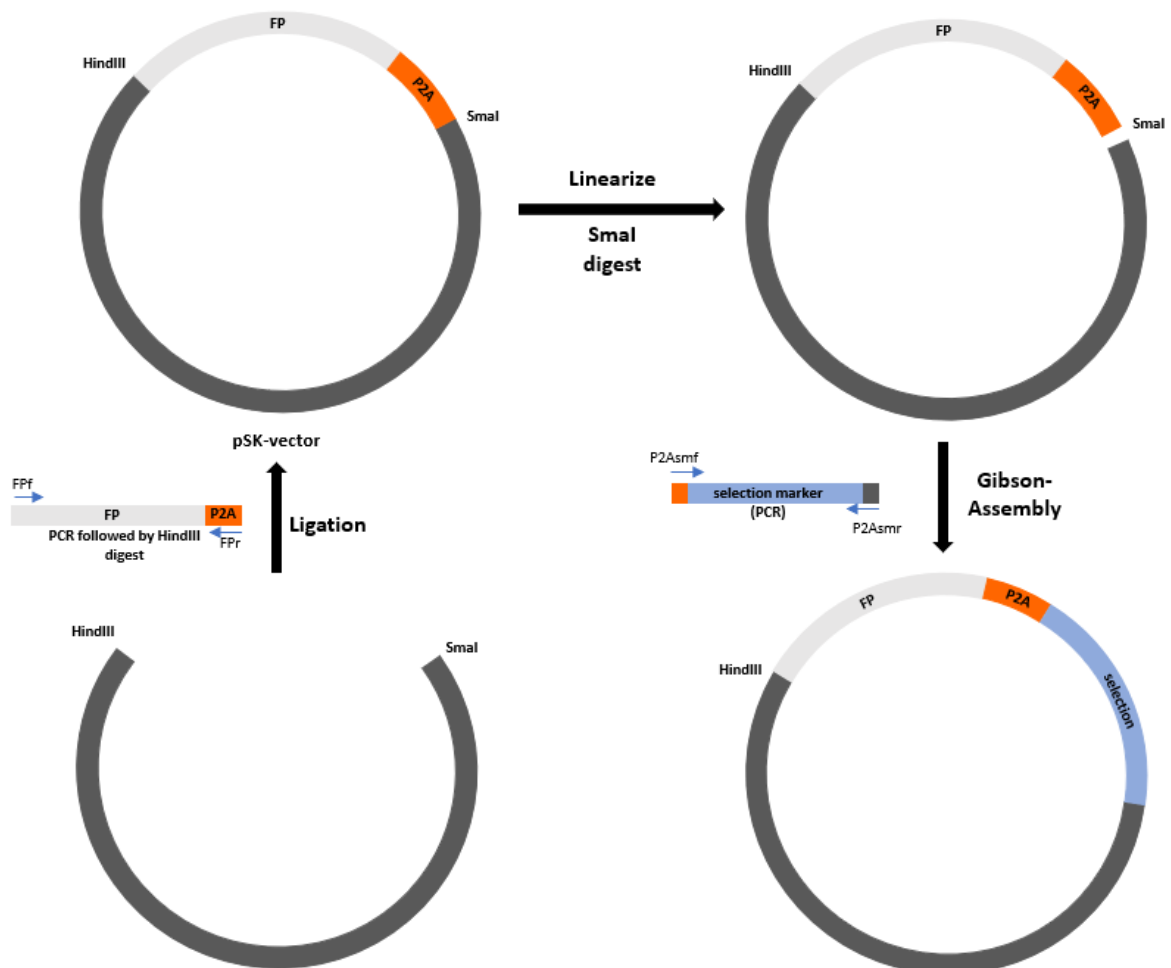
The 2A sequences are a more suitable tool for co-expressing protein coding sequences. The efficiency of the second (C-Terminal) protein expression rate is higher compared to IRES. Especially the P2A sequence shows a really high rate of the downstream protein coding sequence.<sup>188</sup>.



**Fig. 3.6.** Genome editing principle of N-terminal and C-terminal tagging based on developed Knock-In constructs, combined with a selection marker system by ribosomal skipping (P2A). In the case of N-terminal tagging a selection marker sequence would be combined by a P2A sequence with a (fluorescent) tag (FP). This tag would be connected by a linker sequence (yellow) to the N-terminus of the gene of interest. For C-terminal tagging, a linker sequence would connect the end of the gene of interest with a FP followed by a P2A-selection marker sequence.

The specific advantages of 2A sequences are not only transcription but also in-frame translation of all “fused” genes is required. Gene tagging with fluorescent proteins (FP) or other tags can be done at the N-terminus or at the C-terminus of a target gene. (see Fig. 3.6). Tagging at the N-terminus can influence the promoter elements and lead to a lower gene expression if it’s combined with a selection method, like IRES or ribosomal skipping. C-terminal tagging is avoiding this problem and therefore our preferred method. This method will be explained in more detail in the next section.

### 3.3.2 Developing a flexible tagging system.



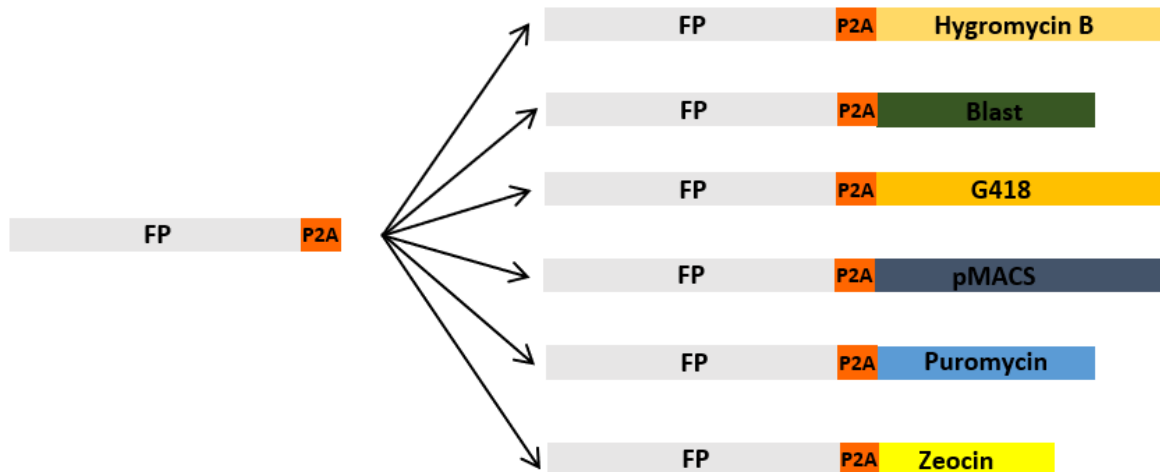
**Fig. 3.7** Schematic cloning strategy for creating different Knock-In cassettes. A psk(-) vector, linearized with HindIII and SmaI. Ligation with a HindIII digested PCR product, including a fluorescent protein sequence (without stop codon) followed by a P2A sequence at the C-terminus, SmaI cutting site stays

in tact. pSk(-)FP\_P2A construct can be linearized by SmaI, selection marker of interest can be integrated by Gibson-Assembly.

I developed a cloning strategy, which allows to combine different fluorescent proteins (FP) with any selection marker system of choice. The cloning vector pBluescript SK(-) (STRATAGENE) has, HindIII (sticky end) and SmaI (blunt end) restriction sites within the multiple cloning site (MCS). PCR's were performed based on a variety of different FP (mAG, mCherry, eBfp and Halo) sequences as templates combined with suitable primers. The FPP2Af/r primers would be the primer pairs for the corresponding constructs ( mAGP2Af/r, mChP2Af/r (same for eBfp), mHaloP2Af/r). The stop codon of the FPs gets removed and a direct connection to the P2A sequence is established. The amplified PCR products were digested with HindIII, and ligated into the prepared pSK(-) backbone (see Fig. 3.7).

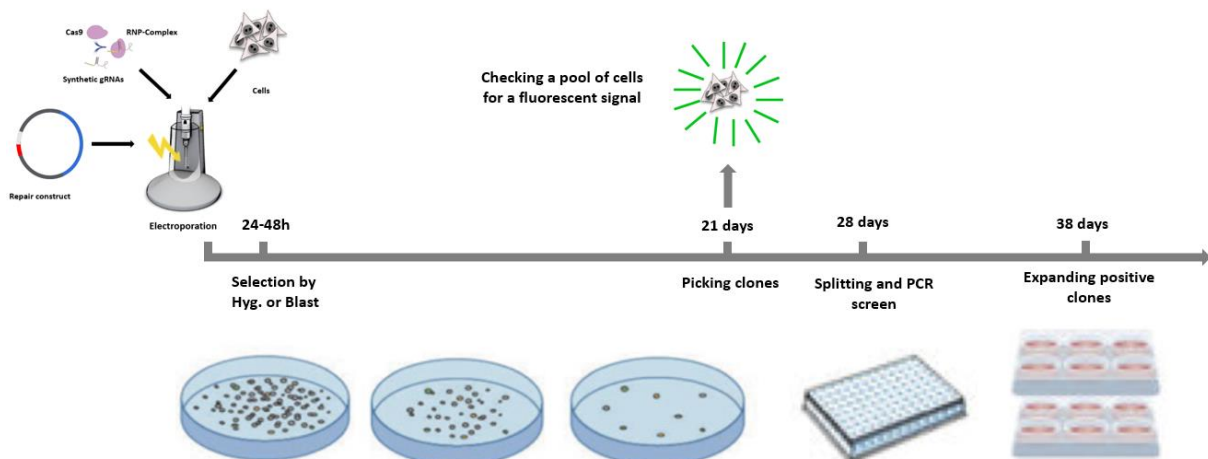
The SmaI cutting site stays intact and is used to linearize the new construct, to add in a selection marker system by Gibson-Assembly. Therefore, new PCR's with 25nt overhangs were performed, based on suitable primer pairs and templates for Blastidicin, Hygromycin B, Puromycin, Neomycin (G418), Zeocin and a selection marker system called pMACS (Miltenyi Biotech) (see Fig. 3.8). The corresponding primers are P2Asmf/r (see Fig. 3.7.) (P2ABlastf/r, P2AHygf/r, P2APurof/r, P2AZeof/r, P2ANeof/r, and P2ApMacsf/r). This cloning strategy offers a high degree of flexibility, each FP can be combined with any a selection marker.

For the Knock-In cell lines, FP proteins like monomeric Azami Green (mAG) , and a mutation from DsRed called mCherry and HALO have been used, combined with selection marker systems such as Hygromycin B and Blastidicin.



**Fig. 3.8.** Illustration of possible cassettes which can be used for CRISPR/Cas9 mediated Knock-In's. The FP's would be mAG, mCherry, eBfp and Halo. The selection markers are Blasticidin, Hygromycin B, Neomycin (G418), pMACS system (Milenty), Puromycin and Zeocin.

These Knock In cassettes can be used to generate a targeting construct for any gene of choice by using Gibson Assembly combined with a suitable PCR strategy. This repair template, together with suitable RNP's can be co-transfected into the cells by Electroporation. After some time, selection can be started, and the cells can grow over a couple of weeks (2-3 weeks) to enrich successfully tagged clones and kill of untagged cells (see Fig. 3.9). After this time, clones can be picked, and verification can be performed by checking the location of a fluorescent signal in the cells, by a PCR assay with primer pairs annealing close to the targeting position and/or by Western-Blot with target protein specific antibodies.



**Fig. 3.9** Schematic illustration of the gene tagging procedure by single steps. RNP complex is getting produced by annealing the specific gRNA (crRNA/tracrRNA) with the Cas9 protein. The formed complex can be co-transfected by electroporation with a repair template. The RNP complex is active and starts to cut the genomic DNA at the designed position immediately after the transfection. The repair template gets integrated at the right position by HR. Selection can be started shortly after the transfection, as soon the cells are no longer stressed (24 h-48h). After 2-3 weeks of selection the cells can be screen and verified.

### **3.4 FP tagging of tagging of Nucleolar bookmarker proteins by CRISPR/Cas9**

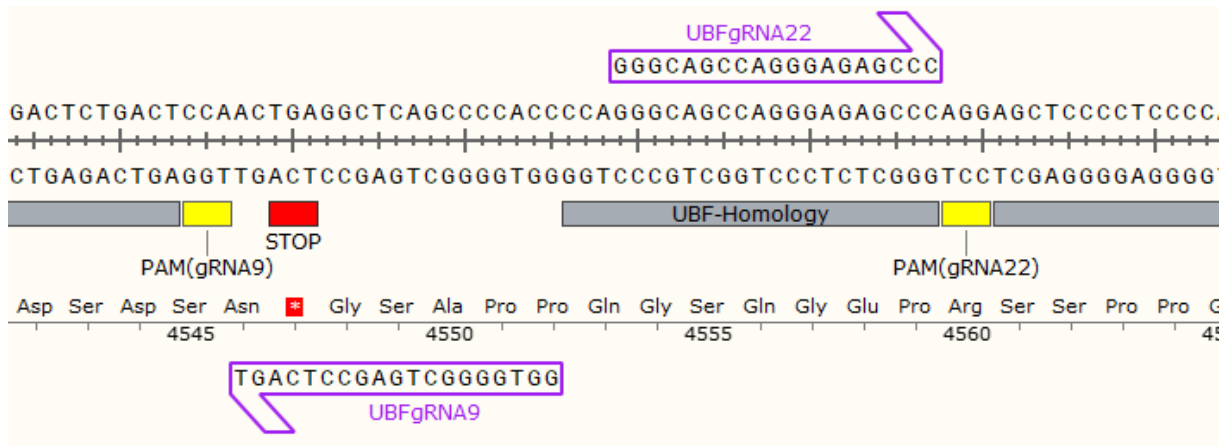
For all Knock-In's, which are part of this thesis a special hTERT-RPE1 cell line was used. In this cell line the Hygromycin B gene (encoded by the hTERT construct, pGRN145) is knocked out. This Knock-Out was created by Dr. Hazel Mangan and kindly provided for my work. In this thesis the cell line will be referred as hTERT-RPE1HKO<sup>135</sup>, this allowed us to use Hygromycin B as a sections marker and also enabled double tagging when combined with Blast or Neomycin.

### **3.5 Endogenous hUBF tagging with monomeric Azami Green (mAG)**

Because of former C-terminal tagging experiments in a cancer cell line based on UBF-eGFP, it was clear that homozygous tagging of UBF is possible<sup>189</sup>. Based on this, UBF tagging in hTERT-RPE1 cells offers a suitable test for the new developed tagging strategy.

The gene encoding human upstream binding factor (hUBF) is located on human chromosome 17. Human UBF is a nucleolar specific HMG-box transcription factor. It exists in two different splice isoforms of 97 kDA (hUBF1) and 94 kDA (hUBF2). It binds extensively over rDNA and therefore can be seen as a FC marker protein<sup>25</sup>.

C-terminal tagging of hUBF by Crispr/Cas9 had already been performed by Dr. Chelly van Vuuren in the cancer cell line HCT116 (Chelly van Vuuren, 2017, Thesis), in the McStay Lab. This tag offers the possibility to observe the behaviour of UBF during the cell cycle as its role as a FC marker by live cell imaging. The repair template (psk-UBF-eGfp) as well as other plasmids encoding two different gRNA's (called gRNA9 and gRNA22) (see Fig. 3.10) to reduce a possible off targeting effect and the Cas9/NickaseD10A protein were co-transfected by Calcium Phosphate precipitation.



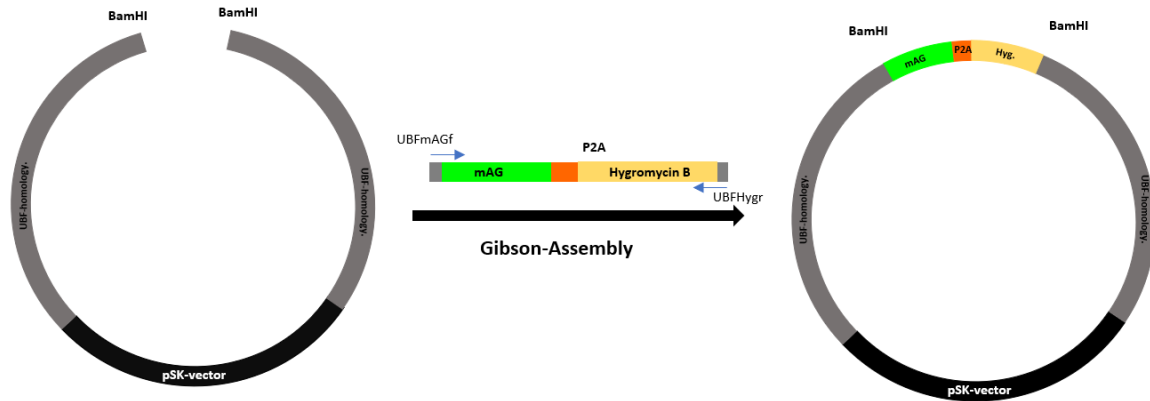
**Fig. 3.10** DNA sequence at the C-terminus of hUBF and targeting positions of two gRNAs (9 and 22). The STOP codon is labelled in red, the PAM sites in yellow. Both gRNA's 9 and 22 are indicated in purple.

Based on this strategy she was able to create a homozygous Knock-In in a cancer cell line (HCT116). As a lesson from this experiment, it can be concluded that hUBF with its role as an essential protein for the cell growth is fully functional as a homozygous C-terminal GFP Knock-In. As part of her thesis work Chelly had shown that the same strategy was unsuccessful in hTERT-RPE1 cells.

As part of my work, same procedure was repeated in hTERT-Rpe1HKO cells with a higher number of clones (120) for screening, but again it was unsuccessful. Therefore, it was necessary to re-think about the tagging strategy for the hTERT-Rpe1HKO cell line. This highlights the low levels of HR in these cells and suggests that it is necessary to combine successful integration with a selection method, based on (P)2A ribosomal skipping.

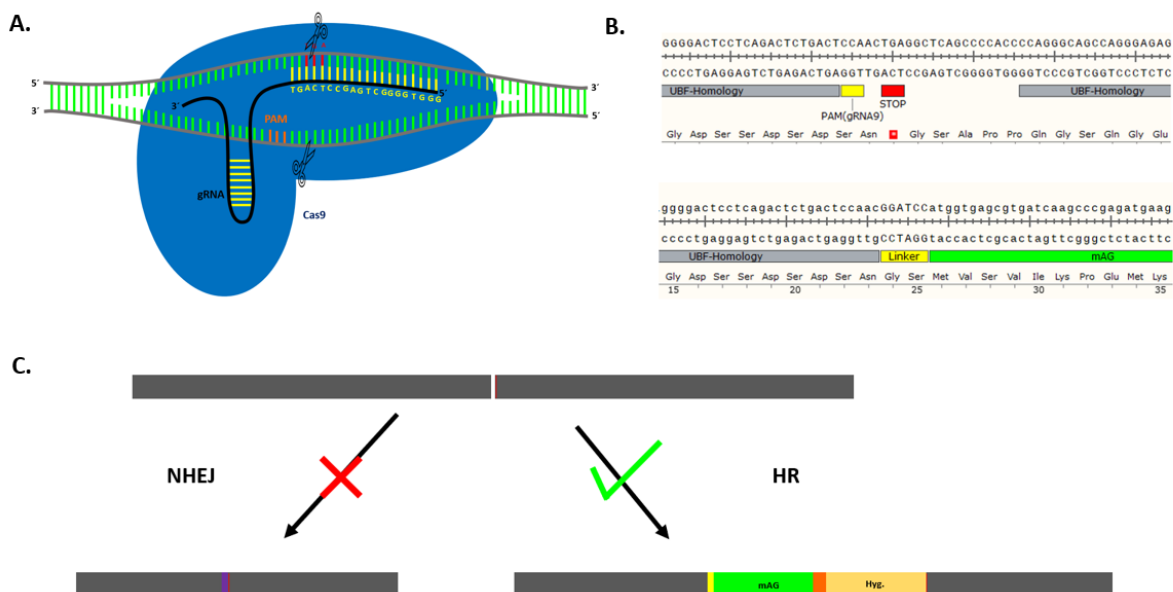
We also decided to combine a more efficient transfection technique (Electroporation) with a single RNP, based on gRNA9. A repair template was designed, which includes a mAG-P2A-Hyge cassette at the C-terminal position of UBF, flanked by UBF homology arms of around 2.5 kb.

The original UBF repair template, which was designed by Dr. Chelly van Vuuren, included eGFP cassette, flanked by BamHI cutting sites. A BamHI digest was performed and the mAG-P2A-Hyge. cassette was amplified by PCR (primer UBFmAGf and UBFHygr) to generate a product including 25 bps overhangs for the Gibson-Assembly (see Fig. 3.11).



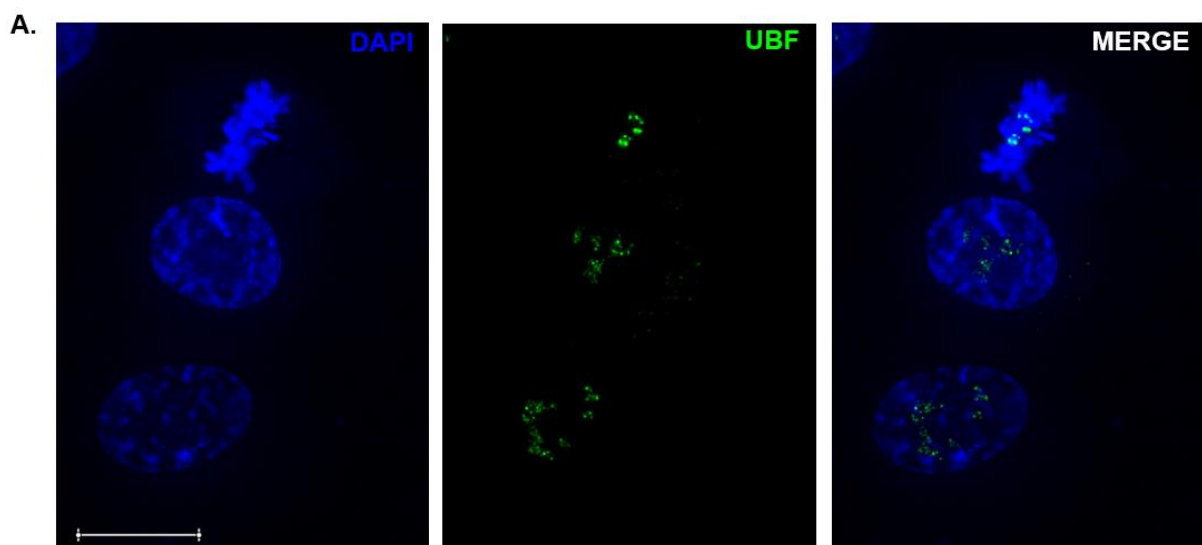
**Fig. 3.11** Illustration of the cloning strategy to create a UBF repair template, based on the original plasmid created by Dr. Chelly van Vuuren. A BamHI digest was performed to release the eGFP cassette, after that the mAG\_P2A\_Hyge cassette got integrated at this position by Gibson Assembly.

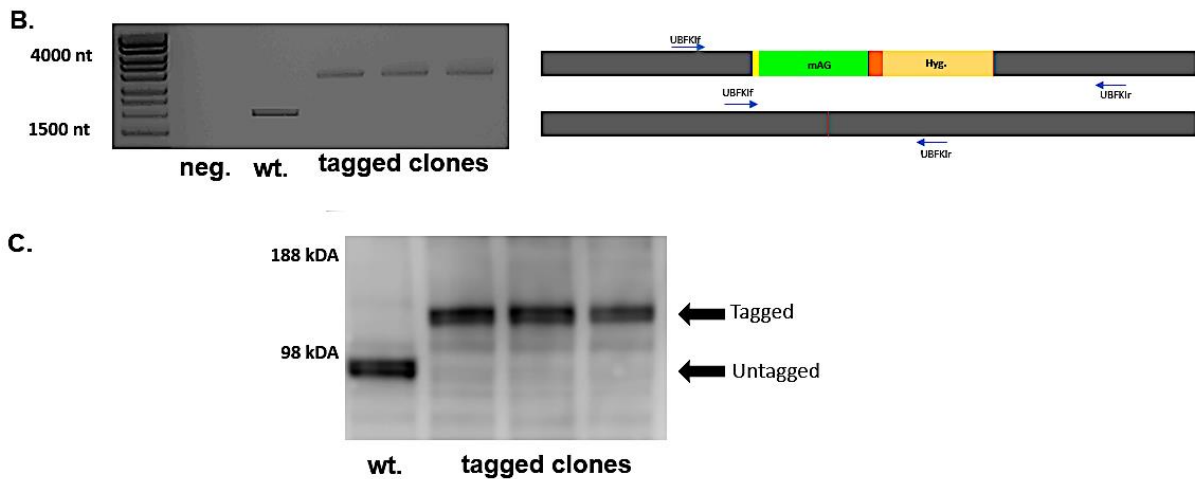
For the producing the RNPs, we used synthetic gRNAs (crRNA+trRNA) from IDT as well as IDT Cas9(HiFi) protein. I developed a manual based on a general description from IDT. Described in detail in the material and method chapter, 2.3 CRISPR/Cas9 genome editing. The components were co-transfected, and the cells could rest for 24 h after the transfection.



**Fig. 3.12** gRNA and sequence design for the UBF tagging strategy. **A.** Crispr/Cas9 complex with the crRNA sequence for UBF (gRNA\_UBF9) the gRNA sequence (as DNA bases) are shown in yellow and black, Cas9 protein in blue, genomic DNA in green and grey, Stop codon in red. **B.** genomic DNA sequence and its translated AS sequence of the original hUBF sequence, PAM sequence is indicated in yellow, STOP codon in red and the homology sequence in grey. In the down part the redesign of the sequence is shown for the C-Terminal tag of the hUBF gene. The PAM sequence is mutated as well as the STOP codon, which were part of the recognition sequence of the gRNA. A short AS linker sequence is shown in yellow, followed by the start of the mAG sequence (green). **C.** illustration of the selection method after successful integration into the genome of the cell. Sequences for mAG are in green, for P2A in orange, for Hyg. in beige, random integrated sequence in purple and stop codon in red.

Selection was started based on Hygromycin B (20  $\mu$ g/ml), for 2.5 weeks. After that time a pool of cells was checked for a green, nucleolar localised, UBF signal. All cells in the pool showed a correct signal localisation (example of observed signal see Fig. 3.13 A.). Therefore, genomic DNA was isolated from single clones and used as templates for PCR amplification. Primers were designed for this analysis, called UBFKIf/r to determine if one or both alleles have been tagged. Successful targeting would lead to a PCR product with a size of 3681 bp, wild type genomic DNA would lead to a product size of 1951 bp. Three out of ten clones were homozygous tagged by the mAG-P2A-Hyg. construct (see Fig. 3.13. B.). These three clones were used for further Western-Blot (WB) analysis to ensure that the observed protein was indeed UBF, by using a specific UBF antibody (Methods 2.2.2). The WB analysis confirmed that the two alleles were successfully tagged. It also showed the typical shift for the two isoforms of UBF, hUBF1 and hUBF2 (as doublet bands) for the wt. sample as well as for the tagged UBF (see Fig. 13. C.).





**Fig. 3.13 A.** Different verification steps for the UBF-mAG Knock-In. UBF-mAG-hTERT-Rpe1 cells were fixed, all cells were shown to express UBF-mAG and this protein was shown to localize correctly during interphase and metaphase. **B.** PCR strategy used to determine whether mAG\_P2A\_Hygromycin sequence had integrated into one or both UBF gene alleles. A primer pair was designed for the PCR amplification of the region surrounding the stop codon of UBF. The genomic DNA samples prepared above were used as template for this PCR experiment. PCR products were subsequently resolved using agarose gel electrophoresis. A single PCR was observed for each sample. PCR of the WT hTERT-Rpe1 genomic DNA resulted in a product (1951 bp) between 1000 and 1500 bp. The PCR product for the successful tag (3681bp) resulted in a product between 3kb and 4 kp. **C.** Western Blot analysis was used to ensure that the observed protein was indeed mAG tagged UBF. Both cells, WT and UBF-mAG cells were counted and lysed in an appropriate volume of NUPAGE LDS Sample Buffer, to yield a concentration of  $12.5 \times 10^3$  cells/ $\mu$ L. 15  $\mu$ L of each lysate was loaded onto the gel. Proteins were separated and transferred onto a membrane, before being probed with the UBF antibody. A shift in UBF mobility could be observed, in the UBF-mAG tagged cell lines. The characteristic UBF doublet (isoforms UBF1 and 2) was observed in the WT cell lysate, as well in the lysate of the tagged cells. This suggested that all UBF, expressed in the generated cell lines was mAG tagged. The Western-Blot analysis also indicate that all three tagged clones have a fully functional C-terminal homozygous UBF-mAG tag. In all, three rounds of UBF tagging showed similar results.

All WB samples are normalized with respect to its cell number. Each lane contains the protein amount of around 2 million cells. In this chapter I was able to show that the new developed tagging strategy is working, based on the example of UBF. The newly developed UBF-tagged cell line shows the same growth characteristics and the morphological appearance as non-tagged hTERT-RPE1 cells. This observation was subjective. Planned cell growth assays for the created cell lines were precluded by the COVID lock-down.

With this newly created cell-line we are in the position to observe the dynamic behaviour of one of the three nuclear sub-compartments, by live-cell imaging.

## **Chapter IV**

## **4. Tagging of nucleolar markers for the DFC and GC compartment**

### **4.1 A wider perspective for tagging other marker proteins**

Unleashing the potential of the approach described in the previous chapter to generate a C-terminally tagged hUBF hTERT-RPE1 cell line, applying this technique to other targets was the next logical step.

In this thesis chapter a strategy will be described for the tagging of Fibrillarin (FIB) and Nop52, marker proteins for the DFC and GC sub compartment of the nucleolus respectively. The aim is to tag representative proteins from all nucleolar compartments and try to combine these tagged markers in a single cell line.

Based on this, a tagged FC marker (UBF-mAG) is now existing in hTERT-Rpe1 cells. Because of the use of hTERT-Rpe1HKO cells, we still have the possibility to use other efficient selection marker systems like Blasticidin to tag other bookmarker proteins for nucleolar sub-compartments. A combination of a FC/DFC but also a FC/GC double tag, would be very helpful to get a more detailed picture for the dynamic behaviour of the nucleolus. Therefore, Fibrillarin is a suitable candidate, as a bookmarker for the DFC sub compartment.

### **4.2 Tagging of Fibrillarin as a bookmarker for the DFC compartment**

#### **4.2.1 A little background about the function of FIB.**

Fibrillarin (FIB) is a component of a small nuclear ribonucleoprotein particle (snRNP) and therefore it is part of the DFC sub-compartment of the nucleolus. It is involved in the 2'-O-methylation of ribose (C/D box protein family). These proteins recognize the target RNA based on complementary sequences and bind it at the correct position. The catalytic subunit of Fibrillarin then becomes active and transfers the methyl group of S-adenosylmethionine to the ribose of nascent pre-RNA. It gets a part of the small subunit (SSU) processome which already forms during active transcription<sup>190</sup>.

The gene encoding Fibrillarin (FBL) is localized on human chromosome 19. The encoded protein contains three different main domains: a N-terminal repetitive domain that is rich in glycine and arginine residues, a central RNA binding domain and a C-

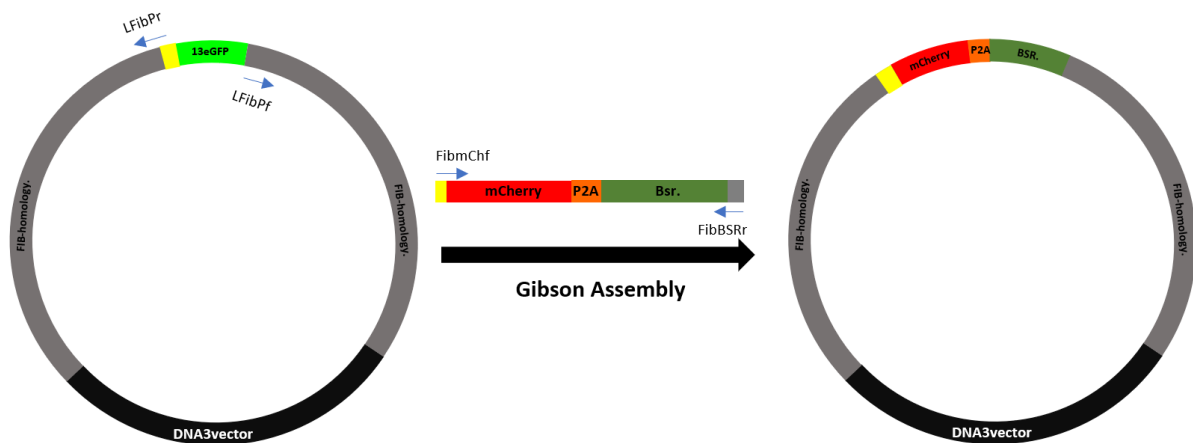
terminal alpha-helical domain. The RNA binding domain and the alpha-helical domain are typical for a methyltransferase (domain) <sup>191</sup>.

#### **4.2.2 Results for tagging Fibrillarin with mCherry**

As already mentioned, FIB is a classical marker for the DFC compartment. Therefore, a new construct was designed as a repair template for a C-terminal FIB tag based on mCherry. The construct itself was based on a mCherry-P2A-Blasticidin Knock-In cassette and had homology arms on either side of around 1 kb. This construct was co-transfected (see Methods, 2.3) together with a suitable RNP complex in the new hTERT-RPE1HKO-UBF-mAG-P2A-Hyge cell line to create a double tagged cell line. This design just included a two amino acid (AS) linker sequence, between the end of the original FIB sequence and the start of the mCherry tag. After 3 weeks of selection a pool of cells was checked and didn't show a red fluorescent signal. However, the cells were Blasticidin resistant. A possible explanation is that the proximity of the C-terminal end of FIB (alpha-helical domain), inhibits the successful folding of the mCherry which would explain the non-existence of a red fluorescent signal. A negative or inhibiting effect of this tag on the methyltransferase activity is also possible. Further characterisation, e. g. by WB or PCR was not carried out as no fluorescent signal was observed.

A review of the literature showed that Fibrillarin had already been tagged at the C-terminus with GFP by CRISPR/Cas9 in hiPSC cells. This Knock-In was heterozygous <sup>192</sup>. The repair template included a 13 AS linker sequence between the C-terminal end of the FIB sequence and the start of the GFP.

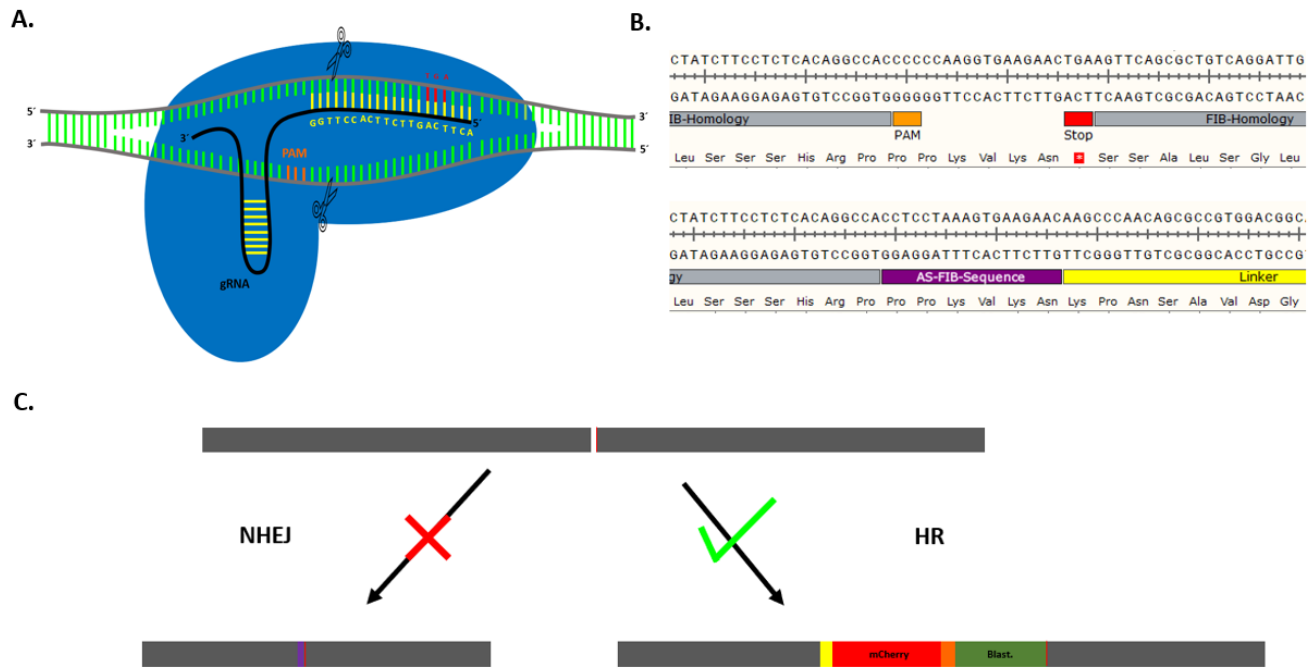
I re-designed the tagging strategy based on the published repair template, called AICSDP-13:FBL-mEGFP (Addgene). As already described for the UBF Knock-In, it is necessary for a repair template to have homology arms on either side of the Knock-In cassette. Because of the lack of suitable cutting sites close to the eGFP sequence in the original repair template from Roberts et al, I had to linearize the construct by PCR amplification. PCR primers were designed (LFibPf/r) annealing inside the linker sequence and at the start of the right homology arm, close to the C-terminal end of the eGFP sequence.



**Fig 4.1** Cloning strategy for the FIB repair template. The repair template for the Knock in is based on the AICSDP-13:FBL-mEGFP construct, because of the lack of suitable cutting sites close to the eGfp sequence we used it as a template to produce a PCR product. The PCR amplification is including the homology arms and the vector. A second PCR was performed based on a mCherry\_P2A\_Blast cassette as template, adding on 25 bps for the Gibson-Assembly. The linker sequence is indicated in yellow, the mCherry in red, the P2A in orange, the Blast. in green, the homology arms (1 kb) in grey and the vector in black.

A second PCR was performed (Primers FibmChf and FibBSRr), based on a mCherry-P2A-Blast cassette to create 25 bps overhangs for Gibson-Assembly (see Fig. 4.1). The two PCR constructs were used for Gibson-Assembly, resulting in a new construct, which was verified by sanger sequencing.

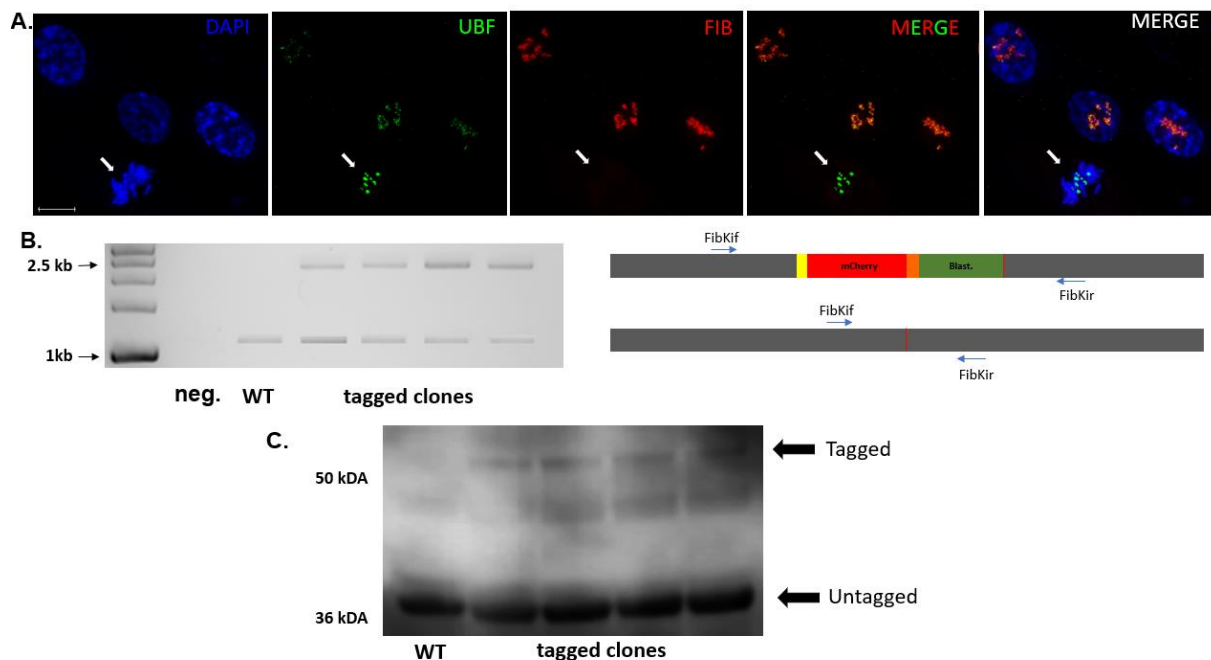
For the RNPs, a published gRNA sequence from Roberts et al., was used, as a synthetic gRNA from IDT combined with IDT Cas9HiFi protein (see. Fig. 4.2 A.) The co-transfection procedure of RNPs and the repair template follow the description in Methods, 2.3.



**Fig. 4.2** gRNA and sequence design for the FIB tagging strategy **A.** Caspr/Cas9 complex with the crRNA sequence for FIB (gRNA\_FIB) the gRNA sequence (as DNA bases) is shown in yellow, the PAM sequence in orange, Cas9 protein in blue, genomic DNA in green and grey and Stop codon in red. **B.** genomic DNA sequence and its translated AS sequence of the original Fibrillar sequence, PAM sequence is indicated in orange, STOP codon in red and the homology sequence in grey. In the down part the redesign of the sequence is shown for the C-Terminal tag of the Fibrillar gene. The PAM sequence is mutated as well as the STOP codon, which were part of the recognition sequence of the gRNA, shown in purple. We also included a 13 AS long linker sequence (yellow) between the mutated STOP codon and the start of mCherry. **C.** illustration of the two possible repair mechanisms (HR and NHEJ) after the CRISPR/Cas9 cut, after successful integration of the template by HR a selection based on Blast can be performed. The linker is shown in yellow, the mCherry in red, the P2A sequence in orange, Blast. Sequence in green, the STOP codon in red (red line) and the homology arms in grey. NHEJ is also possible, as well as the integration of other short sequences, in purple. Cells which performed NHEJ wouldn't survive the process of selection.

The transfected cells were kept under Blastidicin (5 µg/ml) selection, for around 2 weeks. After that time a pool of cells was checked for a red FIB signal. All cells in the pool showed a correct signal localisation (see Fig. 4.3 A.). Therefore, genomic DNA was isolated from single clones and used as templates for PCR amplification. Primers were designed for this analysis, called FibKlf/r to determine if one or both alleles have

been tagged. Successful targeting would lead to a PCR product with a size of 2334 bp, wild type PCR product would have a size of 1122 bp. A first check of 25 clones revealed that just one allele was tagged by mCherry\_P2A\_Blast. Four of these 25 clones were used for further analysis by WB (and PCR) (see Fig. 4.3 B. and C.). The WB, based on a FIB specific antibody (also used by Roberts et al., Methods 2.2.2) showed that wt. FIB as well as tagged FIB is getting expressed. The expression rate of the tagged FIB is very low compared the original wt. (see Fig 4.3 C.). However, especially the WB analysis confirms the results from the original FIB-eGFP Knock-In in the hiPSC <sup>192</sup>.



**Fig. 4.3** Different verification steps for the UBF-mAG/FIB-mCherry double tagged cell line **A.** UBF\_mAG\_P2A\_Hyge/Fib\_mCherry\_P2A\_Blast\_hTERT\_Rpe1HKO cells were fixed and analysed by dark field microscopy. All clones showed a Fibrillarin\_mCherry signal, which is localizing at the right position. It was compared to the UBF\_mAG signal and its function as a FC marker. The signal for Fibrillarin localizes correctly during both interphase and metaphase. **B.** A PCR strategy was used to determine whether mCherry\_P2A\_Blast sequence had integrated into one or both Fibrillarin gene alleles. A primer pair was designed for the PCR amplification of the region surrounding the stop codon of Fibrillarin. Successful integration of the mCherry\_P2A\_Blast sequence would result in a xxx bp PCR product because it includes the whole sequence for the Knock-In cassette. A single PCR was observed for each sample. PCR of the WT hTERT-Rpe1 genomic DNA resulted in a product (1951 bp) between 1000 and 1500 bp. The PCR product for the successful tag (3681bp) resulted in a product between 2.5 kb and 3 kb. **C.** Western Blot analysis was used to ensure that the observed protein was indeed mCherry

tagged Fibrillarin. Both cells, WT and FIB\_mCherry cells were counted and lysed in an appropriate volume of NUPAGE LDS Sample Buffer, to yield a concentration of  $12.5 \times 10^3$  cells/microL. 15 microL of each lysate was loaded onto the gel. Proteins were separated and transferred onto a membrane, before being probed with a FIB specific antibody. A shift in FIB mobility could be observed in the FIB\_mCherry tagged cell lines. This suggested that all Fibrillarin expressed in the generated cell lines was mCherry tagged. The Western-Blot analysis also indicate that all three tagged clones have a fully functional C-terminal heterozygous FIB\_mCherry tag. Three rounds of tagging experiments, based on the developed repair template showed similar results.

It should be also mentioned here that the new created double tagged cell line, showed the typical behaviour during stress response (cap formation). This will be shown in more detail in the following chapter 5. Based on these observations and the Western-Blot analysis, we assume that our heterozygous C-terminal FIB-mCherry tag localizes appropriately.

As a final remark I want to mention here that the Allen Institute is selling the created heterozygous FIB-eGFP hiPSC cell line as a commercial product (See Link). Nevertheless it should be noted that a heterozygous Knock-In is indicating that this tag is not fully functional. The reasons for this are not fully clear. It can be assumed that the created tag is inhibiting the function of FIB.

**Link:** <https://catalog.coriell.org/0/Sections/Collections/AllenCellCollection/CellLines.aspx?PgId=483>

### **4.3 Tagging of a suitable GC bookmarker protein**

In the previous chapters, I could prove that the developed tagging strategy is working, and double tagging is also possible.

The next logical step would be to tag a typical GC bookmarker protein. Also, here the Allen Institute is offering a couple of different commercially available hiPSC cell lines with Nucleophosmin 1 (NPM1) as a GC marker (see Link). The NPM1 was C-terminally tagged, but also heterozygous. Former work and publications in the McStay lab used, Nop52 as a GC marker. Therefore, I decided to create a tagging strategy for this GC marker protein.

### **4.4. Tagging of Nop52 with HALO**

#### **4.4.1 Some background information of Nop52**

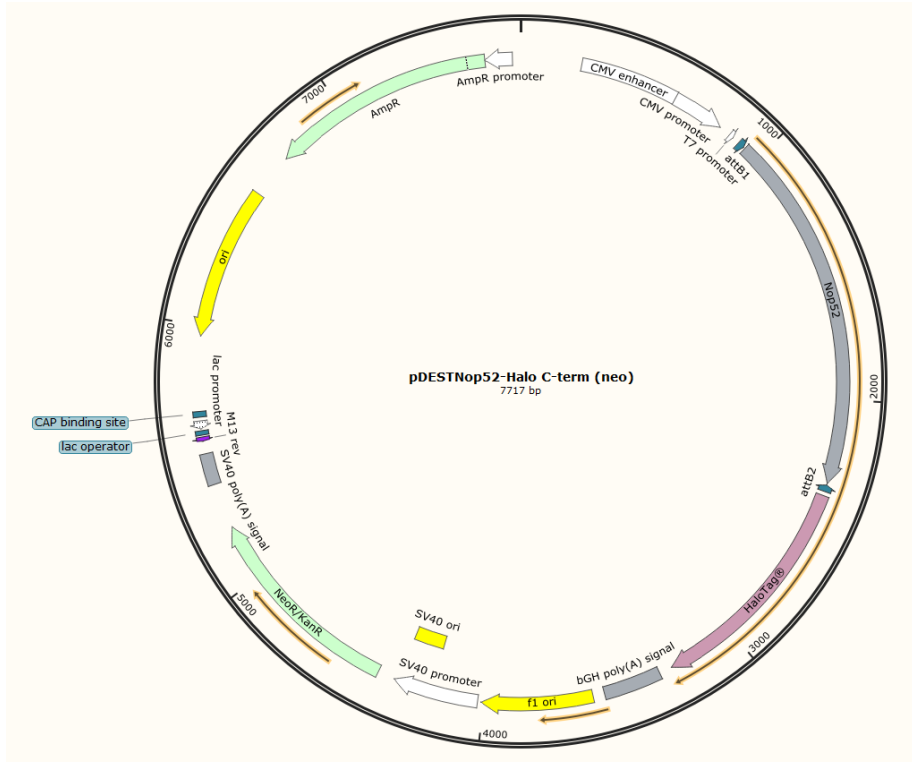
Ribosomal RNA processing protein 1 homolog A (RRP1) or also called Nop52 is encoded by the RRP1 gene, located on chromosome 21. Nop52 is involved in the late stages of nucleogenesis at the end of mitosis and maybe required for the generation of 28S rRNA <sup>193</sup>. It is also an excellent nucleolar GC marker protein <sup>135,167</sup>.

#### **4.4.2 Results for tagging Nop52 with HALO**

HALO is not one of the classical fluorescent proteins, it is a self-labelling bacterial enzyme/tag (HaloTag), more precisely it is a haloalkane dehalogenase: It can bind specific ligands (chloroalkanes) <sup>194</sup>. These ligands can be attached to any fluorophore or molecule of choice (like Biotin). PROMEGA offers a bright range of commercially available ligands. As part of this work, I used the TMR-dye from Promega, a red fluorescent cell permeable dye connected to the mentioned ligand. This tag and the staining allows to observe the “turnover-time” of an endogenous tagged protein.

A Nop52 C-term fusion had not previously been described, so it was necessary to create a suitable expression construct, encoding a Nop52-Halo fusion protein. This is based on the GateWay cloning system, using a pENTR Nop52 (C-term) and a pDEST-

C-term-HALO construct. A useful side effect of the GateWay system is that a specific linker sequence (attB2 see Fig. 4.5) is introduced between the protein of interest and the (fluorescent) tag. This is based on the so called “att”-sequences, which are fundamental for the GateWay cloning system to exchange DNA sequences between the pDEST- and pENTR- backbone constructs.

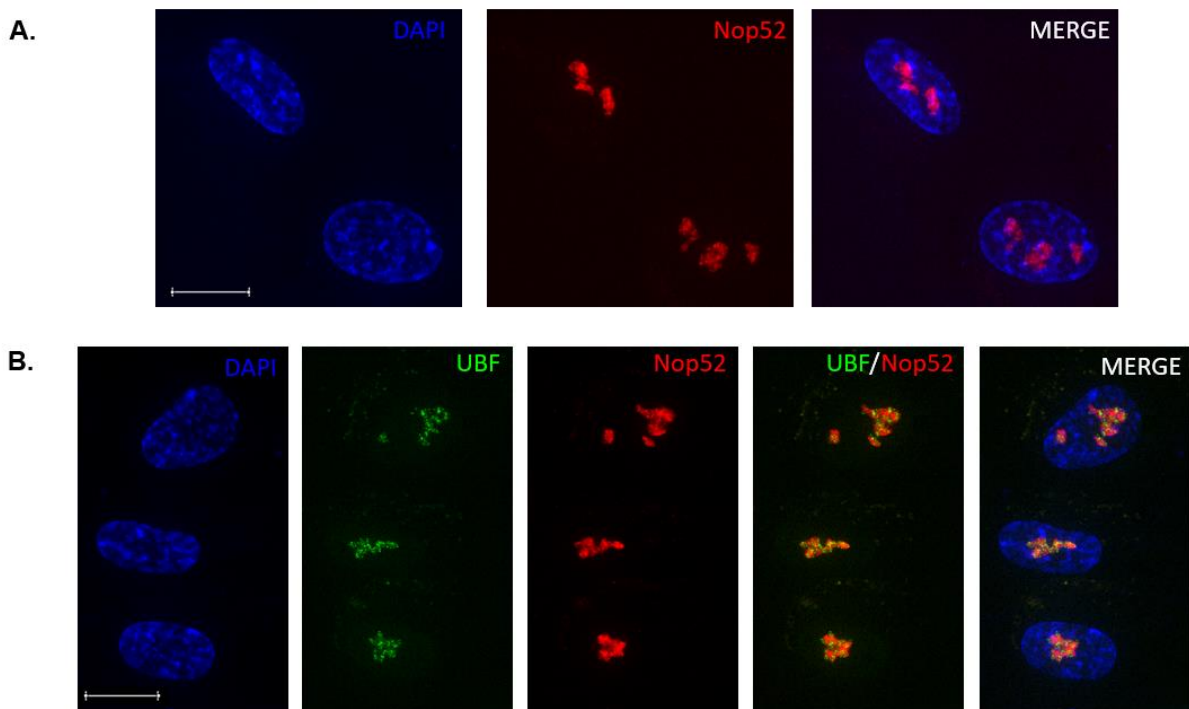


**Fig. 4.5** Map of the C-terminal Halo tagged Nop52 expression construct.

We transfected the new created fusion construct (pDEST-Nop52-Cterm-Halo) into hTERT-Rpe1HKO cells as well into the UBF-mAG Knock-In cell line, by a Calcium Phosphate precipitation.

We stained these cells with the cell permeable TMR-HALO dye from Promega, a red fluorescent dye which can specifically bind to HALO, following the manufacturer protocol (Methods 2.5). After fixing the cells with 4%-PFA and staining with DAPI, the cells were imaged. A transiently expressed red fluorescent nucleolar signal could be observed, correctly localizing to nucleolar GC compartments. Especially for the UBF-mAG tagged cell line, Nop52-Halo fusion signal surrounded the UBF signal and so showing the typical FC/GC architecture (see Fig. 4.6). Based on these observations,

we concluded that C-terminally Halo tagged Nop52 is localizing the correct way and is likely to be fully functional.

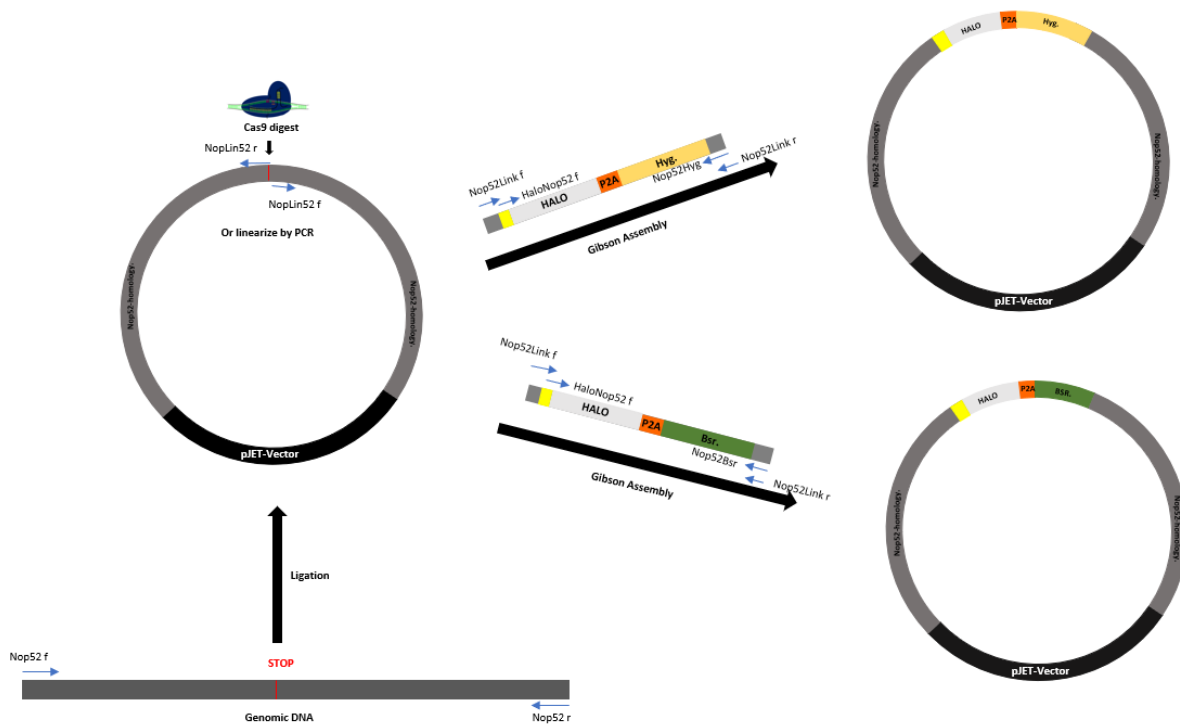


**Fig. 4.6** The transient expression of a Nop52-Halo construct **A.** Fixed hTERT\_Rpe1\_HKO cells with transient expression of pENTR\_Nop52\_Cterm\_HALO, stained before fixing with TMR-HALO dye. The signal is localizing as nucleolar signal. **B.** Fixed UBF-mAG tagged hTERT\_Rpe1\_HKO cells, with transient expression of pENTR\_Nop52\_Cterm\_HALO stained before fixing with TMR-HALO dye. The signal is localizing as nucleolar signal, and its bookmarker function for the GC compartment, relatively to the UBF-mAG signals and its function as FC marker.

From the previous experience with the FIB Knock-In, I decided to implement a linker sequence between the C-terminal end of the endogenous Nop52 gene and the start of the HALO sequence. Because of the promising results, based on the expression construct (pDEST-Nop52-Cterm-Halo), I used the same 18 amino acid “attB2” linker sequence for the endogenous NOP52 tagging strategy.

For the repair template I designed primers (primers Nop52 f/r), to amplify homology arms of around 1 kb flanking the stop codon of Nop52. Genomic DNA from hTERT-RPE1 cells was used as template. This PCR product was ligated into a pJET-vector (see Fig. 4.7).

It is necessary to linearize the new created construct, to implement a suitable Knock-In cassette. This can be done by a Cas9 “*in vitro*” digest or also by a pure PCR strategy. Both methods are working. An *in vitro* digest driven by a CRISPR/Cas9 complex can be seen as a first indicator for cutting efficiency of the designed gRNA.



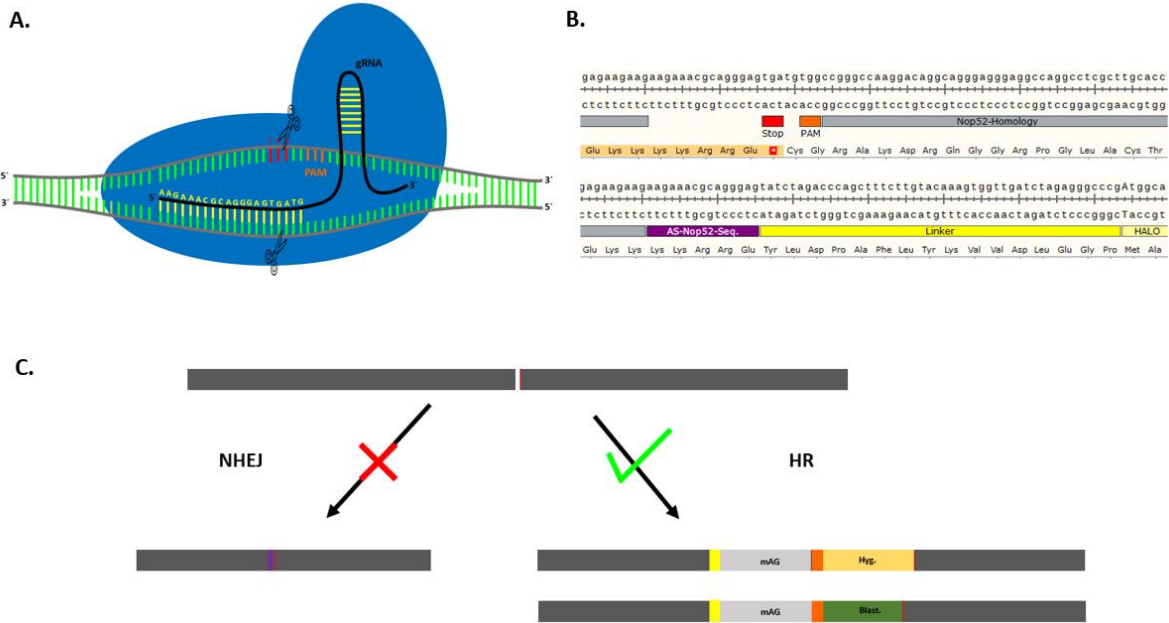
**Fig. 4.7** Cloning strategy for the Nop52 repair templates. Genomic hTERT-Rpe1 DNA was used to generate homology arms (1 kb) flanking the stop codon. It was ligated into a pJET-vector and linearized again by a Cas9 digest or by PCR (primers NopLin52f/r). For the PCR linearized homology arms it is just necessary to amplify Knock-In cassettes one time by PCR (primers HaloNop52f and Nop52Hyg./Bsr.R). For the Cas9 linearized construct it is necessary to amplify the Knock-In cassettes two times with the indicated primer pairs (HaloNop52f and Nop52Hyg/BsrR and then a second round with Nop52Linkf/r). The PCR products are getting cloned into the linearized Nop52 homology construct by Gibson-Assembly.

The design of the inserts were based on a Halo-P2A-Hyg./Blast. cassettes combined with an 18 AS long linker sequence as well as a 25bp overhang for the Gibson-Assembly at both ends. Therefore, two PCR amplifications must be performed. The first PCR reaction is based on Nop52Halof and Nop52Hyg./Bsr.R as primer pairs, for adding the linker sequence to the Knock-In cassette. Based on this product, a second amplification is necessary to add the 25 bp overhang for the Gibson-Assembly, using Nop52Linkf/r as primer pair. The fully assembled Knock-In cassettes (including linker and overhang) can be now used for the Gibson-Assembly, together with the Cas9

linearized homology construct. Resulting in the final two repair constructs, called pJET-Nop52-Halo-P2A-Hyg./Bsr.

Because Cas9 “in vitro” digests can be very inefficient and sensitive, the construct can be also linearized by PCR amplification. Therefore, Nop52Linf/r primer pair was designed, which is also adding a part of the linker sequence to this product. The primer pairs HaloNop52f and Nop52Hyg./Bsr.R can be used to amplify the Knock-In cassettes. The 25bp overhang sequence will be within the linker sequence, and can be used for the final cloning step resulting in the same repair constructs (pJET-Nop52-Halo-P2A-Hyg./Bsr.).

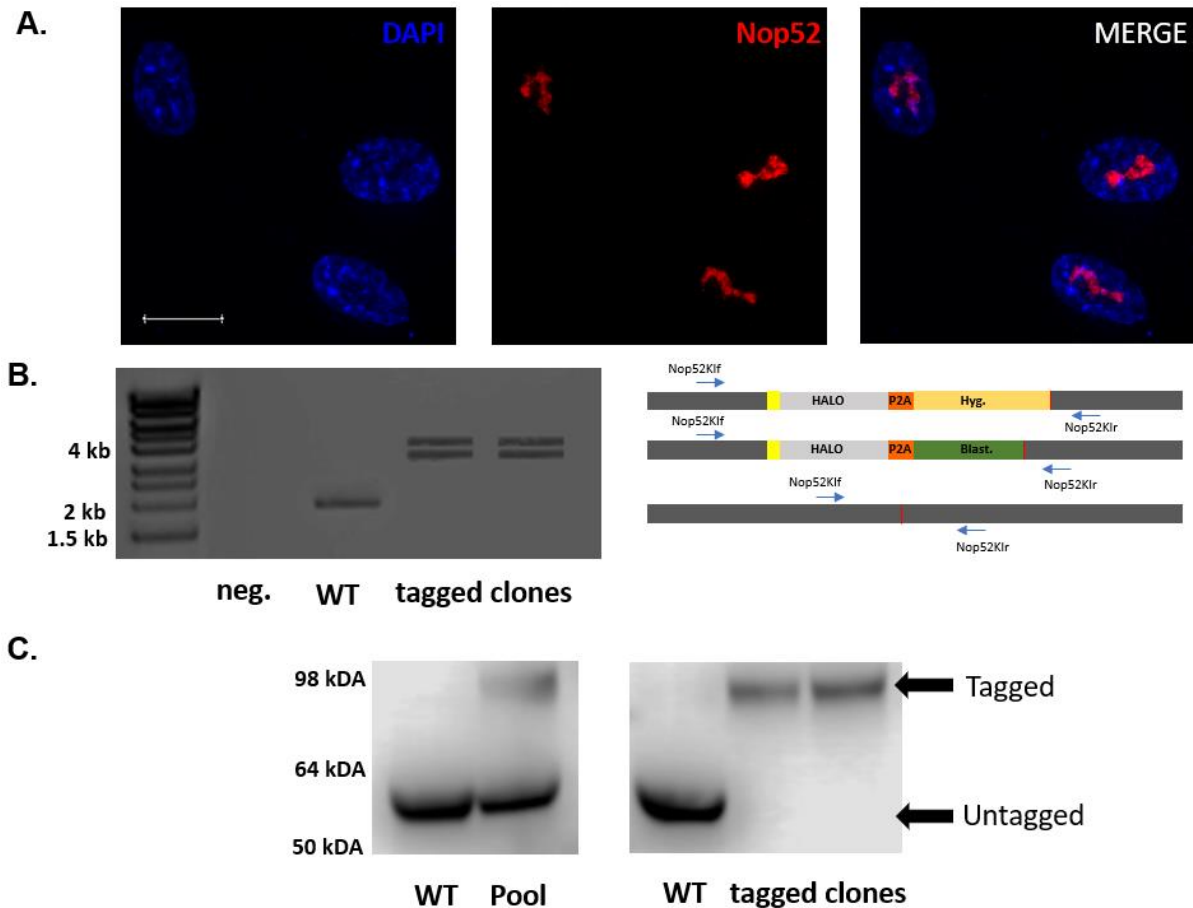
For the following experiments I used the repair constructs, based on the Cas9 “in vitro” digest.



**Fig. 4.8** gRNA and sequence design for the Nop52 tagging strategy . **A.** Crispr/Cas9 complex with the crRNA sequence for Nop52 (gRNA\_Nop52) the gRNA sequence (as DNA bases) is shown in yellow, the PAM sequence in orange, Cas9 protein in blue, genomic DNA in green and grey and Stop codon in red. **B.** genomic DNA sequence and its translated AS sequence of the original Nop52 sequence, PAM sequence is indicated in orange, STOP codon in red and the homology sequence in grey. In the down part the redesign of the sequence is shown for the C-Terminal tag of the Nop52 gene. The PAM sequence is mutated as well as the STOP codon, which were part of the recognition sequence of the gRNA, shown in purple. We also included a 16 AS long linker sequence (yellow) between the mutated STOP codon and the HALO sequence. **C.** Illustration of the two possible repair mechanisms (NHEJ and

HR) after the CRISPR/Cas9 cut, after successful integration of the template by HR a selection based on Hyge. or Blast can be performed. The linker is shown in yellow, the HALO in grey, the P2A sequence in orange, Blast. Sequence in green and Hyge. sequence in beige, the STOP codon in red (red line) and the homology arms in grey. NHEJ is also possible, as well as the integration of other short random sequences, in purple.

A gRNA sequence was designed, cutting within the Stop codon of Nop52 (see Fig. 4.8 A.). The Stop codon got mutated, resulting in a 18 amino acid long (attB2) linker sequence (see Fig. 4.8 B.). The RNP assembly and the transfection were performed as previously described, for UBF and FIB. The first transfection round was performed with hTERT-Rpe1HKO cells, combined with pJET-Halo-P2A-Hyg. construct as repair template. 24 h after the transfection HygromycinB selection was performed (20 µg/ml) for 2.5 weeks. After this time a pool of cells was checked by TMR-Halo staining and by WB, with a specific Nop52 antibody (see Fig. 4.9 C. left image). The staining and the following imaging showed that the red signal is localizing the right way in the hTERT-Rpe1HKO cells. However, the WB revealed that the observed TMR-Halo signals was indeed tagged Nop52, but the (WB) signal was weak compared to the wt. version of the protein. It also showed that the isolated pool of cells for WB analysis was heterozygous tagged.

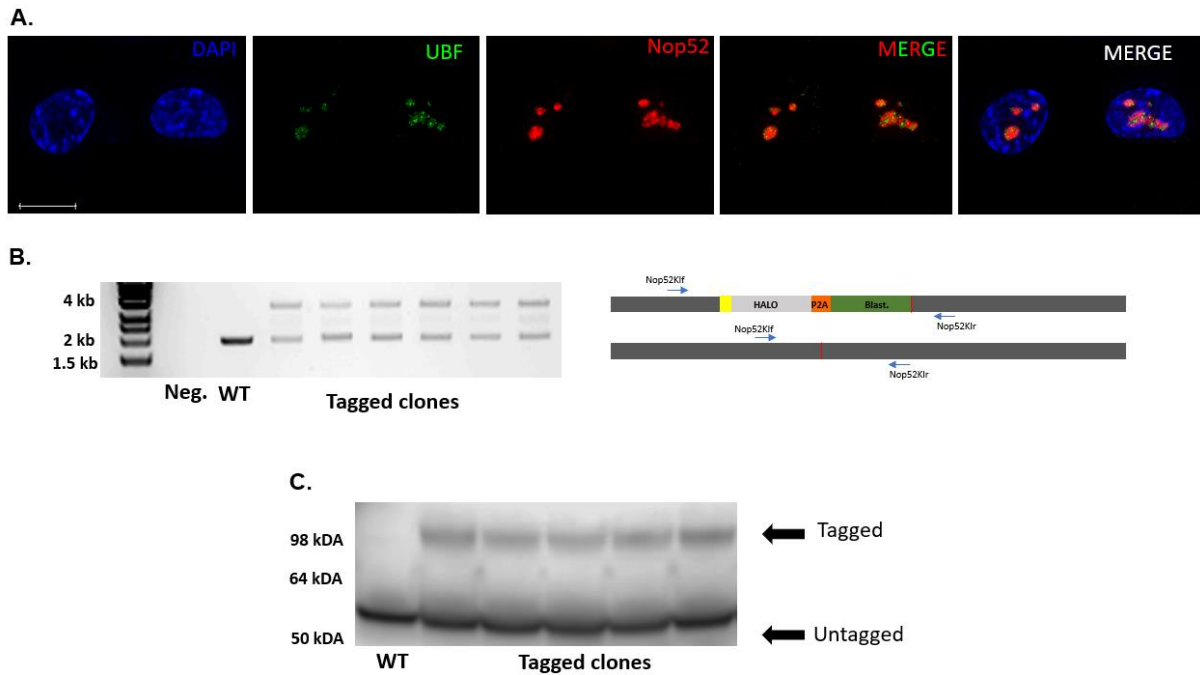


**Fig. 4.9** Different verification steps for the Nop52-Halo tagged cell line **A.** Nop52-Halo-P2A-Hyge/Blast-hTERT-Rpe1HKO cells were stained with TMR-Halo dye and fixed, analysed by dark field microscopy. All clones showed a Nop52-Halo signal, which is localizing at the right position as a nucleolar signal. **B.** A PCR strategy used to determine whether Halo-P2A-Hyge. or Halo-P2A-Blast. sequence had integrated into one or both Nop52 gene alleles. A primer pair was designed for the PCR amplification of the region surrounding the stop codon of Nop52. Successful integration of the Halo-P2A-Hyge or Halo-P2A-Blast sequence would result in 3893 bp/3287bp PCR products. However, PCR of the WT gene would result in a 1864 bp product. The genomic DNA was isolated and used as template for the PCR assay. PCR products were subsequently resolved using agarose gel electrophoresis. PCR of the WT hTERT-Rpe1 genomic DNA resulted in a product (1951 bp) between 1.5 kb and 2 kb. The PCR product for the successful tags (3287bp/3893bp) resulted in products between 3 kb and 4 kb. **C.** Western Blot analysis was used to ensure that the observed protein was indeed Halo tagged Nop52. Both cells, WT and Nop52 cells were counted and lysed in an appropriate volume of NUPAGE LDS Sample Buffer, to yield a concentration of  $12.5 \times 10^3$  cells/ $\mu$ l. 15  $\mu$ l of each lysate was loaded onto the gel. Proteins were separated and transferred onto a membrane, before being probed with a Nop52 specific antibody. A shift in Nop52 mobility could be observed in the Nop52\_Halo tagged cell lines. A first check on Nop52-Halo tagged cells showed that it was a heterozygous tag, WT Nop52 and Nop52-Halo got expressed. After a second round of tagging two clones were analysed and showed that just Halo tagged Nop52 got

expressed. This is showing that the C-terminal tag of Nop52 with Halo is fully functional. Two rounds of tagging showed similar results.

Therefore, a second round of tagging was performed based on the isolated pool of tagged cells. pJET-Nop52-Halo-P2A-Bsr. was used as a repair template, to tag the other wt. allele of Nop52. After 3 weeks of Blasticidin selection (5 µg/ml), two clones could be picked. The two clones showed an intense TMR-Nop52 signal (Fig. 4.9 A.) localizing the correct way. PCR analysis based on genomic DNA from the two clones, (primer pair Nop52Klf/r) revealed that the two alleles are tagged. Each of the alleles is tagged by a distinct selection marker (see Fig. 4.9 B.), therefore resulting in PCR products of different sizes. This result got confirmed by WB. analysis, surprisingly the expressed homozygous tagged Nop52 is weaker compared to his wt. version. All WB samples (wt., pool and clones) are normalized with respect to the cell number, each lane/sample is including the protein amount of 2 million cells (see Fig. 4.9 C). It could be also observed, by live-cell imaging that the TMR-Nop52 signal is detectable for around 5 days. However, based on this tagging strategy and the showed results it could be proved that a homozygous C-terminal Nop52 Halo Knock-In is possible to create and therefore fully functional.

To create a double tagged FC/GC bookmark cell line, the already verified homozygous UBF-mAG tagged cells were used for another round of tagging. Because Hygromycin B was already in use for the UBF Knock-In, Blasticidin was the selection marker of choice. The transfection, and verification procedure was the same as already described for the previous Nop52-Halo tag (see Fig. 4.10).



**Fig. 4.10** Different verification steps for the UBF-mAG/Nop52-Halo double tagged cell line **A.** Nop52\_Halo\_P2A\_Hyge/Blast\_hTERT\_Rpe1HKO cells were stained with TMR-Halo dye and fixed, and analysed by dark field microscopy. All clones showed a Nop52-Halo signal, which is localizing at the right position as a nucleolar signal. UBF\_mAG and its function as a FC marker is localizing within the GC signal of the Nop52\_Halo **B.** A PCR strategy used to determine Halo\_P2A\_Blast. sequence got integrated into the Nop52 gene. A primer pair was designed for the PCR amplification of the region surrounding the stop codon of Nop52. Successful integration of the Halo\_P2A\_Blast sequence would result in a 3287bp PCR product. However, PCR of the WT gene would result in an 1864 bp product. The genomic DNA was isolated and used as template for the PCR assay. PCR products were subsequently resolved using agarose gel electrophoresis. PCR of the WT hTERT-Rpe1 genomic DNA resulted in a product (1951 bp) between 1.5 kb and 2 kb. The PCR product for the successful tags (3287bp/3893bp) resulted in products between 3 kb and 4 kb. **C.** Western Blot analysis was used to ensure that the observed protein was indeed Halo tagged Nop52. Both cells, WT and Nop52\_Halo cells were counted and lysed in an appropriate volume of NUPAGE LDS Sample Buffer, to yield a concentration of  $12.5 \times 10^3$  cells/microL. 15 microL of each lysate was loaded onto the gel. Proteins were separated and transferred onto a membrane, before being probed with a Nop52 specific antibody. A shift in Nop52 mobility could be observed in the Nop52\_Halo tagged clones. This is showing that all screened clones are expressing heterozygous C-terminal tagged Nop52\_HALO. Two rounds of tagging showed similar results.

All 5 isolated clones were heterozygous. However, Because of the former results, shown in this chapter, we can be certain that Nop52, which is C-terminally tagged by

HALO is fully functional. However, it is quite possible that if more clones had been screened, there would have been a homozygous one.

#### **4.3 Discussion and future perspective**

In this chapter I was able to show, that CRISPR/Cas9 tagging of genes in hTERT-RPE1 cells is possible. The important key steps are the transfection efficiency, a suitable way to deliver the CRISPR/Cas9 complex (deliver as RNPs) and above all the connection of successful integration at the right position with selection. This is based on the developed “Knock-In cassettes”. The different genes could be tagged at the C-terminus (hUBF, FIB, Nop52) with three different FP (mAG, mCherry, HALO), based on this I created 4 cell lines: UBF-mAG, UBF-mAG/FIB-mCherry, Nop52-HALO, UBF-mAG/Nop52-HALO using HygromycinB and Blasticidin as selection marker systems. Based on this method of creating Crispr/Cas9 mediated Knock-In’s it is not necessary to use FACS and risk contamination of the used cell lines. However, the developed CRISPR/Cas9 Knock-In method is simple, fast, and safe for a former tagging resistant cell line. I want to state here that all screened clones and all observed pools for the created cell lines always showed the right signal localisation. So it is almost certain that isolated clones are successfully tagged.

However, one argumentative weak point that must be mentioned here is the non-existent sequencing of the genomic DNA to verify correct integration of the repair template. This is based on some problems with the sequencing company, as well as the lock-down that was active at the time and the resulting time management. However, if we take a closer look at the individual Knock-In lines, we can say that especially the UBF and Nop52 tags are fully functional. This is based on the homozygous nature of these Knock-In’s. The visualized nucleolar sub-compartments (UBF for FC, Nop52 for the GC) showed normal behaviour as well as the cells itself did that. Especially, verification by Western-Blot analysis with specific antibodies for the respective proteins showed that the tagging was successful. A suitable-sized shift could be observed for the tagged proteins and most importantly no wt. version of the targeted protein was left. It should be pointed out that sequence verification of the resulting cell lines was not performed.

The tagging of FIB resulted in a heterozygous Knock-In. The tagged FIB showed the typical behaviour, which can be expected (especially cap formation, as stress response). Verification by Western-Blot revealed also successful tagging. However, the expression of the tagged allele was low compared to the wt. one. Based on these results, which are comparable to the one from Roberts et al., it can be assumed that the created FIB Knock-In is not fully functional. It is possible that the created FIB-mCherry Knock-In does not have the same functionality like its original wt. counterpart. The Allen Institute offers FIB-eGFP tagged hiPSC cells as commercial product. A possible way to prove if a homozygous FIB Knock-In is possible would be to use one selection marker for a single allele, like it was shown for Nop52-Halo.

Based on the methodologies I have developed, other members of the McStay lab were able to tag another FC maker protein TCOF-I at the C-terminus with an mCherry-P2A-Blast sequence. The knock-in technique that I have developed can be also modified with other useful sequences. If a knock-in cassette gets combined with an Auxin inducible degradation (AID2) sequence it will be possible to knock-down the endogenous levels of successfully (homozygous) tagged proteins. This strategy was applied to Ki67, a successful homozygous tagged Ki67-AID-mCherry-P2A-Blast. cell line could be created. Future experiments based on this cell line will reveal the influence of Ki67 to nucleolar dynamics and formation.

As already mentioned in the introduction of Chapter 3, a discussion about suitable cell lines and their role as model systems for experiments should be taken more seriously. Many bio-science related research groups, avoid using immortalised or primary cell lines because they are more difficult to handle in the aspect of genetic manipulation. I hope I have been able to show that the developed method works reasonably efficiently to endogenously tag any protein of interest.

However, this method can be more optimized. Future Knock-In's based on developed method could use linear gBlock-Genes (IDT), which can have a size currently up to 3000bps as a repair template, instead of a circular repair template construct. This would allow a higher concentration of pure repair template within the cell nucleus and therefore, it could lead to a higher amount of homozygous tagged clones within a transfected pool of cells.

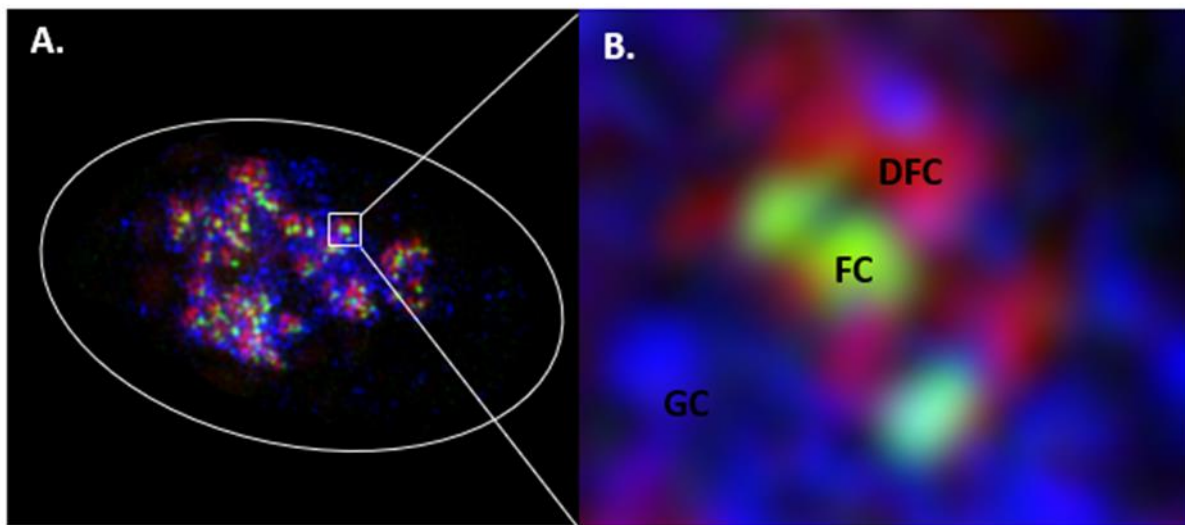
# Chapter V

## 5. The dynamic behaviour of the nucleolus

### 5.1 Different aspects about the nucleolus

Nucleoli are the most prominent membrane less structures within the nucleus of eukaryotic cells. They belong to a group of membrane-less organelles (MLOs) and as such, they are dynamic structures with highly mobile constituents that can diffuse between the nucleolus and nucleoplasm.

As previously described in the introductory chapter, nucleoli have a tripartite structure in higher eukaryotes. The three substructures or sub compartments are functionally “distinct”. The fibrillar centres (FC) contain non-transcribed rDNA and rDNA chromatin associated factors. rDNA transcription occurs at the interface between FCs and the dense fibrillary component (DFC). Early processing of precursor ribosomal RNA (rRNA) takes place in the DFC, whereas the granular component (GC), surrounding the FC/DFC, is the location for late processing of released pre-rRNAs and assembly of ribosome units. The typical tripartite structure in human cells depends on active transcription of rDNA <sup>135</sup> (See Fig. 5.1).



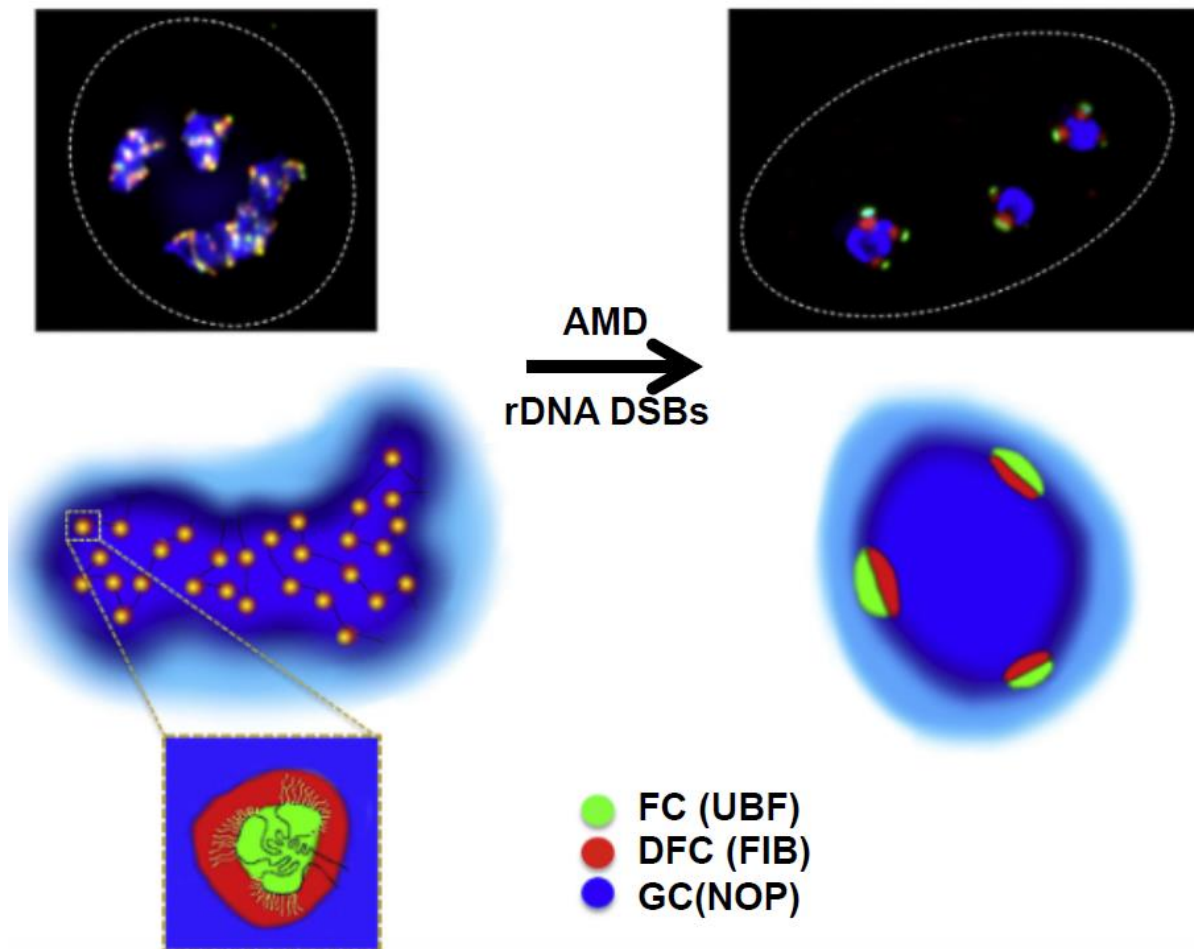
**Fig. 5.1:** A closer look on the nucleolus and its sub compartments **A.** Tripartite structure of a nucleolus visualized by IF. **B.** The FC in green (UBF as bookmark) is surrounded by the DFC in red (FIB as bookmark), the GC in blue hosting FC/DFC structures. UBF, FIB and Nop52 are bookmarkers for the FC, DFC and GC compartments. (IF image taken from <sup>135</sup>)

This organization reflects the vectorial progression of ribosome biogenesis from DFC (early processing) to GC (late processing).

During mitosis, the nucleolus breaks down and transcription driven by the Pol-I machinery stops. Many FC proteins localize to the secondary constrictions. Other nucleolar proteins, localising to the DFC and GC nucleolar compartments, are found at the surface of metaphase chromosomes <sup>195</sup>. After cell division, nucleolar proteins are evenly distributed between daughter cells, which is facilitated by their localisation at the periphery of mitotic chromosomes <sup>196</sup>. After metaphase, nucleolar proteins at this chromosome periphery compartment merge to generate pre-nucleolar bodies (PNBs). Interestingly, Ki67 is essential for maintaining the chromosome periphery department and when depleted cells can divide but are slow to generate nucleoli <sup>148,197</sup>. The assembly of the nucleolus starts during telophase by activation of rDNA transcription of several nucleolar organizer regions (NORs) and by progressive recruitment of early and late processing proteins on rRNA transcripts. PNB proteins are recruited to these reforming nucleoli, multiple small nucleoli then fuse to generate larger nucleoli surrounded by peri-nucleolar chromatin (PNH) <sup>136</sup>.

Live cell imaging experiments in HeLa cells showed that the concentration of these factors increases in nucleoli and decreases in PNBs. This is a biphasic process with transfer of early processing factors associated with the FCs being complete within 15-20 minutes. Complete transfer of the GC factors required at least 90 minutes. Interestingly the traffic between PNBs and nucleoli was unidirectional for the DFC protein fibrillarin and bidirectional for GC proteins such as NPM and Nop52 <sup>156</sup>. These experiments were based on the transient expression of the relevant proteins, encoded on suitable plasmids.

The induction of nucleolar stress reveals a further aspect of the dynamic nucleolus. Stress can be induced using pharmacological inhibitors of rDNA transcription by RNA Pol I. Inhibitors include Actinomycin D (AMD). Nucleolar stress can also be induced by introduction of DNA double strand breaks (DSB) within the rDNA. Such DSBs lead to ATM dependent inhibition of transcription. In both cases, nucleolar stress is associated with nucleolar segregation.



**Figure 5.2:** Nucleolar cap formation as part of the stress response. Figure adapted from <sup>167</sup>.

AMD is a peptide antibiotic from *Streptomyces parvulus*. It belongs to the group of actinomycin's and consists of two cyclic peptides linked by an oxidised Phenoxazine unit <sup>198</sup>. At low doses, AMD inhibits DNA transcription by acting as a cross-linker. The phenoxazine group plays an important role here as it specifically binds to the guanine and cytosine nucleotide-rich sites of the DNA. Its effect is not bound to a specific phase of the cell cycle <sup>199</sup>.

As a result of the stress response the FC and DFC compartment depart from their typical internal localization to form dense nucleolar caps at the nucleolar periphery of spherical nucleoli (GC). This would make the rDNA breaks accessible to the DNA repair machinery, which is located external to the nucleolus <sup>166</sup>. Single NORs form a single cap at the periphery of the nucleolus, because of the stress response (see Fig. 5.2) <sup>54</sup>. However, it is still unclear what mechanism is at work and what exactly serves

as the driving force for the cap formation. In principle, two assumptions could be possible:

The individual DFC/FC compartments act as a homogeneous unit which is separated from its environment, like a liquid phase which is then pushed to the side of the nucleolus forming caps. Thus, the LLPS mechanism could offer an explanation. Another possible view, the individual rDNA units or FC's are the starting point for the separation, so there is no homogeneous unit, but rather a sum of individual small units that are individually pulled to the side of the nucleolus, forming caps. The DFC compartment would follow because of its binding affinity to FC's.

However, during the last couple of years, the role of liquid-liquid phase separation (LLPS) in formation of MLOs has been increasingly recognized. LLPS has a role in the formation and internal organization of the nucleoli to functional substructures. The lack of nucleolar membrane has long been intriguing biologists and (bio)physicists alike: Questioning the bio-physical nature of the nucleolus. It should not be just seen as a soup of different macromolecules following the laws of surface tension and viscosity. Studies in frogs found that nucleoli in *Xenopus laevis* oocytes behave like liquid droplets *in vivo*, as well as (partial or modelled) when reconstituted *in vitro*. Based on these experiments it was suggested that nucleoli form through liquid-liquid phase separation (LLPS) of the nucleolar compartments in the nucleoplasm (Berry et al., 2018; Brangwynne et al., 2011; Feric et al., 2016). It has also been shown that there is a correlation between the volume distribution of such nucleoli and a diffusion-limited aggregation process. This is based on a constant influx of different factors <sup>162</sup>. In addition, it should be also mentioned that the size of nucleoli in the worm *C. elegans* embryos was found to be dependent on the concentration of nucleolar components in the nucleoplasm which is consistent with the liquid like nature of the nucleolus <sup>202</sup>. It is also suggested that the nucleolar sub compartments within the nucleolus (GC), the FC and DFC form via liquid-liquid phase separation <sup>201</sup>. However, these observations and the assumptions derived from them are in contradiction to observations made in *Drosophila* embryos. Wieschaus et al, could show that nucleolar formation is precisely timed and follows the activation of rDNA transcription. Therefore, the best way to

describe nucleolar formation and behaviour is a combination of active recruitment of different factors and LLPS<sup>170</sup>.

Another study, which had a closer look on nucleolar behaviour in living cells supported the idea of LLPS behaviour of nucleoli. As part of this publication, they describe the nucleolus as a “colloidal solution”, which is containing solid-like granules (FC/DFC). They also could observe nucleolar fusing *in vivo*, a specifically assigned LLPS behaviour. Furthermore, it was noted that active processes (ATP-dependent) are involved in maintaining the nucleolus-nucleoplasm interface. They could observe an irregular shape of nucleoli after ATP depletion, using 2-deoxyglucose and trifluoromethoxy-carbonylcyanide phenylhydrazine<sup>203</sup>. A caveat to these experiments is that they relied on transient expression of NPM and FBL FP-fusion proteins, with no control over protein expression levels or evidence for full functionality.

There is no doubt that *in vivo* experiments deliver the most reliable answer to nucleolar behaviour and the role of the LLPS model. Live cell imaging is the method to visualize nucleolar behaviour in its natural environment, the cell nucleus. The cited studies showed “natural” behaviour in the context of *in vivo* experiments, based on overexpression of different nucleolar key proteins, which are part of the single sub compartments. But these studies didn’t show the normal expression rate or behaviour of the proteins, controlled by the cell itself. Also, the used (human) cell lines, which are all cancer cells, are not the most reliable model system to target the question of normal or natural behaviour. Non-transformed (immortalized) cell lines with a normal karyotype, like hTERT-RPE1 cells, offer a better model system for cell culture-based experiments.

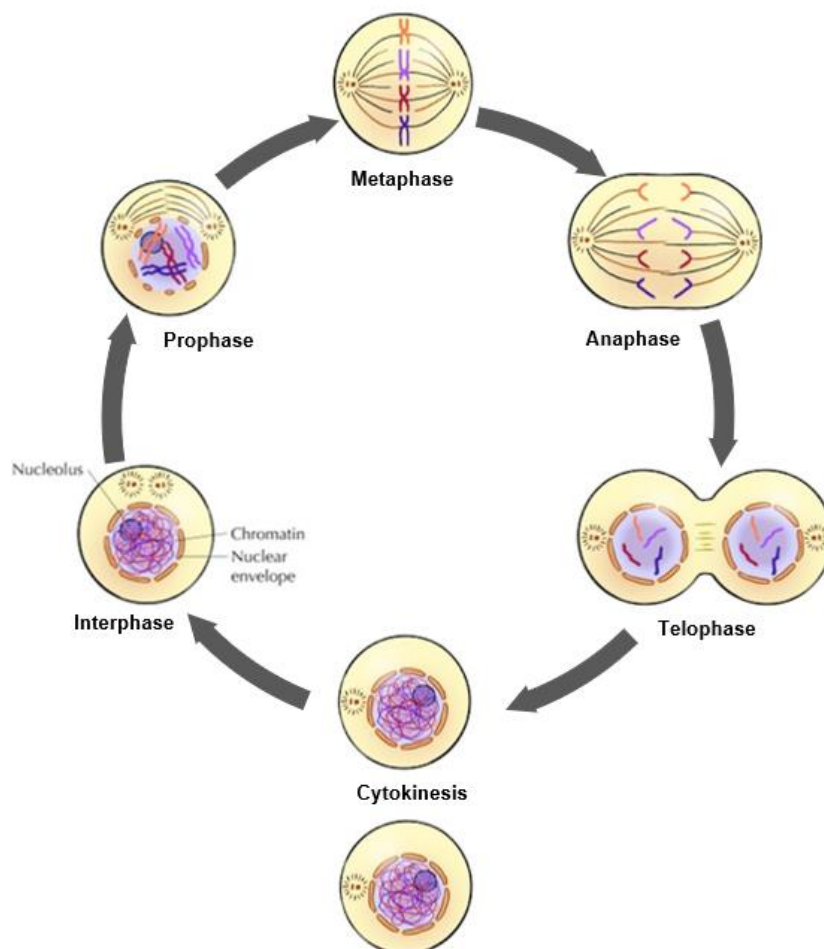
The CRISPR/Cas9 tagging of FC, DFC, and GC markers in hTERT-RPE1 cells offers a new platform for analysis nucleolar dynamics in live cells. In the previous chapters I described a way of tagging genes in hTERT-Rpe1 cells. Using the methodologies I have developed, it was possible to C-terminally tag specific markers for the single sub compartments (UBF (FC), FIB (FC) and Nop52 (GC)) of the nucleolus. For UBF and Nop52 and their role as specific markers, I was able to show that these tags are fully functional.

In this chapter I describe the dynamic behaviour of the nucleolar sub compartments during formation and nucleolar cap formation as a result of Pol-I inhibition.

## **5.2 Results**

### **5.2.1 The *in vivo* perspective on the nucleolus and its sub-compartments**

The Knock-In methods and cell lines I have developed provide an ideal tool for monitoring nucleolar behaviour. Formation and stress response could be observed by live-cell imaging. Key proteins from each of the nucleolar sub-compartments were tagged at the C-terminus by different FP proteins. As previously described in Chapter 3 and 4, I was able to create a homozygous UBF-mAG Knock-In in hTERT-Rpe1HKO cells. This cell line was used to create two different double tags: UBF-mAG combined with FIB-mCherry and UBF-mAG combined with Nop52-Halo. Also, a homozygous Nop52-Halo Knock-In was created, providing proof that C-terminally tagged Nop52 is fully functional. Thus, the tags for UBF and Nop52 are fully functional, the tag for FIB however is heterozygous and therefore probably not fully functional. This is also the case for the double tags. Using the double tagged cell lines described in chapter 3 and 4 it was possible to observe nucleolar formation and stress response from a FC/DFC (UBF/FIB) and a FC/GC (UBF/Nop52) perspective.



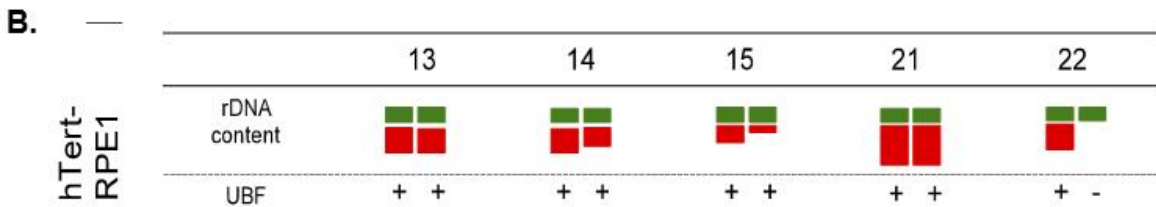
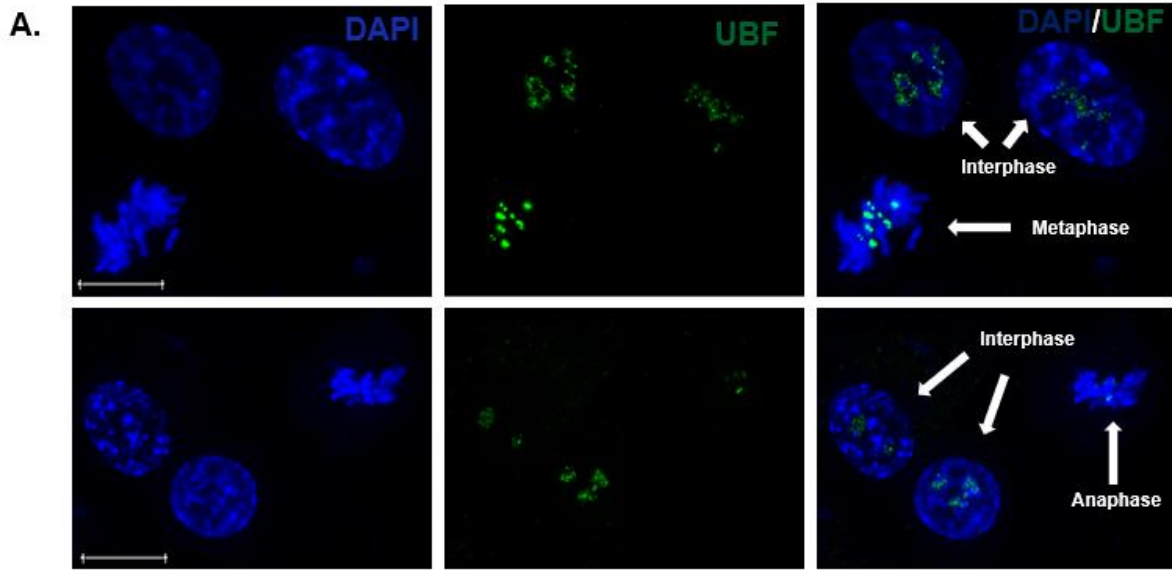
**Fig. 5.3** A simple illustration of the cell cycle.

With regard to cell cycle, a method to visualise chromatin in live cells would be beneficial. This would facilitate identification of various cell cycle stages, including pro-, meta-, ana- and telophase (see Fig. 5.3).

These stages can be identified by different chromatin staining or visualization methods. DAPI can be mentioned as the most prominent chromatin dye. However, this staining can be just used for fixed (dead) cells. Other prominent techniques such as live cell chromatin dyes, with fluorescent abilities in the “far red” do not work with hTERT-RPE1 cells. This will be discussed below.

In the following section I will argue that the FC protein UBF, which is bound to rDNA throughout the cell cycle can be used as a surrogate for identifying cell cycle stage. UBF is a marker for the FC compartment of the nucleolus, because of its binding affinity to rDNA. Therefore, UBF remains bound to the rDNA and thus to the respective acrocentric chromosomes during the entire cell cycle.

During metaphase each NOR is visible as a single UBF focus that disperses into multiple FC foci associated with DFC and localised with a larger GC upon the resumption of transcription.

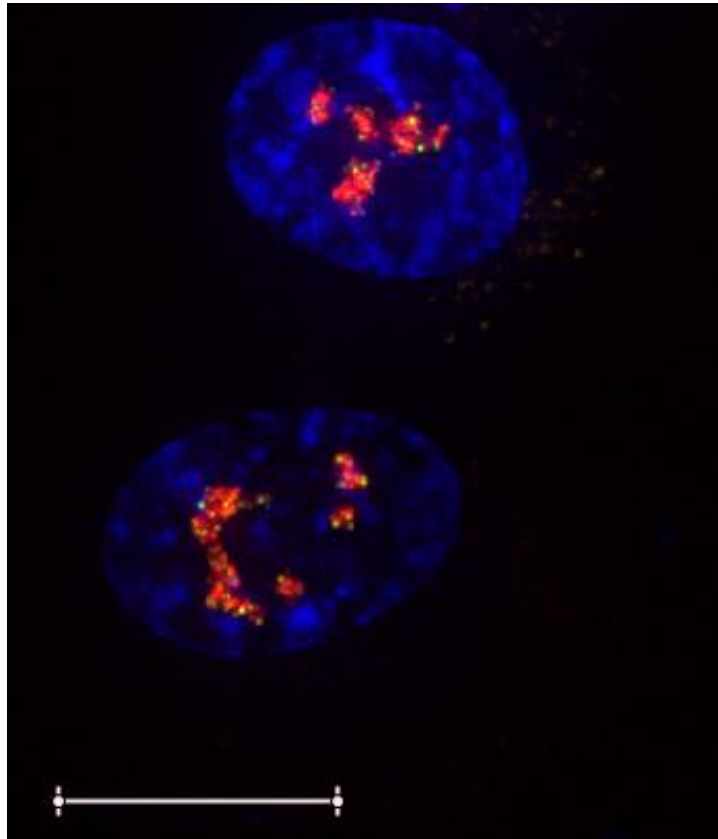


**Fig.5.4** UBF behaviour during meta- and anaphase. **A.** Fixed UBFmAG-hTERT-RPE1HKO cells during different stages of the cell cycle. Each panel shows cells with a condensed chromatin state and the characteristic UBF cap formation. In the upper panel a metaphase cell is present, but also two interphase cells with UBF as part of the nucleoli. In the other panel a (daughter) cell during anaphase can be seen and also two cells during interphase. **B.** NOR ideogram for hTERT-RPE1 cells, targeting rDNA and UBF loading level on the acrocentric chromosomes. Figure B. adapted from <sup>51</sup>.

Immuno-FISH experiments performed on metaphase spreads prepared from hTERT-RPE1 cells revealed different loading levels of UBF on the single acrocentric chromosomes (see Fig. 5.4 B) <sup>51</sup>. However also in fixed and in living cells it is possible to observe a difference between cells in metaphase or anaphase. It is obvious that there is a difference in the number and intensity of these caps for the two stages (see Fig. 5.4 A, two panels). However, an exact count of the individual spots is usually not possible in living or fixed cells, because of the 3-dimensional structure of the condensed chromosomes and the (low) resolution of the microscope. Usually there should be 18 foci present for an anaphase cell and but during metaphase, the sister chromatids are not fully resolved.

Nevertheless, the cap formation, can be used as an indicator for specific stages of the cell cycle during live cell imaging

### 5.2.2 Nucleolar formation from the FC/DFC perspective by live-cell imaging.

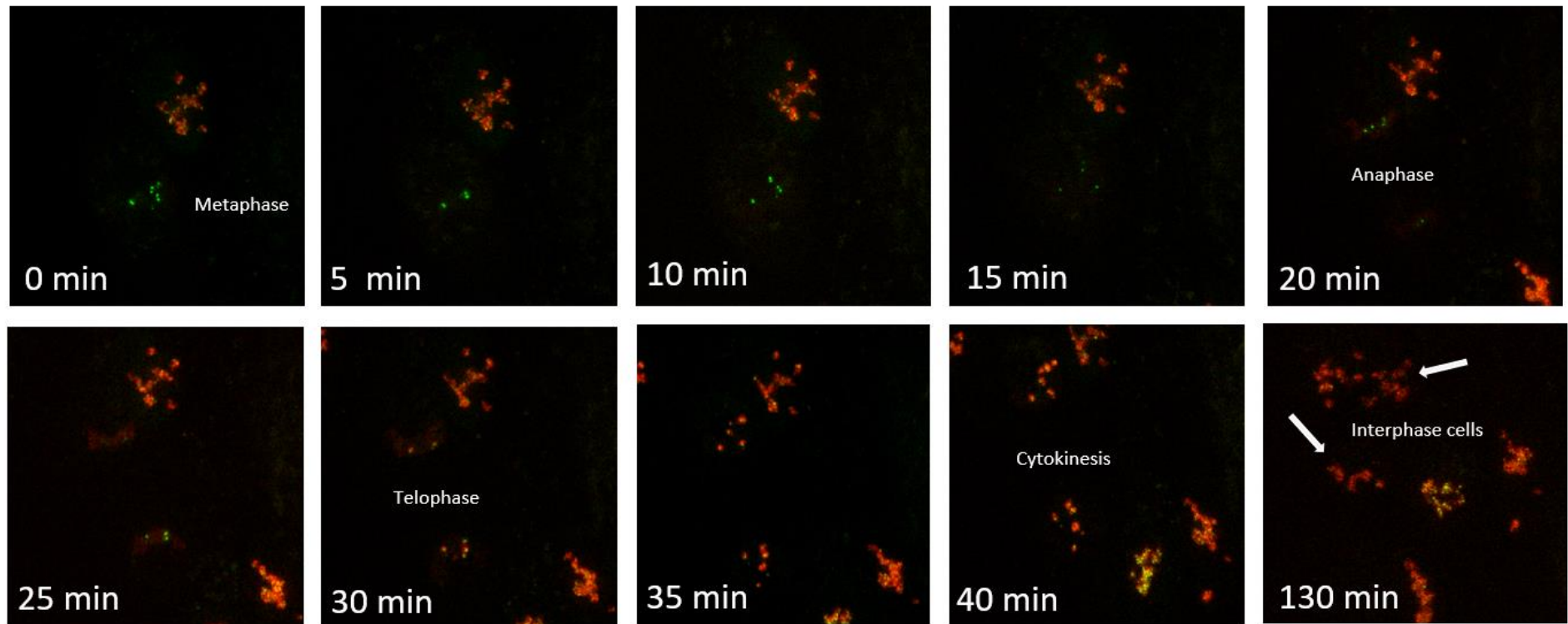


**Fig. 5.5** Image of fixed and DAPI stained UBFmAG-FIBmCh-hTERT-RPE1 knock in cells, showing the localisation of the FC (UBF green) and DFC (FIB red) bookmarks within the nucleus, stained by DAPI (blue). The bar is indicating 15 microns.

The following results from live cell imaging experiments with a UBFmAG and FIBmCherry double tagged cell line represent the first experiment to verify the functionality of the two tags within a single cell line. The following results will be also used to compare them with published data.

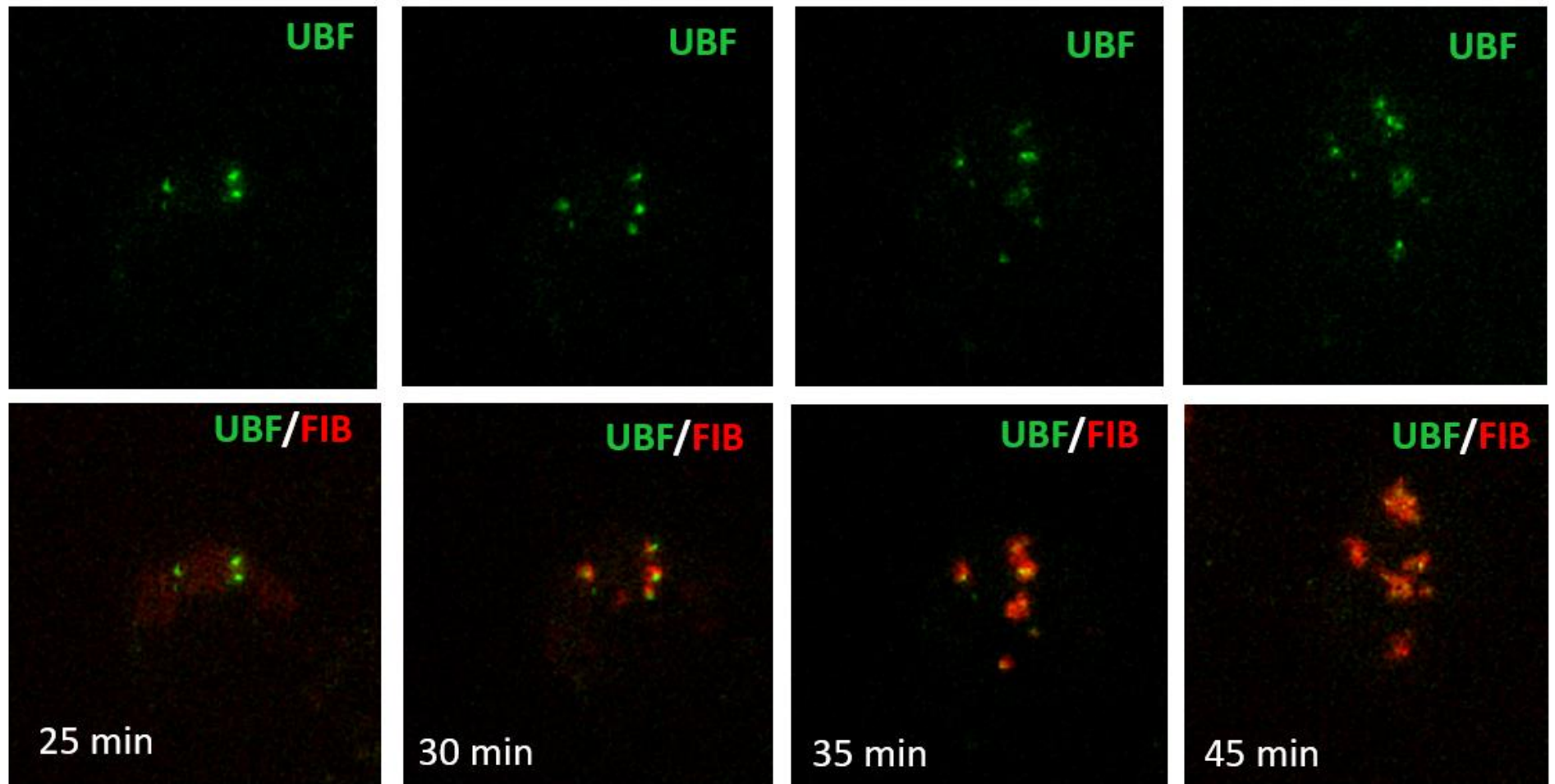
UBF and its role as a FC marker, based on its DNA binding abilities is a double useful marker in the aspect of live cell imaging. It is indicating the FC compartment and it also offers the possibility to visualize specific cell cycle stages, such as prophase, meta- and anaphase. During live cell imaging it is possible to identify metaphase cells based on the typical UBF-mAG caps (see Fig. 5.6, timepoint 0 min). During that stage multiple foci with different fluorescent signal intensities can be observed. 20 minutes later the cell is going through anaphase and is forming two daughter cells. Every new forming

daughter cell is still showing the typical UBF-mAG caps. FIB however is part of the PNBs and can be observed as a red cloud like signal (see Fig. 5.6 timepoint 20 min).



**Fig. 5.6.** Illustration of the formation event of the FC and DFC nucleolar sub-compartments in hTERT-Rpe1-UBFmAG-FibmCherry Knock-In cell line. UBF (in green) foci can be observed for meta-, ana- and early telophase. When rDNA transcription reappears, FIB (in red) is getting recruited and can be observed as red signal, which is surrounding the UBF. White arrows are indicating the daughter cells for the last time point. The different time points (min) are indicated for every image in white.

After anaphase the cell is entering telophase (see Fig. 5.6, timepoint 30 min), rDNA transcription reassumes, and FIB is getting recruited. Therefore, FIB is not just a marker for the DFC compartment, it is also indicating the start of transcription. In case pseudo NORs would be present in a tagged cell line, these would just show a FC cap signal, but they would not be transcriptional active.

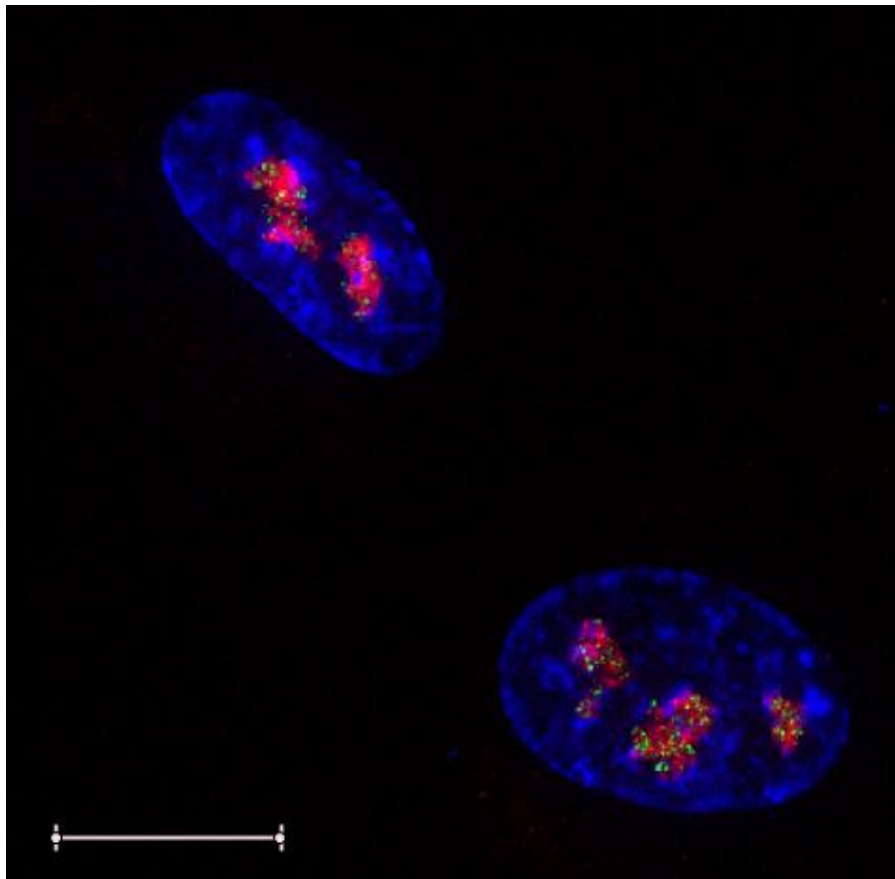


**Fig. 5.7** Enlargement of the formation event (from Fig. 5.6) After the separation of the chromosome pairs. The upper panel shows UBFmAG, the panel below is a fused signal of UBFmAG and FIBmCherry. The timepoint at 30min shows the start of transcription, indicated by the recruitment of FIBmCherry. The following timepoints show growing FC/DFC compartments of the nucleoli.

However, a closer look on the UBF caps reveals, that these caps divide apart (see Fig. 5.7 timepoints 25, 30, 35 and 45 min) shortly after transcription reassumes. A single cap turns into multiple FC/DFC units (see Fig. 5.6 and 5.7 timepoints 30, 35, 40 and 130 minutes) when the cell is going through cytokinesis and entering interphase.

These observations and the timing are typical for this double tagged cell line and could be observed in 6 live cell imaging experiments.

### **5.2.3 Nucleolar formation from the FC/GC perspective.**

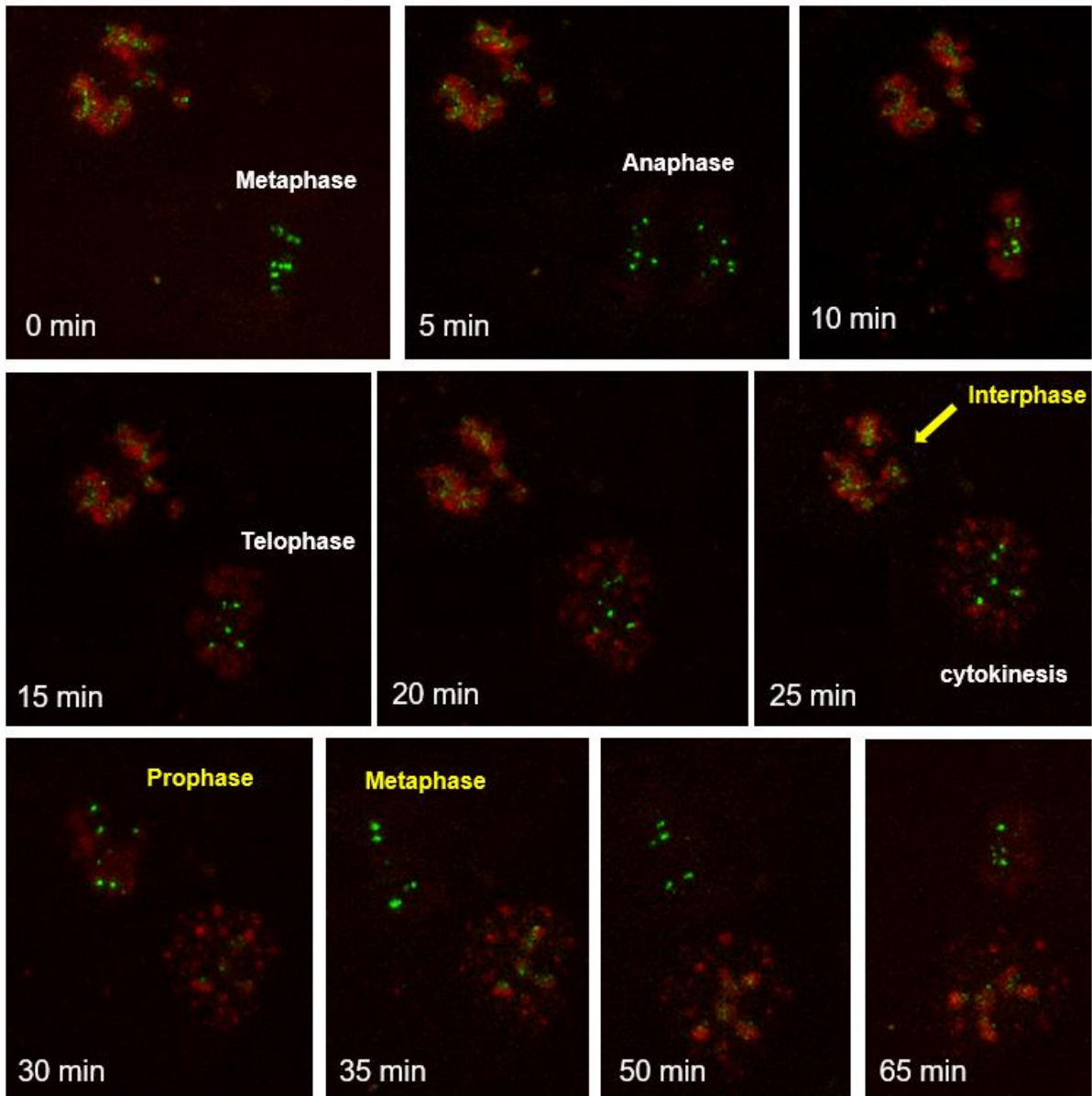


**Fig. 5.8** Image of fixed and DAPI stained UBFmAG-Nop52Halo-hTERT-RPE1 knock in cells, showing the localisation of the FC (UBF green) and the GC (NOP52 red) nucleolar sub compartments within the nucleus, stained by DAPI (blue). The bar is indicating 15 microns

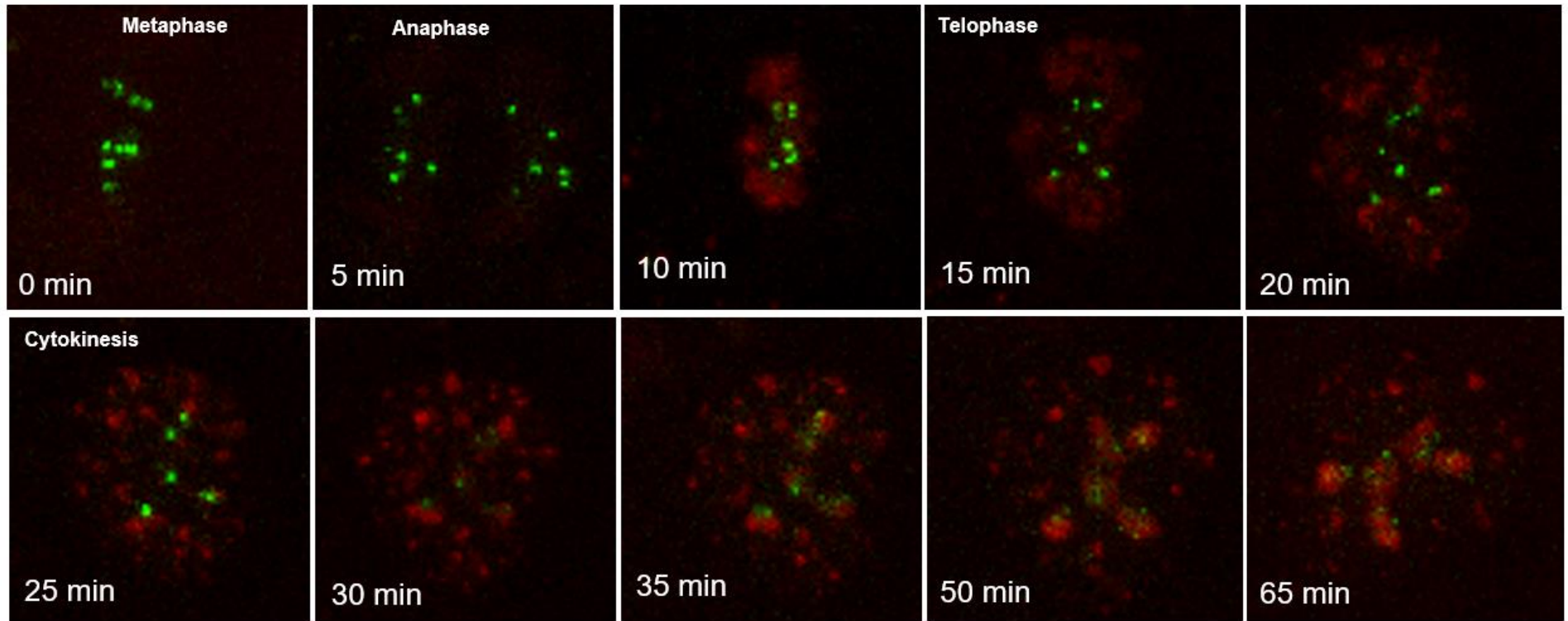
Another perspective to nucleolar formation offers the UBFmAG-Nop52Halo-hTERT-RPE1 cell line. With Nop52 as a tagged protein and its role as a GC marker we are able to visualize the nucleolar behaviour directly. As previously described UBF caps

indicate cells which go through specific cell cycle steps, such as pro-, meta- and anaphase. In the following figure (Fig. 5.9) we can see nucleolar formation, but also nucleolar disruption. Both events are precisely timed and take place within minutes.

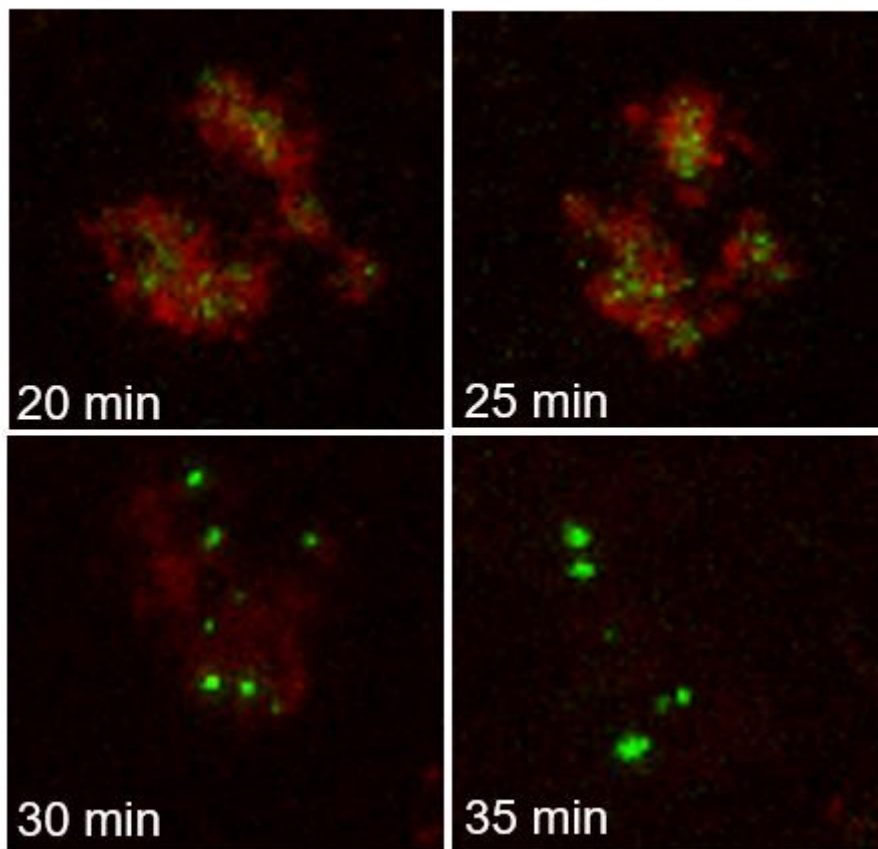
A metaphase cell can be used again as a start point (see Fig. 5.9 timepoint 0 min). 5 minutes later the cell is entering anaphase. Here we also can see again an obvious difference between the cap number and their intensity between the metaphase cell and the forming daughter cells (see Fig. 5.9 timepoints 0 and 5). The following timing is similar to the UBF-mAG, Fib-mCherry double tagged line and is shown in Fig. 5.9 and Fig. 11. The UBF foci split up to form multiple FC/DFC units, shortly after that Nop52 is getting recruited and forming the GC compartment around these units. Also, like UBF-mAG, the Nop52-Halo tag has a useful dual function. It is a GC marker for the nucleolus and based on that it is visualizing the full volume of the nucleolus itself. Because it gets included in peri chromosomal bodies (PNBs) during specific stages of the cell cycle (see Fig 5.9 and 5.10, timepoints 10-65) <sup>156</sup>. These PNBs are distributed over the whole volume of the nucleus during the early stages of the cell cycle, like ana- and telophase. Therefore, it is hard to see what a small arising nucleolus or a PNB is. However it is obvious at a later stage that the nucleolus is growing in size and signal intensity (see Fig. 5.9 and 5.10, timepoints 30-65). These observations overlap with experiments based a Gfp-Nop52 expression construct in HeLa cells <sup>156</sup>.



**Fig. 5.9** Illustration of the formation event of the FC and GC nucleolar sub-compartments in hTERT-Rpe1-UBFmAG-Nop52Halo Knock-In cells. UBF (in green) foci can be observed for meta-, ana- and early telophase. (Just one daughter cell is shown here). When rDNA transcription reappears, Nop52 (in red) is getting recruited from PNBs and nucleoli start to form. Nucleolar dismantling also takes place and can be observed at the time point 30 min. The different time points (min) are indicated for every image in white.



**Fig. 5.10** Closer look on the formation event shown in the figure above (Fig. 5.9). The cell is going through different cell cycle steps, which take place during nucleolar formation (meta-, ana-, and telophase). The yellow arrow indicates the daughter cell which is shown in the following images. The single timepoints (minutes) are indicated as white numbers.

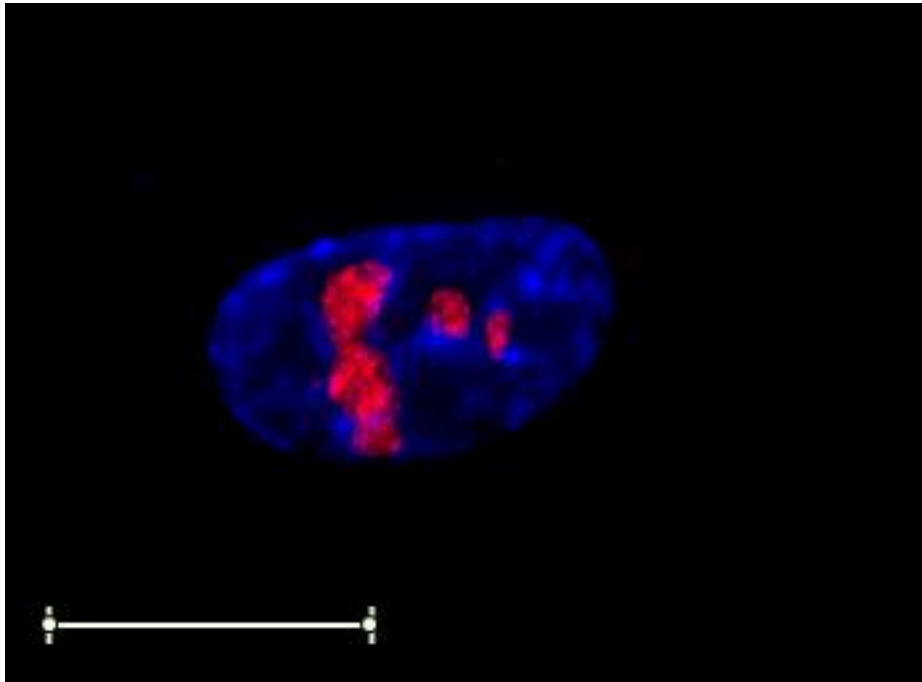


**Fig. 5.11** A closer look on nucleolar disruption as a cell cycle event. The image is an enlargement from Fig. 5.9. Cells enter prophase, the UBF-mAG signal fuses together, forming the characteristic caps. Nop52-Halo gets delocalized from the former nucleolus and can be observed as a weak background signal in the nuclear matrix.

However, we also can observe nucleolar dismantling, which takes place in late prophase. Active Pol-I transcription stops, and chromatin condensation follows. The UBF signal(s) fuses together and can be observed as intense green foci, whereas the typical nucleolar (Nop52) signal dismantles and can be seen as a red cloud like signal in the nuclear matrix (see Fig. 5.11 timepoints 20-35 min). From this point on, the cells go through metaphase, followed by anaphase and telophase. rDNA transcription gets reactivated during telophase and nucleolar formation takes place. The separation of the two daughter cells gets complete during telophase (cytokinesis) and the new cells enter interphase.

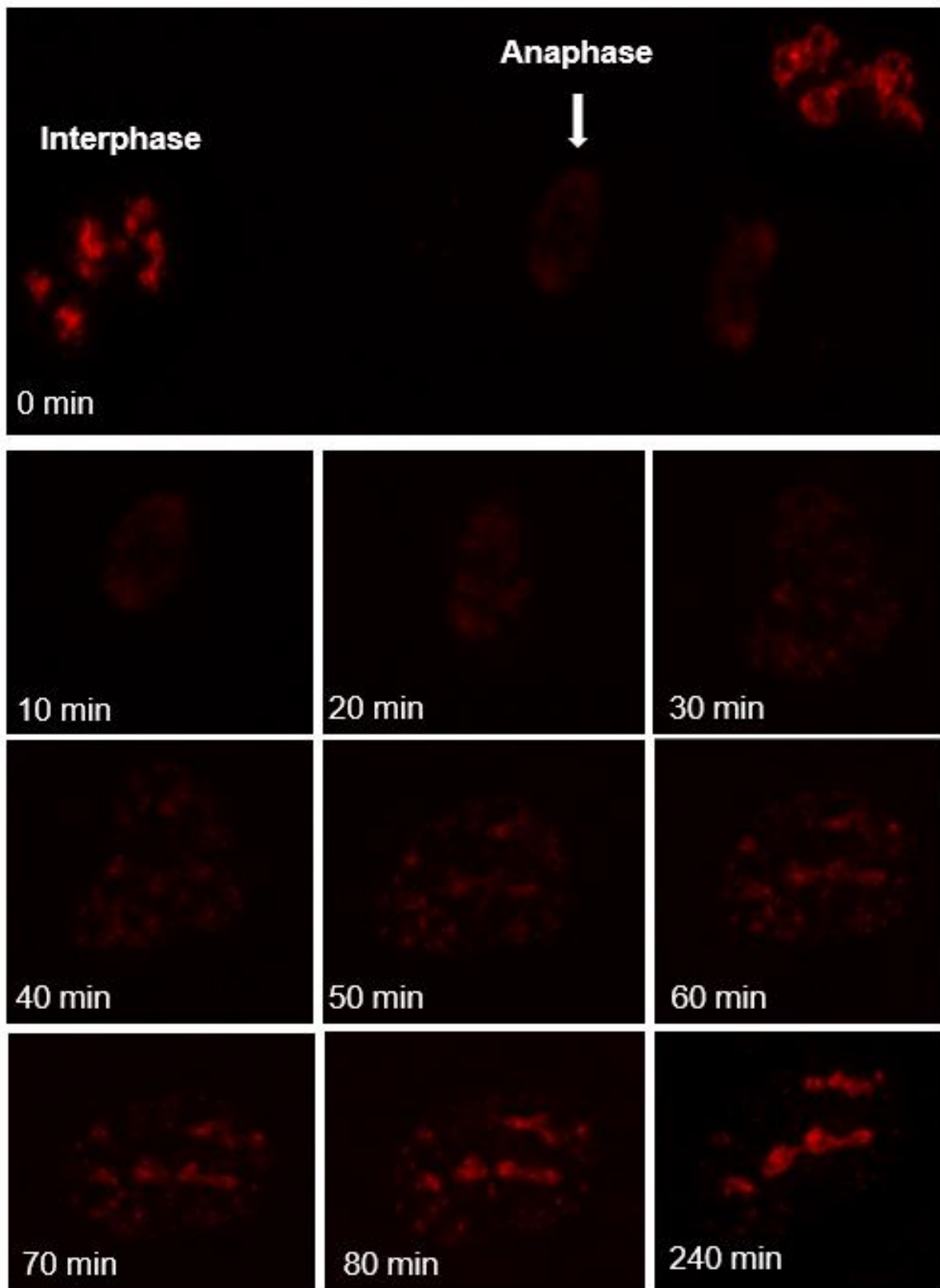
The described live cell observations could be observed multiple times during different live cell imaging sessions.

#### **5.2.4 Nop52 and its dynamic behaviour during the cell cycle.**



**Fig. 5.12** Homozygous Nop52Halo-hTERT-RPE1HKO cell Image. Nop52Halo showing the localisation of the nucleoli in red within the nucleus, which is stained by DAPI (blue). Bar is indicating 15 microns.

The use of the homozygous Nop52-Halo tagged cell line offers a more intense (TMRred) signal and therefore improved quality during live cell imaging. This cell line can be used to observe nucleolar formation and fusion events. A closer look, on the Nop52 localisation during ana- and telophase reveal that Nop52 is a part of the peri chromosomal compartment (Fig. 5.13: timepoint 0 and 10 min). As telophase progresses, PNBs start to form and can be seen in the images as more intense spots and areas. Also, nucleolar formation takes place around the decondensing transcriptionally active NORs, this could be seen in Fig. 5.9 and Fig 5.10 timepoint 35 and following. The nucleoli start to grow and during this process they recruit the required Nop52 from the PNBs (Fig 5.13 timepoints 50-240 min). The nucleolar signal intensity is increasing over the time, whereas the signal intensity of the PNBs is decreasing. This timing and the observed events are the same shown before with the UBFmAG-Nop52Halo-hTERT-RPE1HKO cell line.

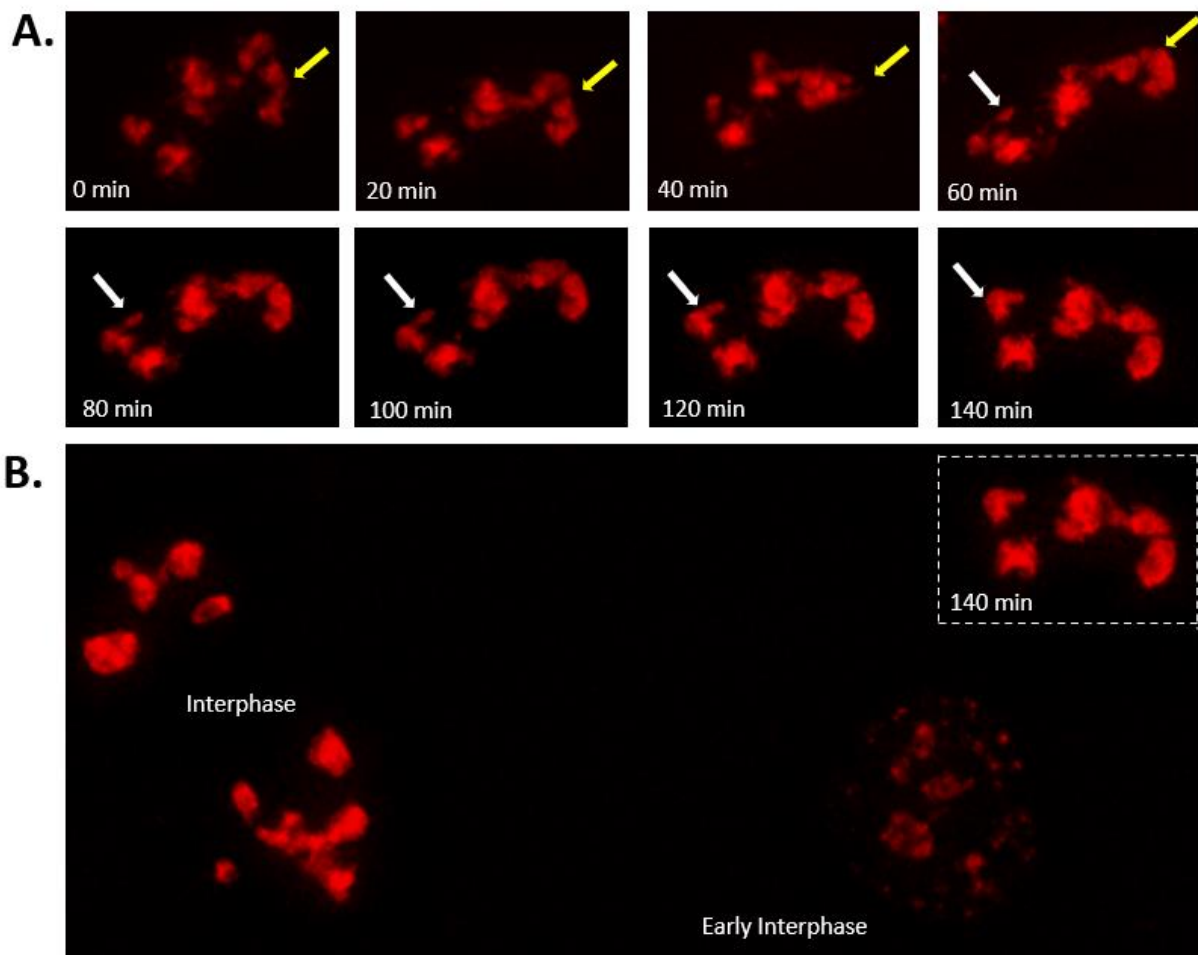


**Fig. 5.13** A closer look on Nop52 and nucleolar formation. Nop52 and its localisation during anaphase and telophase. Images were taken from a homozygous NOP52HALO-hTERT-RPE1HKO cell line. The observed Nop52 signal, and its function as a GC bookmark showed for the first time points a

homogeneous distributed signal in the peri chromosomal space. The yellow arrow indicates the observed daughter cell. Different time points in minutes are indicated for every image in white.

I would like to mention here that interestingly the Nop52-Halo-TMRed signal is stable over through multiple cell divisions over 4-5 days after the treatment for live cell imaging conditions.

However, it is also interesting to have a closer look on a nucleolus or nucleoli during interphase. These observations allow us to have a closer look for the dynamic behaviour of nucleoli during normal conditions (interphase).



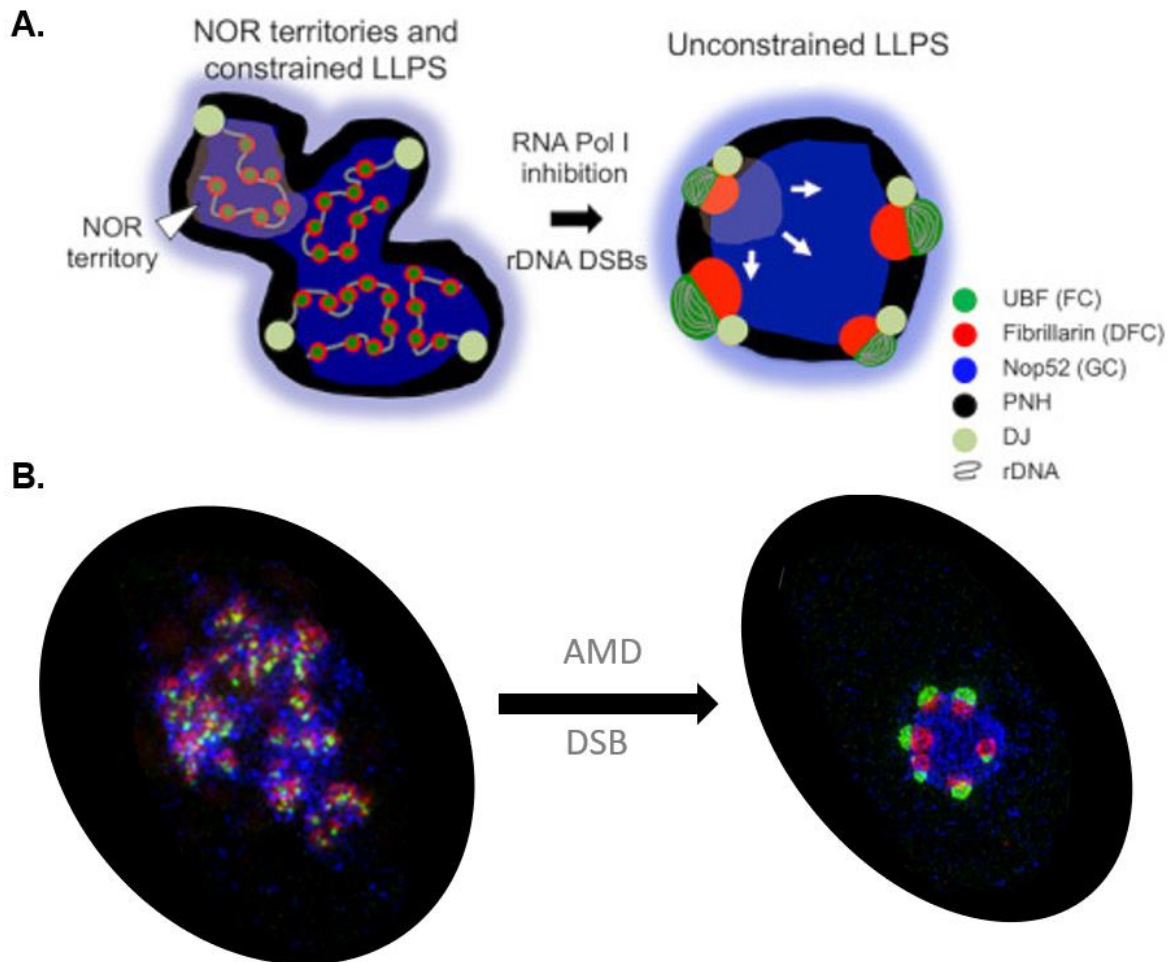
**Fig. 5.14** A closer look on Nop52 and nucleolar dynamics of a single interphase cell **A**. The yellow and white arrows indicate an accumulation or fusion event of nucleoli also a separation event can be observed (yellow arrow). **B**. whole observed field for the last timepoint, different cells can be seen. The images were taken from a homozygous hTERT-RPE1-Nop52Halo clone. The different time points (min) are indicated for every image in white.

One striking feature is the shape or the appearance of the nucleoli in hTERT-Rpe1 cells under normal conditions. They do not look like compact droplets; they look more like wispy clouds. Nucleoli in hTERT-Rpe1 cells are also highly dynamic, they can change their shape and they can fuse and separate from each other. This can be observed in the Figure above (Fig 5.14 A.). The dynamic events are indicated by a yellow and white arrow. The yellow arrow indicates a shape changing and fusion event, two nucleolar compartments aggregate with each other and create a new compartment. The white arrow points out a classical fusion event, a small nucleolar piece fuses together with a bigger nucleolus. In the last image (Fig. 5.14 B) we can see the observed cell in relation to other cells, especially one cell which in early Interphase building up its nucleoli. Some of the Nop52Halo is still localized within the PNBs, indicating the area of the nucleus. These observations overlap with former experiences based on the expression of suitable fusion constructs, encoding a FP and Nop52<sup>156</sup> and on IF images<sup>189</sup>. However, after a while the Nop52 nucleolar recruitment is over and the PNBs are no longer visible.

The observed nucleolar shape and the dynamic events during interphase could be observed on a regular base during live cell imaging. The results shown are representative.

It is striking that the shape of the nucleolus or the single nucleoli does not show a droplet-spherical like structure. However also the nucleolar size is depending from the number of active NORs (5 or 10 acrocentric chromosomes (diploid cells), indicated by the presence of UBF), which can vary depending on the cell cycle stage. These typical forms can be observed as part of the stress response, which will be described in the next section.

### 5.3 Nucleolar stress response of the FC/DFC and FC/GC compartments



**Fig. 5.15** Model for nucleolar organization in normal growth conditions and under nucleolar stress reorganization, this can be induced by inhibiting Pol-I transcription or by rDNA double strand breaks. **B.** IF images of hTERT-RPE1 cells, UBF in green, FIB in red and Nop52 in blue. Figure adapted from <sup>135</sup>.

Nucleolar stress response is resulting in the cap formation of the FC/DFC compartments and in a spherical shape of the nucleolus itself (GC compartment). This morphological reaction allows the DNA repair machinery to get access to the rDNA. The stress response can be triggered by the inhibition of the Pol-I transcription machinery, or by rDNA double strand breaks. Based on Immuno-FISH experiments showing that each FC/DFC caps has a single DJ signal (see Fig. 5.15) it was proposed that they represent individual NORs<sup>54</sup>.

The following experiments show the kinetics of nucleolar segregation/cap formation in live cells. This was achieved using the two double tagged cell lines UBF-mAG/Fib-mCherry and UBF-mAG/Nop52-Halo.

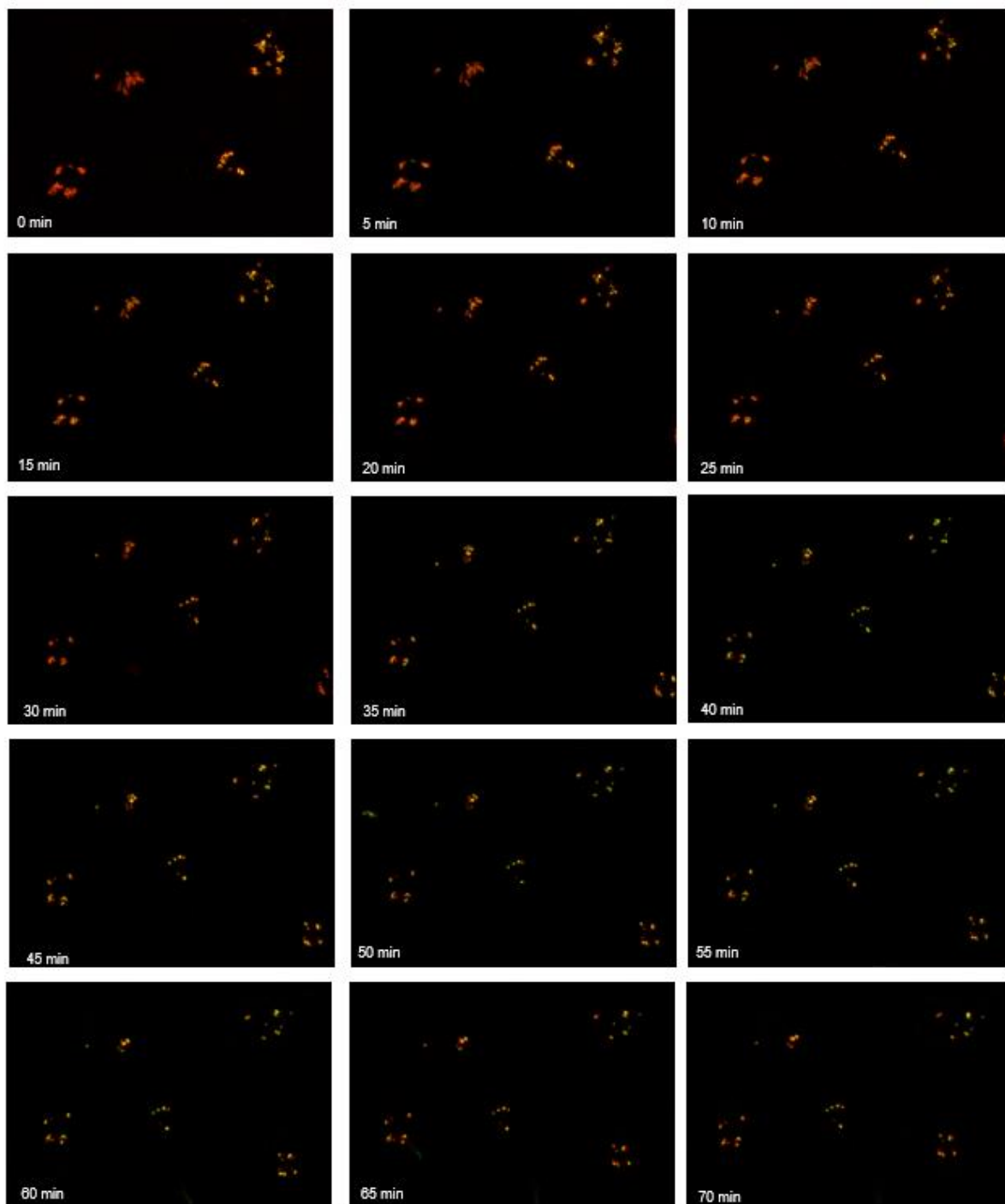
To induce a stress response in the tagged cell lines, I chose to use a low concentration of ActinomycinD (ActD or AMD) (50ng/ml). AMD is a Pol-I inhibitor and catalysing nucleolar cap formation. AMD is an intercalating agent with selectivity for guanine and cytosine nucleotide rich sites of DNA, at low doses. Higher doses would even inhibit DNA-replication. In the following images (Fig 5.16 and enlarged in Fig. 5.17) we can see the behaviour of the FC/DFC sub compartments, by using the double tagged UBF-mAG/FIB-mCherry (UBFmAG-FIBmCh-hTERT-RPE1HKO) cells.

The first image in all following figures (Fig. 5.16-5.20) shows an untreated interphase cell. Directly after the (first) image was kept, AMD was added.

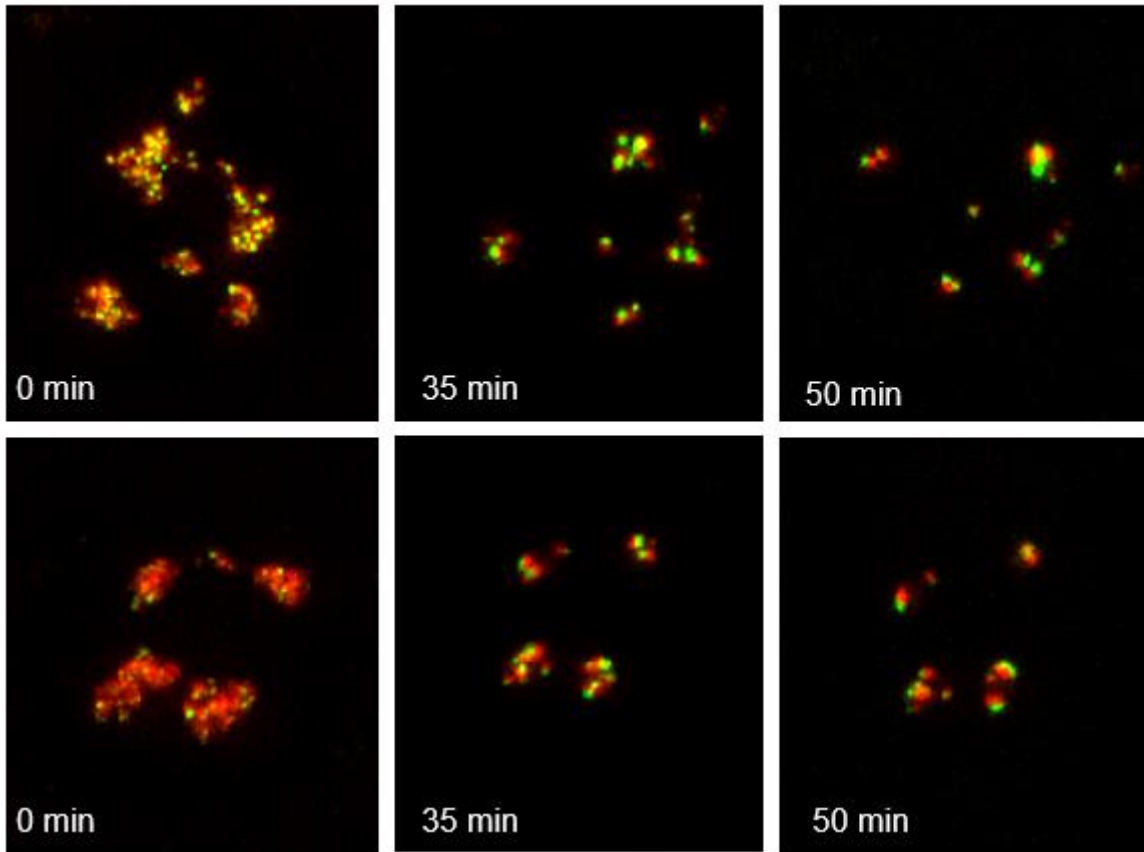
As soon as the reagent has been added, the cap formation begins. UBF-mAG (FC) fuses together and is moving to the periphery of the nucleolus. The cap formation itself takes around 50 min (at 50ng/ml AMD). FIB-mCherry (DFC) is following this behaviour and is forming together with UBF-mAG the typical bipartite FC/DFC caps, localizing at the edge of a usually spherical nucleolus. In figure 5.16 four different cells show this typical reaction after the AMD treatment. It can be also seen in more detail in figure 5.17. As previously mentioned in Chapter 4, there is no homozygous C-terminal Knock-In for FIB existing. This stress response was also one more verification step to check the function of the new created cell line.

However, the position and the dynamic behaviour of the nucleolus cannot be visualized with this cell line (UBFmAG-FIBmCh-hTERT-RPE1HKO).

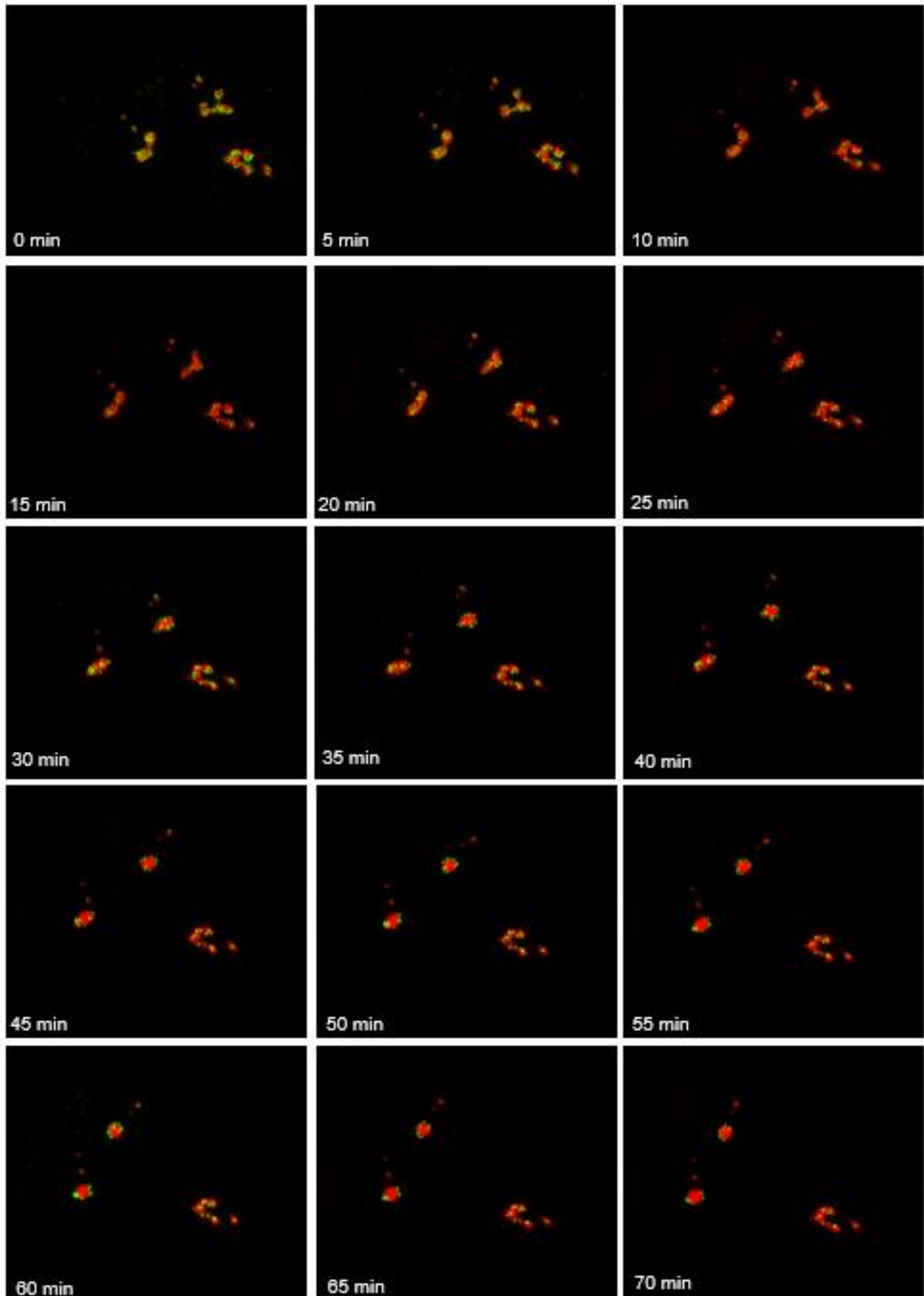
Therefore, the UBFmAG-Nop52Halo-hTERT-RPE1HKO cell line can be used because it indicates the behaviour of the FC and GC compartment. Here, we can observe that the UBF-mAG fusion takes place within the nucleolus, or to put it more precisely within the fluorescent 2-dimensional Nop52 signal. This can be seen for three different cells (see Fig. 5.18). For two of the cells the nucleus rounds up to a spherical droplet like structure (in more detail see Fig. 5.19). One other cell is forming caps, but the nucleolus is not changing its shape.



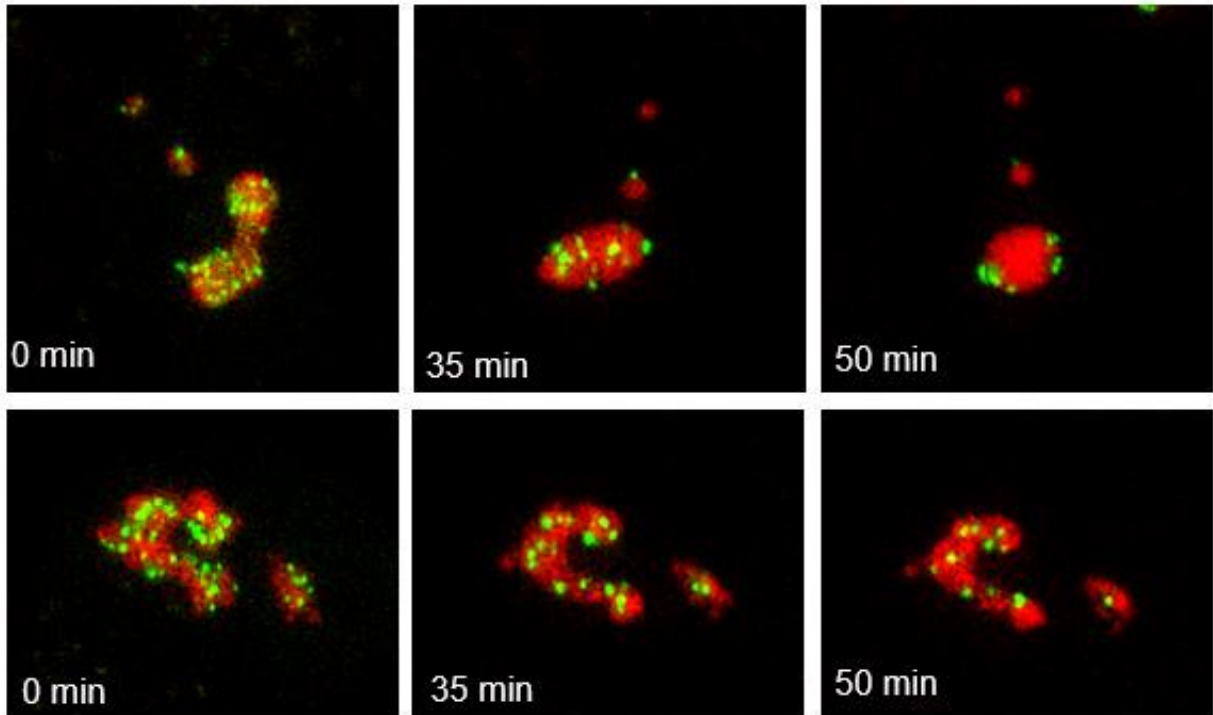
**Fig. 5.16** Cap formation in of four hTERT-Rpe1-UBFmAG-FIBmCherry cells. Time points for every image are indicated in white.



**Fig. 5.17** Enlargement of the nucleolar stress response in a single cell under AMD treatment for UBFmAG-FIBmCherry tagged cells, marking the FC/DFC compartments.



**Fig. 5.18** Cap formation in three different hTERT-Rpe1-UBFmAG-Nop52Halo cells. Time points for every image are indicated in white.



**Fig 5.19** Enlargement of the nucleolar stress response in a single cell under AMD treatment for UBFmAG-Nop52Halo tagged cells, marking the FC/GC compartments.

#### **5.4 Discussion and future perspective.**

The fact that UBF binds across the rDNA arrays offers the possibility to use it as an indirect marker for specific cell cycle stages. It is forming characteristic caps when the chromatin is condensing during cell cycle stages like pro-, meta-, ana- and telophase. However, when chromatin is switching back into an uncondensed state the UBF caps are dissolving and form the typical FC/DFC units within the nucleolus. Therefore, UBFmAG has the useful role to indicate start point for live cell imaging sessions, observing nuclear formation.

During the activation of Pol-I transcription the recruitment of FIB, with its role in early rRNA processing (DFC) takes place and is followed by other factors, responsible for later rRNA processing steps. FIB is part of the SSU (terminal ball) and is also responsible rRNA methylation as a core component of box CD snoRNPs.

NPM is used as a classical GC or nucleolar marker because of its role in the late processing of rRNA. However, the exact role in the nucleolus is still unknown. As part of this thesis Nop52 was used instead of NPM. NPM is often used as a “classical” nucleolar GC marker (cell lines Allen Institute). However, it fulfils other roles in centromeres and as a histone chaperone <sup>204</sup>. For this reason, NPM should be questioned as its classical role as a GC marker.

Based on Nop52 I was able to create a fully functional (homozygous) Knock-In. Currently there is no homozygous NPM Knock-In known. This GC (Nop52) tag is crucial to observe the “real” behaviour of the nucleolus inside a normal human cell, such as hTERT-RPE1. The intelligent tagging strategy, using suitable FC, DFC, and GC markers allowed me to observe nucleolar behaviour from different perspectives. The UBF and its role as a FC marker showed, because of the cap formation condensed chromatin, based on this it was possible to observe nucleolar formation. FIB however with its role as a DFC maker, based on the early processing of new transcribed rRNA could be also used as an indicator for the start of transcription. Nop52 however with its role as a GC marker also showed the based on the localisation within the PNBs the size of the cell nucleus during specific stages of the cell cycle <sup>156189</sup>. However, based on the observed timing for the two cell lines and the labelled nucleolar sub compartments it can be concluded that the recruitment of FIB indicates the activation

of Pol-I. Shortly after, the nucleolar formation of the GC compartment, with Nop52Halo as its maker, takes place. A full perspective about the nucleolus, the three sub compartments and the exact timing could deliver a triple tagged cell line with UBF, FIB and Nop52 as targets.

UBF and Nop52 were useful in a multifunctional way, of course they are indicating their nucleolar sub compartments (FC/GC) but they also indicate indirectly the chromatin stages. UBF by cap formation and Nop52 as part of the PNBs. Prominent chromatin dyes for live cell imaging, like SiR-DNA from SPIROCHROME do not work with hTERT-RPE1 cells. Even with just 10% of the lowest recommended concentration, the cells get stuck during anaphase and interphase cells show after 3-4 hours cap formation as part of the stress response. These observations do not really surprise, if we think about the function of the imaging mechanism of this dye. It is building up its fluorescent abilities in the far red when it intercalates between the single DNA strands. However, this dye got prominent for live cell experiments with different cancer cell lines, but it should be kept in mind that this dye is influencing the genome stability and the transcription activity. Based on the described mechanism it is not surprising that this dye is not working for normal cells such as immortalized and primary cell lines <sup>205</sup>.

An alternative to visualize chromatin could be the use of tagged H2B, encoded on an expression construct. However, the expression of tagged H2B also influences the cell cycle and so the doubling time. The best way to solve this problem of the direct vitalization of chromatin would be to tag a “chromosome coating” protein. Ki67 is one of these proteins, and based on the developed tagging strategy other members of the McStay lab, were able to create a C-terminally tagged homozygous Ki67 Knock-In.

The partial observation of the FC/DFC and FC/GC formation showed that it is a well-organized and timed process. This agrees with already published results, based on the transiently expression of nucleolar key proteins in different cancer cell lines <sup>156,203,206</sup> or even in other model organisms <sup>162,200 150</sup> . It can be also stated that single NORs form single nucleoli, which aggregate/fuse together during the cell cycle. These NORs occupy a territory within the nucleolus <sup>135</sup>. This is based on experiments with mouse hybrid cells, which are including single human (acrocentric) chromosomes. These single human chromosomes can be tagged by CRISPR/Cas9 and shuttled back into

human cells. Based on this tag it is possible to visualize single NORs by a technique called SABER FISH (signal amplification by exchange reaction) <sup>135</sup>. The single NORs could be observed as UBFmAG foci during the live cell imaging, the caps can be observed during pro-, meta-, and anaphase as well as during the stress response as caps at the edge of a nucleolus. However, the observation of single NOR (UBFmAG caps) is also a question of the resolution of the used imaging system. High resolution microscopy could be a good complement technique for future imaging. Also, the created Ki67-AID-mCherry-P2A-Blast Knock-In could help to answer the question of NOR territories within an (interphase) nucleolus and their role in nucleolar dynamics and formation.

It is striking that nucleoli in hTERT-RPE1 cells are not “droplet-like” they look more like clouds. Also high resolution IF images of a hTERT-Rpe1 nucleolus shows the same impression <sup>135</sup>. The defined indicators for the LLPS mechanism are also present, fusion events could be observed <sup>170</sup>. The typical droplet structure is not present in normal proliferating cells. The typical spherical droplet shape is based on inactive Pol-I transcription. But even when active transcription is suppressed, indicated by FC cap formation on the edge of the nucleolar periphery, not all cells show the typical droplet structure (Fig. 5.19). As part of my lab work it was also possible to isolate single nucleoli from the tagged cell lines, these isolated nucleoli can be embedded in low melting agarose or fixed on poly-lysine slides and still show a good visible fluorescent signal. Future experiments based on these isolated nucleoli could address the issue of cap formation in response to transcriptional inhibition. In other words, is nucleolar segregation intrinsic to nucleoli. It should be also mentioned here that the resulting cap formation based on a low concentration treatment with AMD (50ng/ml) is reversible. It is possible after the removal of AMD in the imaging media that the cap formation is decondensing resulting in normal proliferating cells.

The AMD treatment revealed that the created FIBmCherry Knock-In is functional because it is forming the typical caps. It also showed in the UBFmAG-Nop52Halo cell line that most of the nucleoli start to form spherical round structures. These observations prove that active Pol-I transcription is influencing the dynamic behaviour of the nucleolus. This classical and typical expected nucleolar shape (spherical) could only be observed as a consequence of transcriptional inhibition. The round nucleolus would represent the perfect shape with respect to thermodynamics that depends

entirely on a LLPS mechanism. A round liquid structure/droplet is the perfect and lowest energy state a liquid can get. However, it was possible to observe a “viscous” droplet behaviour.

Most cells of the UBFmAG/Nop52Halo tagged cell line showed this behaviour during AMD treatment. The nucleolus starts to round up during AMD treatment. the UBFmAG signal(s) gather together within the nucleolar interior (see Fig.5.20) and move then to the nucleolar periphery next to their DJ anchor in perinuclear heterochromatin (see Fig. 5.15) <sup>135</sup>. This led to the conclusion that the nucleolus can round up, as soon as the NOR territories (FC/DFC units) are leaving the nucleolar space and would also agree with observations done in oocytes from *Xenopus laevis* (see Fig. 5.21). This would also support the idea of LLPS behaviour.

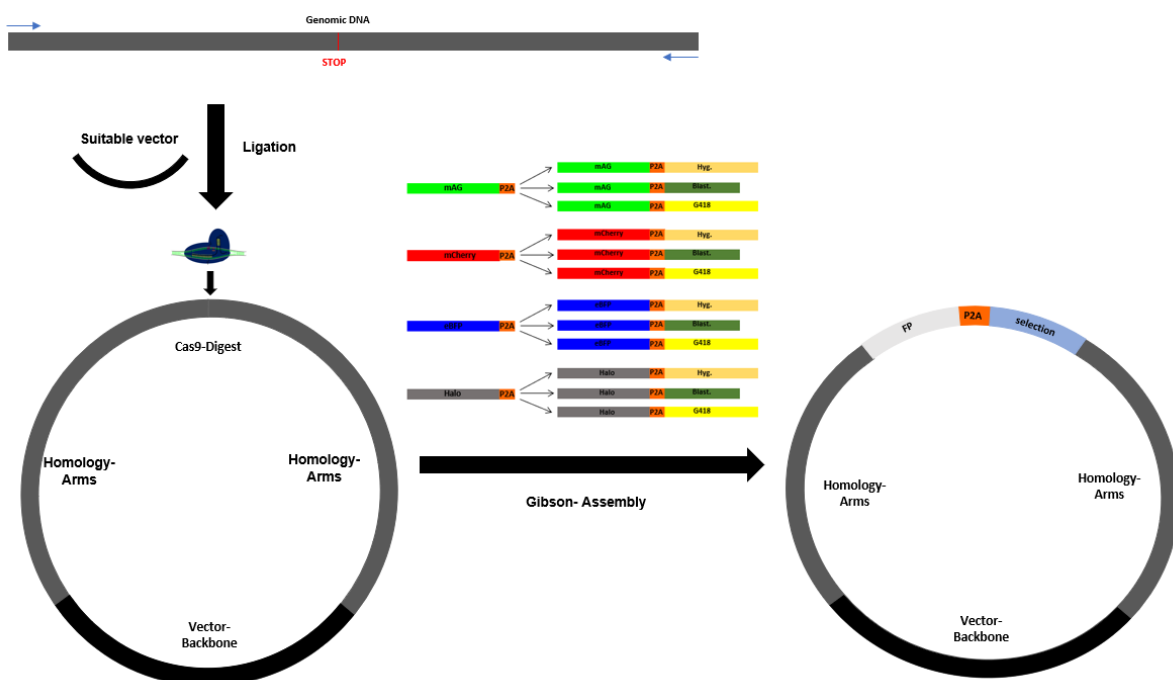
Future experiments based on the created KI67-AID-mCherry-P2A-Bast Knock-In in combination with tagged Nop52 could give a deeper understanding of the nucleolar behaviour *in vivo*. Interesting results in the aspect of *in vitro* behaviour could deliver isolated nucleoli based on the developed cell lines. However, it can be stated that the endogenous tagging of relevant genes in normal cell lines are the best way of targeting question related to cell behaviour and relevant highly discussed mechanisms such as LLPS.

# Chapter VI

## 6. Final Discussion

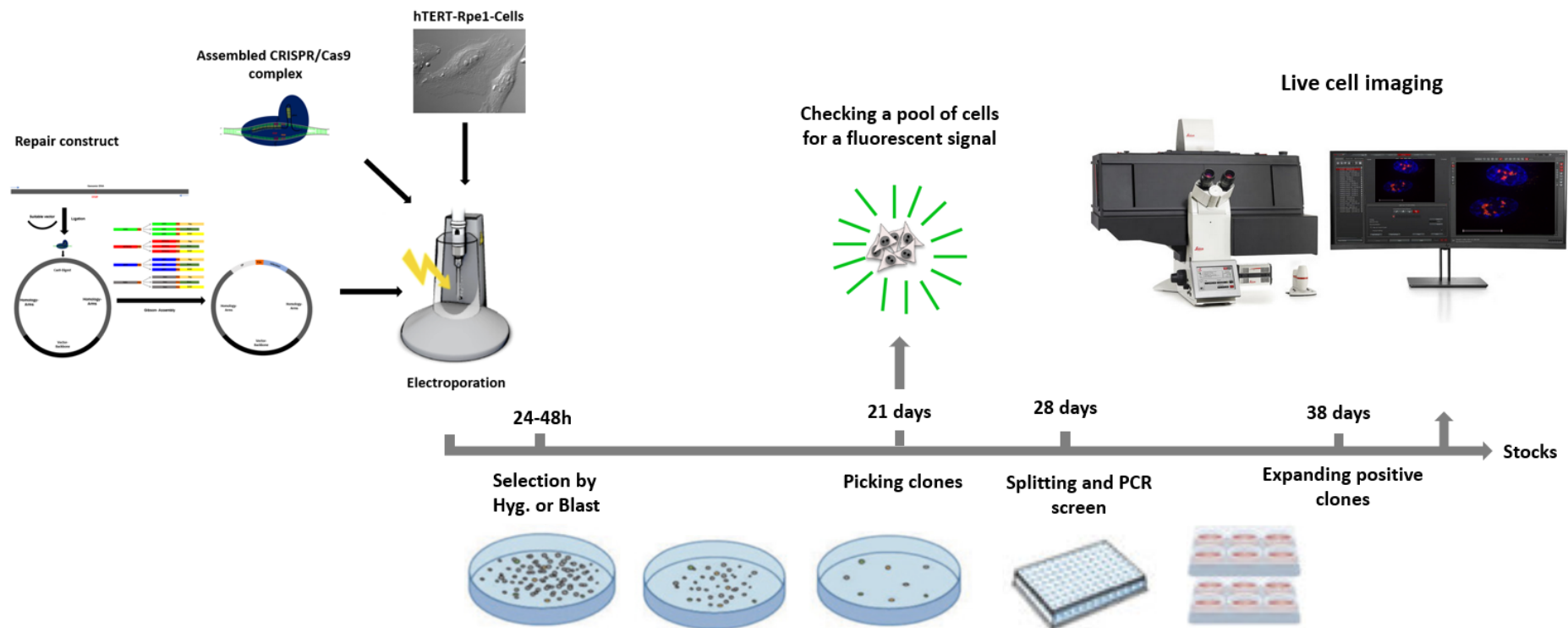
### 6.1 The development and application of the CRISPR/Cas9 system in hTERT-RPE1 cells, a non-transformed cell line

The work described in this thesis represents a significant methodological advance in CRISPR/CAS9 FP-tagging of endogenous genes in a previously refractive key human cell line, hTert-RPE1. This cell line is based on retinal pigment epithelial cells from a human female individual and was immortalised using hTert<sup>179</sup>. hTert-RPE1 cells are non-transformed with a normal and stable diploid karyotype. They are a commonly used model system for studying cell cycle progression as they maintain normal checkpoints and contact inhibition of growth. Furthermore, they are very well characterised with respect to their NORs<sup>51</sup>. The low intrinsic HR activity in these necessitated the development of an advanced CRISPR/Cas9 knock-In methodology. The key feature is connecting integration of targeting constructs at the appropriate genomic position with expression of a selectable marker. This linkage is achieved through the use of viral P2A sequences. In principle, two selection strategies are possible: Tagged cells can be positively selected, for example by use of the pMACSystem (Milentyl) (see chapter 3) or alternatively, non-tagged cells can be removed through the use of antibiotics like Blastidicin (Blast.), HygromycinB (Hyg.) G418, Zeocin or Puromycin.



**Fig. 6.1.** Developed cloning strategy to generate repair constructs for C-terminal CRISPR/Cas9 tagging. Genomic DNA can be used to amplify homology arms of around 0.5-2 kb by PCR (primers in blue) at each side of the STOP codon. The PCR product can be cloned into a suitable vector by ligation. The new generated construct can be linearized by a specific Cas9 digest, using a suitable gRNA for the target of choice. The linearized homology-arm construct can be combined with any knock-In cassette of choice by Gibson-Assembly. These cassettes are the key component and consist of different fluorescent proteins or tags like mAG, mCherry, eBFP and HALO. These FP proteins can be combined, by Gibson-Assembly, with every selection marker system of choice by a ribosomal skipping sequence (P2A). Selection marker systems of choice were Blast, Hyg. and G418 but also other antibiotics like Zeocin or Puromycin can be used. Also, the pMACSystem for isolating tagged cells by using magnetic beads targeting the pMACSystem-surface protein is possible.

Targeting constructs in which the FP tag and the selection strategy are linked represent the key component for successful tagging in hTERT-Rpe1 cells. The strategy I have developed allows a high degree of flexibility with respect to choice of target gene and the variety FP/selection combinations. Targeting constructs can be produced with minimal cloning steps (see Fig. 1). Intact targeting constructs are co-transfected together with the Cas9/gRNA RNPs into hTERT-Rpe1 cells. Selection can be performed 24-48 h after the transfection, and clones selected after 3 weeks. Various verification steps can be performed to ensure that successful tagging was achieved. These include PCR, western blotting and sub-cellular localisation of the FP tagged protein. Critically, hetero- or homozygous tagging need to be distinguished. Homozygous tagging is a good measure of full functionality of the tagged protein. Post-verification, live cell imaging or other cell-based experiments can be performed (see Fig. 2). While this tagging strategy was developed for and deployed in hTert-RPE1 cells, I have full confidence in the utility of this strategy for tagging in primary human cell lines and in human induced pluripotent cells. In the case of primary cells which have a limited replicative lifespan use of puromycin and blasticidin would be ensure rapid selection of positive clones. Indeed, I envisage it possible to carry out simultaneous tagging of two different targets.



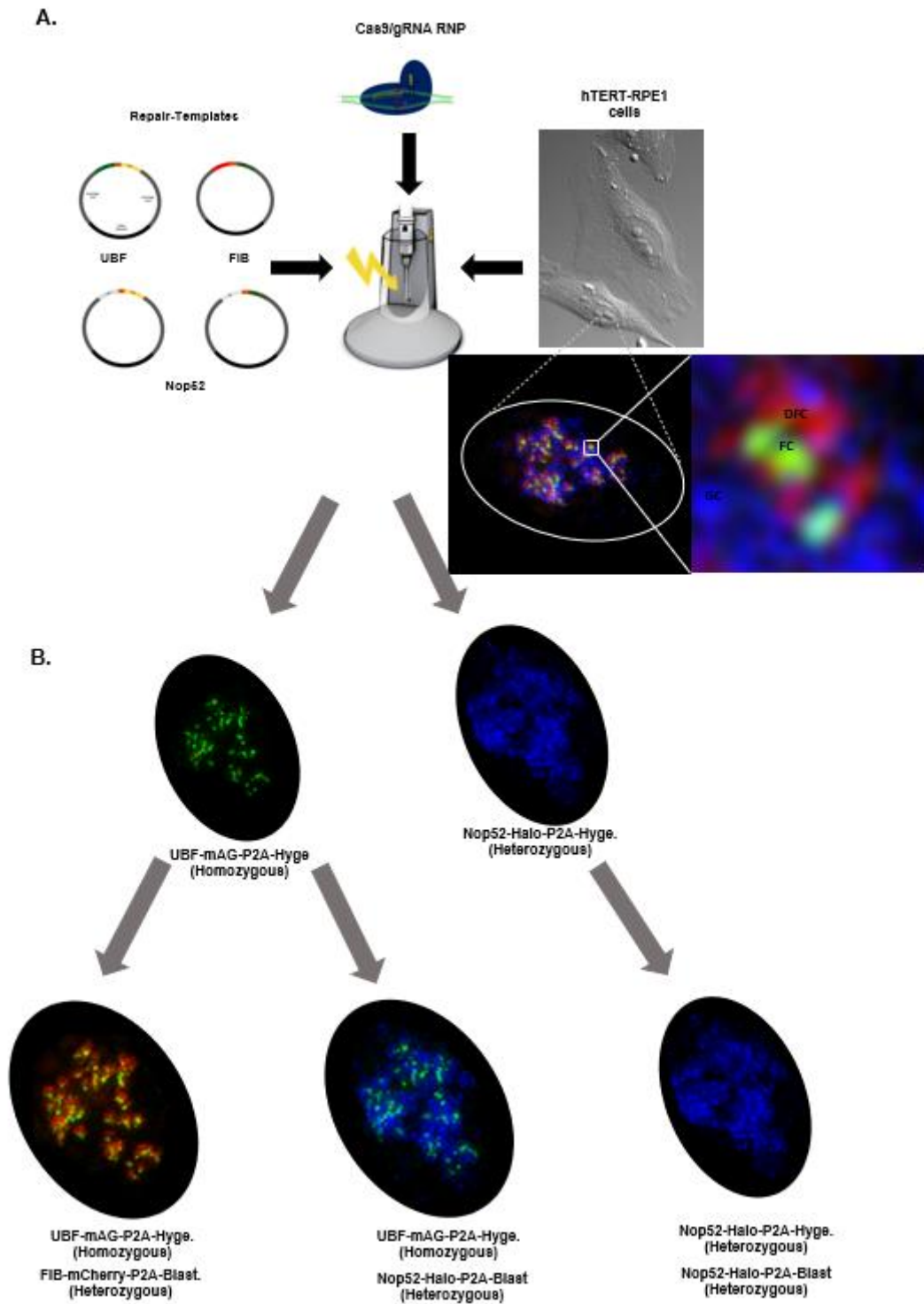
**Fig. 6.2.** The principal workflow of generating different hTERT-Rpe1 tagged cell lines, based on the developed method. The generated repair templates, which are they key component for this method, are transfected together with assembled CRISPR/Cas9 complexes into hTERT-Rpe1 cells. After 24-48 hours these cells can be selected with one of the marker systems over a time of 2-4 weeks. The resulting clones normally show the expected fluorescent signal, produced by a successful FP tag of the gene. These clones have to be verified by PCR and Western-Blot (WB). After these verification steps stocks can be produced and live cell imaging can start.

By deploying the advanced tagging protocols that I developed during the course of my thesis work, I was able to generate 4 different cell lines marking the different nucleolar sub compartments in a variety of combinations: a UBF-mAG (FC), a UBF-mAG-FIB-mCherry (FC/DFC), a UBF-mAG-Nop52-Halo (FC/GC) and a single Nop52-Halo (GC) cell line. The generated tags for UBF and Nop52 were homozygous, the double tagged cell lines based on FIB-mCherry and Nop52-Halo were heterozygous.

The CRISPR/Cas9 system is much simpler and more reliable than previous nuclease systems such as TALENS or Zinc fingers. It is now a relatively well-established system that will be a standard technique in many molecular biology related laboratories in the coming years.

One institute that already distributes CRISPR/Cas9 tagged hiPSC cell lines is the Allen-Institute in Seattle. However, almost all cell lines offered are heterozygous, giving cause to doubt about the functionality of the tagged protein. Also, the design of the respective tags is not always optimal, an N-terminal Halo tag at UBF can serve as an example here. The N-terminal dimerization domain of UBF, as well as the tagging itself at this position can influence the gene expression as well as the dimerization and the folding of UBF. In my opinion large scale tagging projects place too much emphasis on the numbers of proteins tagged at the expense of assessing the functional impacts.

To date, CRISPR/Cas9 tagging of endogenous genes with FP's is the optimal method for observing the true behaviour of a protein in its normal cellular environment. With successful homozygous tagging, we can be confident in full functionality of the fusion protein. Thus, it is then possible to observe the natural behaviour of the target protein in its normal cellular environment and at normal levels of expression. In some cases, obtaining heterozygous rather than homozygous clones may be a reflection of targeting efficiency rather than loss of functionality. This eventuality can be overcome through the use of two targeting constructs that differ only in their selection cassette



**Fig. 6.3** General overview about the generated cell lines. **A.** Simple illustration of the single key components to generate tagged hTERT-RPE1 cell lines, to visualize the nucleolus and its sub compartments (FC, DFC, GC see zoom in). The single repair templates for UBF(FC), FIB(DFC), and Nop52(GC) are illustrated with the corresponding Knock-In cassettes. **B.** The generated tagged cell lines, based on the different nucleolar markers. A homozygous UBF-mAG-P2A-Hyge. was the first generated cell line and the base for the following double tags: heterozygous FIB-mCherry-P2A-Blast and heterozygous Nop52-Halo-P2A-Blast. To proof that Nop52-Halo is fully functional a single tagged

Nop52-Halo-hTERT-RPE1 was created, the first allele was tagged with a Halo-P2A-Hyge. cassette and the second allele with a Halo-P2A-Blast cassette, resulting in a homozygous Nop52-Halo Knock-In.

As previously mentioned, I was able to create a homozygous C-terminal UBF-mAG tag and a homozygous Nop52 Halo tag in hTERT-Rpe1 cells, the cell lines showed normal behaviour, which is indicating that the created tag is fully functional. The UBF-mAG based double tags, which are Fib-mCherry and Nop52-Halo were heterozygous. For the Nop52-Halo knock-In we know that this tag is also fully functional. In the case of Fib-mCherry we must have a closer look. The design is based on a template offered by the Allen-Institute. This institute is also selling different mono, double and triple tagged lines, all these cell lines have a heterozygous Fib-eGFP tag. My Fib-mCherry tag is also heterozygous and shows the typical expected behaviour. Fib is localizing the correct way during the different cell cycle stages and forms the characteristic caps at the nucleolar periphery, driven by Pol-I inhibition. Based on this observed behaviour during many live cell imaging sessions it can be concluded that Fib-mCherry is not fully functional. A future alternative could be the tagging of another snoRNPs component, which like Fib are located in the DFC compartment.

Using the approach, I have developed, other members of the McStay lab were able to create more CRISPR/Cas9 tagged cell lines. The Knock-In cassettes got combined with an AID sequence to knock-down any target protein of choice. Based on this, Ki67 could be tagged with an AID-mCherry-P2A-Blast knock-In cassette, resulting in a homozygous tag. Ki67 is a GC marker, and it plays a role in nucleolar formation and dynamics. Currently the McStay lab are testing the hypothesis that Ki67 is responsible for maintaining NOR territories within the nucleolus. An endogenous knock-down of Ki67 would lead to a different nucleolar behaviour because “separate” single NOR territories would no longer exist within the nucleolus. However, Ki67 also offers the possibility to vitalize chromosomal behaviour during specific stages of the cell cycle. Therefore, it can be partially used as a chromosomal dye.

As already mentioned in chapter 5, DNA live cell imaging dyes, such as SiR-DNA from SPIROCHROME do not work in hTERT-RPE1 cells. Also TCOF-I, a protein which is involved in rDNA transcription could be also tagged based on a mCherry-P2A-Blast knock-In cassette, the resulting cell line is homozygous.

Thus far I have used antibiotic selection to generate FP tagged clones, however, it is also possible to use linked cell surface markers to select tagged clones. I previously mentioned the pMACS system, which encodes a surface protein which can be used to isolate targeted cells using antibodies coupled to magnetic beads.

A recent publication based on a technique called surface engineered fluorescence assisted kit with protein epitope enhanced capture (SNEAK PEEC) could show a method which allows to select for homozygous tagged clones <sup>207</sup>. This technique combines CRISPR/Cas9 genome editing with cell-surface display. SNEAK PEEC links each targeted knock-in with the expression of a distinct exogenous protein epitope on the cell surface, which allows identification using fluorescent binders of each epitope. The system uses two DNA repair templates, each one for a single allele. Each template contains a left homology arm, a C-terminal tag, a self-cleaving peptide (2A), a cell-surface display epitope and a right homology arm. Based on these two different epitopes it is possible to select cells by FACS, using epitope specific antibodies. Homozygous tagged cells would show the presence of these two different epitope proteins on the cell surface. Based on this technique a specific selection for homozygous clones is possible. This technique was developed and performed in transformed HEK293-F cell.

## **6.2 Live Cell imaging short-term priorities**

The successful development of a CRISPR/Cas9 tagging technique and the resulting tagged hTERT-Rpe1 cells, allow us now to address aspects of “typical” nucleolar behaviour in fine detail.

Based on the created cell lines, the different sub-compartments of the nucleolus can be visualized, and their behaviour can be observed *in vivo* by live-cell imaging. This behaviour includes dynamic aspects of the nucleolus in cells and their behaviour during key stages of the cell cycle, specifically nucleolar disruption and formation.

The double tagged cell lines UBF-mAG/FIB-mCherry (FC/DFC) and UBF-mAG/Nop52-Halo (FC/GC) can show the behaviour of the different sub compartments. The tagged FIB can be used as an indicator for active transcription, whereas the tagged Nop52 indicating the volume of the nucleolus itself. UBF on the other hand and its role as a marker for the FC compartment can be also used to indicate different stages of the cell cycle, especially stages of condensed chromatin. The chosen marker proteins for the single nucleolar compartments are multifunctional and have thus proven to be very useful especially for live cell imaging and all other *in vivo* based experiments. The pre-Covid performed live cell imaging observations showed that nucleolar formation is a well-timed process and overlaps with findings made in cancer cell lines. These findings were based on the overexpression of suitable nucleolar constructs (discussed in Chapter 5).

The endogenous tagging of Nop52, with its function as a nucleolar GC marker, has never been done before. It revealed that the nucleolus is not a round compact droplet like structure rather it appears irregular in shape, reflecting its organisation into NOR territories<sup>135</sup>. Formation of such nucleoli involves association of multiple acrocentric p-arms and can be observed by live cell imaging, as shown in figure 5.9.

Nucleolar reorganisation, characterised by cap formation, initiated by reversible inhibition of Pol-I transcription can be observed by live cell imaging. Critical questions remain unanswered, not least what is the mechanism driving nucleolar segregation.

Two general models for nucleolar segregation can be envisaged, ‘intrinsic’ and ‘extrinsic’. In the ‘intrinsic’ model, rDNA and associated FC and DFC components are excluded from the interior to the surface of nucleoli. In the ‘extrinsic’ model these same

component are actively pulled from the nucleolar interior. While there is no definitive evidence for one model over another, there is some recent support for the 'extrinsic' model. When expressed in human cells the restriction endonuclease AsiSI can induce DSBs at thousands of sites throughout the genome, including one site in each rDNA repeat. Using a U2OS cell model expressing AsiSI, Marnef *et al* colleagues presented evidence rDNA break localization at the nucleolar periphery is not a direct consequence of transcriptional repression but rather is an active process involving the nuclear membrane. This model does not explain how nucleoli segregate in response to stresses that do not involve generation of rDNA DSBs such as pharmacological inhibition of RNA Pol I transcription by AMD or BMH21. Furthermore, the study is compromised both by the presence of DSBs throughout the entire nucleus, and by the fact the experiments were performed in a cell line with a highly unusual karyotype with respect to NORs<sup>51</sup>. I favour the 'intrinsic' model suggesting that is it driven by LLPS.

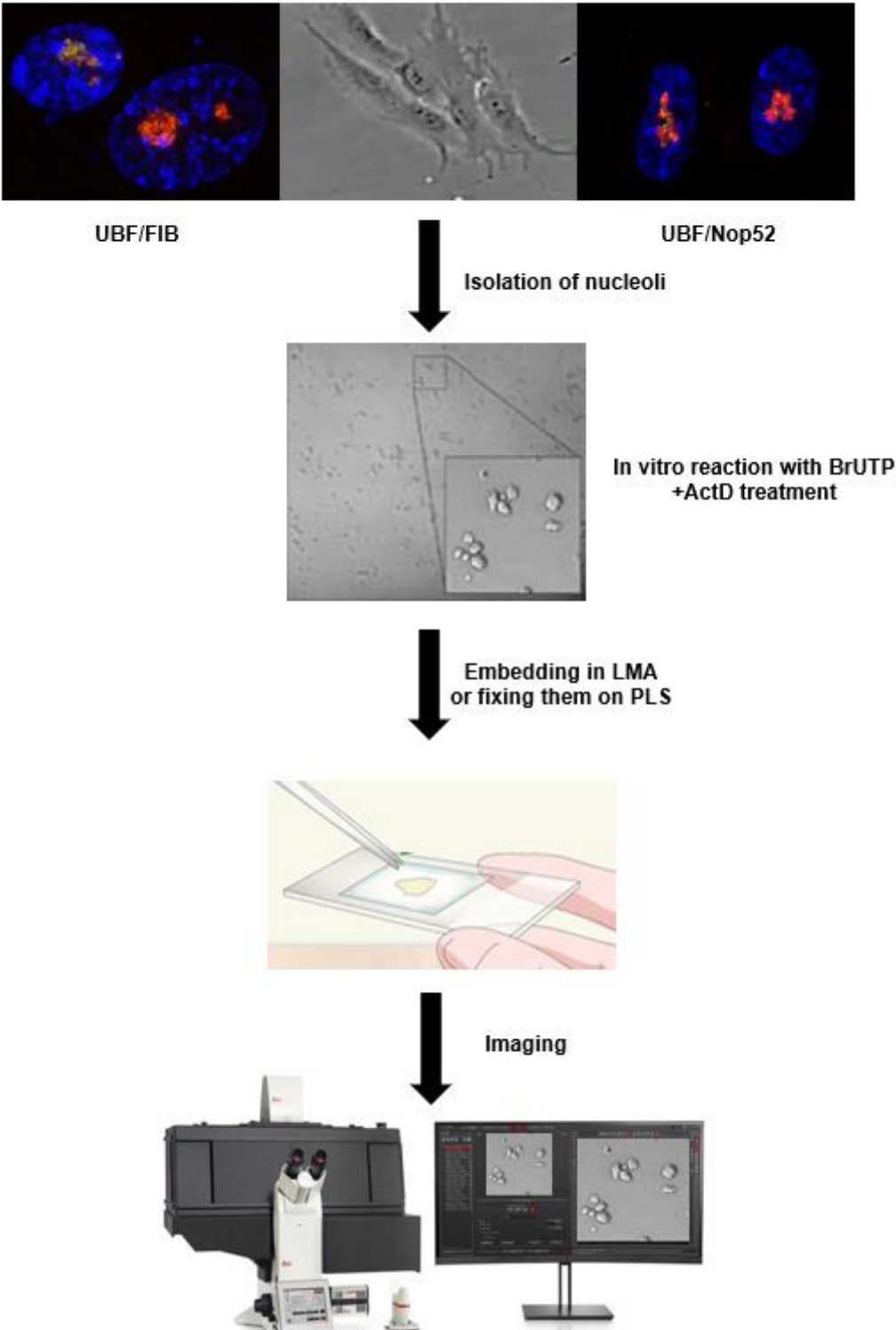
When cap formation takes place, the FC/DFC compartment move to the periphery of the nucleolus and form caps, one per NOR. Also the nucleolus itself change shape to a spherical droplet like structure. I believe that inhibition of transcription results in the gradual gathering of rDNA, FC and DFC components. Eventually all these component from a single NOR territory give rise to a single nucleolar cap located adjacent to its DJ anchor in peri-nucleolar heterochromatin. Notably, these caps are always bipartite in structure with FC components that by default face outward to the nucleoplasm and DFC components face inward the GC nucleolar interior. I believe that these characteristics are compatible with a view that the morphology of stressed nucleoli is dominated by LLPS. A detailed examination of nucleolar segregation in live tagged cells using super-resolution microscopy, should aid in differentiating between intrinsic and extrinsic models. In the following paragraphs I propose experiments with isolated nucleoli that could rule out extrinsic models for establishing and/or maintaining nucleolar segregation.

There are effective protocols for isolating nucleoli from human cells. Furthermore, these isolated nucleoli have been shown to support ongoing transcription of rDNA by RNA Pol I<sup>29</sup>.

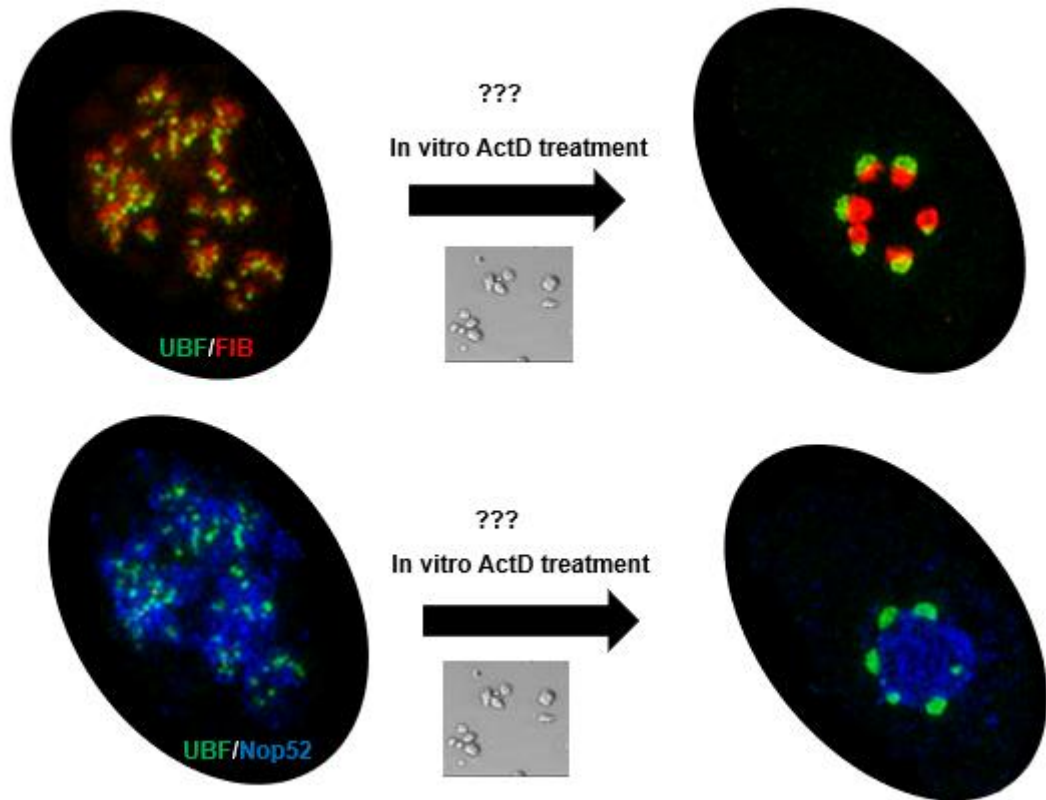
In my own experiments I have shown that nucleoli isolated from the double tagged cell lines (UBF-mAG/FIB-mCherry and UBF-mAG/Nop52-Halo) retain their fluorescence even when embedded in low melting point agarose. This raises the possibility of

visualising isolated nucleoli from AMD treated tagged cell lines to evaluate if maintenance of a segregated morphology is intrinsic to nucleoli. The more exciting prospect, outlined in figure 6.4, would be to inhibit ongoing transcription in isolated nucleoli and observe nucleolar segregation *in vitro*.

**A.** **hTERT-RPE1**  
**double tagged cells**



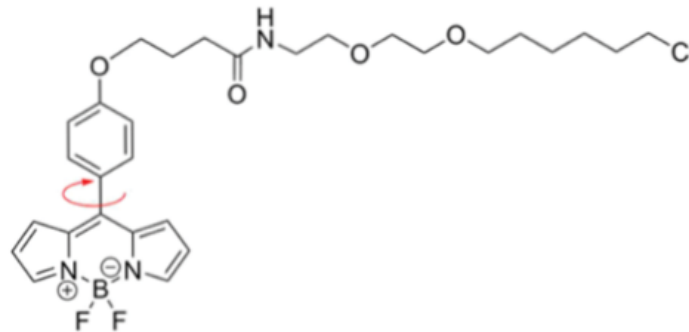
**B.**



**Fig. 6.4** Schematic procedure for in vitro experiments based on isolated nucleoli. **A.** The two double tagged cell lines UBF-mAG/FIB-mCherry as well as UBF-mAG/Nop52-Halo could be used for these experiments. After the isolation the single nucleoli can be used for in vitro reactions followed by an embedding in low melting agarose (LMA) to visualize them. As an alternative these nucleoli can be also fixed onto poly-lysined slides (PLS). **B.** imaging could reveal that nucleolar cap formation is active even in isolated nucleoli. Figure is based on hypothetical considerations

### **6.3 Measuring the biophysical properties of nucleolar compartments during stress**

An imaging dye based on BODIPY offers another useful technique for obtaining a deeper insight into single nucleolar sub compartments and their (biophysical) nature. This dye is called BODIPY-Halo Tag (see Fig. 6.5), it offers the possibility to measure cellular micro-viscosity based on the principle of rotor-based organelle viscosity imaging (ROVI).



**Fig. 6.5** The BODIPY-dye for the Halo-tag system. On the left side of the image the boron-diphyrrin (BODIPI) fluorescent dye with its indicated rotation axe, on the right upper part the Halo-binding ligand is indicated.

The rotation frequency of this specific fluorescent dye depends on the cellular environment or the “packing density” around the dye, this rotation frequency correlates with the fluorescent live time, based on a mathematical (Foerster-Hoffmann) model <sup>208</sup>. Therefore, it’s possible to make live-cell measurements based on the fluorescent live time of the dye, which has a direct correlation to cellular micro viscosity. The *in vivo* application of this method to normal transcriptional active nucleoli and to nucleoli after cap formation could be very revealing.

A second molecular viscosity sensor method is the so called “crowding” sensor <sup>209</sup>. The key principle is based on a Foerster resonance emission transfer (FRET) protein pair. The FRET principle relies on the distance dependent transfer of energy from a donor molecule/protein to an acceptor molecule/protein. A typical FP FRET pair would be mCerulean/mCitirine as donor/acceptor combination.



#### **6.4 Final sum up**

There is no doubt that the CRISPR/Cas9 system will become one of the standard technologies in the coming years. Genome editing will play an increasingly important role in both academic research and industry. Within the scope of this work, I was able to show that genome editing is possible without any problems in a former CRISPR “resistant” cell line (hTERT-Rpe1). Former *in vivo* experiments, targeting the behaviour of specific genes and their encoded proteins, were based on overexpression constructs in different cancer cell lines. It makes sense to think about whether cancer cell lines are still the appropriate model system for fundamental questions in molecular genetics. Even today, many CRISPR/Cas9 related publications are still being carried out in these cancer cell lines. The developed tagging system can be also used for any other cell line which is similar to hTERT-RPE1 cells (normal diploid karyotype etc.) like primary cells or pluripotent stem cells. The FP tagging of genes by CRISPR/Cas9 in hTERT-RPE1 cells allows the observation of real (endogenous) behaviour based on the own cell-controlled expression rate, in a normal non-cancer cell line. A homozygous Knock-In should always be the desired goal, as it guarantees the full functionality of the tagged gene. If the above points are met, it is the best and only way to observe natural and above all normal behaviour *in vivo*. Based on my developed method other genes such as KI67 and TCOF-I have been tagged by other members of the McStay lab. My developed cell lines can be used to answer the question of *in vitro* activity of nucleoli and the ability to form caps outside of the natural cellular environment. Nucleolar key genes can be tagged with any useful tag of FP to create cell lines for any kind of experiment. Halo for example is a very useful tag because it can be combined with different fluorescent dyes, but it also allows to use other ligands for pull down experiments. The single treatment with a Halo dye also allows conclusions about the lifetime of a protein in the cell. The combination of two or even three dyes and the successive use of them can reveal the turnover number of a Halo tagged gene. A UBF-Halo tag could be an interesting base for this kind of experiment.

One of the most relevant topic areas in the coming years will be probably a closer look on the LLPS hypothesis. The nucleolus as the biggest membrane less organelle is the perfect tool to answer this question. The endogenous tagging of nucleolar markers are the fundamental and first step to get a more detailed view about nucleolar behaviour especially in the aspect of the LLPS hypothesis.

I hope the tagging system I developed in the course of this work can make a useful contribution to this debate in the future.

## References:

1. Katoh, Y. *et al.* Practical method for targeted disruption of cilia-related genes by using CRISPR/Cas9-mediated, homology-independent knock-in system. *Mol. Biol. Cell* **28**, 898–906 (2017).
2. Miyamoto, T. The Microtubule-Depolymerizing Activity of a Mitotic Kinesin Protein KIF2A Drives Primary Cilia Disassembly Coupled with Cell Proliferation. *Cell Rep* **176**, 139–148 (2015).
3. Pederson, T. The nucleolus. *Cold Spring Harb. Perspect. Biol.* **3**, 1–15 (2011).
4. Boisvert, F. M., Van Koningsbruggen, S., Navascués, J. & Lamond, A. I. The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* **8**, 574–585 (2007).
5. Schwarzacher, H. G. & Wachtler, F. Nucleolus organizer regions and nucleoli. *Hum. Genet.* **63**, 89–99 (1983).
6. Montgomery, T. H. *Montgomery, T. H. (1898). Comparative Cytological Studies with Especial Regard to the Morphology of the Nucleolus. 1.1.2. (1898).*
7. Litt, M., Kay, E. R. M. & Dounce, A. L. ISOLATION AND PROPERTIES OF LIVER CELL NUCLEOLI \* ( From ~ he Department of Biochemistry , The University of Rochester School of Medicine , i ~ ne. 127–145 (1955).
8. Desjardins, R., Smetana, K., Steele, W. J. & Busch, H. Isolation of Nucleoli of the Walker Carcinoma and Liver of the Rat Following Nuclear Disruption in a French Pressure Cell. *Cancer Res.* **23**, 1819–1823 (1963).
9. Vincent, W. S., Baltus, E., Lovlie, A. & Mundell, R. E. Proteins and nucleic acids of starfish oocyte nucleoli and ribosomes. *Natl. Cancer Inst. Monogr.* **23**, 235–253 (1966).
10. BROWN, D. D. & GURDON, J. B. Absence of Ribosomal Rna Synthesis in the Anucleolate Mutant of *Xenopus*. *Proc. Natl. Acad. Sci. United States* **51**, 139–146 (1964).
11. Birnstiel ML, W. H. Localization of the ribosomal DNA complements in the nucleolar organizer region of *Xenopus laevis*. *Natl Cancer Inst Monogr.* **23**, 431–447 (1966).

12. Henderson, A. S., Warburton, D. & Atwood, K. C. Location of ribosomal DNA in the human chromosome complement. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 3394–3398 (1972).
13. Ritossa, F. M. & Spiegelman, S. Localization of Dna Complementary To Ribosomal Rna in the Nucleolus Organizer Region of *Drosophila Melanogaster*. *Proc. Natl. Acad. Sci.* **53**, 737–745 (1965).
14. MCCONKEY, E. H. & HOPKINS, J. W. the Relationship of the Nucleolus To the Synthesis of Ribosomal Rna in Hela Cells. *Proc. Natl. Acad. Sci. U. S. A.* **51**, 1197–1204 (1964).
15. Perry, R. P. the Cellular Sites of Synthesis of Ribosomal and 4S Rna. *Proc. Natl. Acad. Sci.* **48**, 2179–2186 (1962).
16. Reeder, R. H. & Roeder, R. G. Ribosomal RNA synthesis in isolated nuclei. *J. Mol. Biol.* **67**, 433–441 (1972).
17. Roeder, R. G. & Rutter, W. J. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**, 234–237 (1969).
18. Miller, O. L. & Beatty, B. R. Visualization of nucleolar genes. *Science (80-. )*. **164**, 955–957 (1969).
19. Bernhard, W., Haguenu, F., and Oberling, C. L'ultra-structure du nucléole de quelques cellules animales, r'év'el'ee par le microscope 'electronique. *Exp. 1.1.2* (1952) doi:10.1017/cbo9781139629485.007.
20. Sirri, V., Urcuqui-Inchima, S., Roussel, P. & Hernandez-Verdun, D. Nucleolus: The fascinating nuclear body. *Histochem. Cell Biol.* **129**, 13–31 (2008).
21. Prieto, J. L. & McStay, B. Pseudo-NORs: A novel model for studying nucleoli. *Biochim. Biophys. Acta - Mol. Cell Res.* **1783**, 2116–2123 (2008).
22. Raška, I., Shaw, P. J. & Cmarko, D. Structure and function of the nucleolus in the spotlight. *Curr. Opin. Cell Biol.* **18**, 325–334 (2006).
23. Raska, I. Does the synthesis of ribosomal RNA take place within nucleolar fibrillar centers or dense fibrillar components? (1995).
24. Sirri, V., Hernandez-Verdun, D. & Roussel, P. Cyclin-dependent kinases govern formation and maintenance of the nucleolus. *J. Cell Biol.* **156**, 969–981

- (2002).
25. Jantzen, H. M., Admon, A., Bell, S. P. & Tjian, R. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* **344**, 830–836 (1990).
  26. Learned, R. M., Cordes, S. & Tjian, R. Purification and characterization of a transcription factor that confers promoter specificity to human RNA polymerase I. *Mol. Cell. Biol.* **5**, 1358–1369 (1985).
  27. J., B. & Dobрева G. TIF-IA , the factor mediating growth-dependent control of ribosomal ... *EMBO Rep.* **Bd 1**, 171–5 (2000).
  28. Moorefield, B., Greene, E. A. & Reeder, R. H. RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4724–4729 (2000).
  29. Andersen, J. S. *et al.* Nucleolar proteome dynamics. *Nature* **433**, 77–83 (2005).
  30. King, T. C., Sirdeskmukh, R. & Schlessinger, D. Nucleolytic processing of ribonucleic acid transcripts in procaryotes. *Microbiol. Rev.* **50**, 428–451 (1986).
  31. Gaal, T. *et al.* Colocalization of distant chromosomal loci in space in e. Coli: A bacterial nucleolus. *Genes Dev.* **30**, 2272–2285 (2016).
  32. de Boer, H. A., Gilbert, S. F. & Nomura, M. DNA sequences of promoter regions for rRNA operons rrnE and rrnA in E. coli. *Cell* **17**, 201–209 (1979).
  33. Chiaruttini, C. & Milet, M. Gene organization, primary structure and rna processing analysis of a ribosomal rna operon in lactococcus lactis. *J. Mol. Biol.* **230**, 57–76 (1993).
  34. LaFauci, G., Widom, R. L., Eisner, R. L., Jarvis, E. D. & Rudner, R. Mapping of rRNA genes with integrable plasmids in Bacillus subtilis. *J. Bacteriol.* **165**, 204–214 (1986).
  35. Petes T.D., B. D. Simple Mendelian Inheritance of the Reiterated Ribosomal DNA of Yeast Author ( s ): Thomas D . Petes and David Source : Proceedings of the National Academy of Sciences of the United States of America , Published by : National Academy of Sciences. **74**, 5091–5095 (2020).
  36. Hamperl, S. *et al.* Chromatin states at ribosomal DNA loci. *Biochim. Biophys.*

- Acta - Gene Regul. Mech.* **1829**, 405–417 (2013).
37. Elion, E. A. & Warner, J. R. An RNA polymerase I enhancer in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **6**, 2089–2097 (1986).
  38. Warner, J. R. Synthesis of ribosomes in *Saccharomyces* These include :  
Synthesi ' s Ri ' bosomes i ' n Sacc-harornyces. **53**, 256–271 (1989).
  39. French, S. L., Osheim, Y. N., Cioci, F., Nomura, M. & Beyer, A. L. Structural and vibrational properties of solid Mg(OH)<sub>2</sub> and Ca(OH)<sub>2</sub>.pdf. *Mol. Cell. Biol.* **23**, 1558–1568 (2003).
  40. Kobayashi, T., Hidaka, M., Nishizawa, M. & Horiuchi, T. Identification of a site required for DNA replication fork blocking activity in the rRNA gene cluster in *Saccharomyces cerevisiae*. *MGG Mol. Gen. Genet.* **233**, 355–362 (1992).
  41. Kobayashi, T. Strategies to maintain the stability of the ribosomal RNA gene repeats: Collaboration of recombination, cohesion, and condensation. *Genes Genet. Syst.* **81**, 155–161 (2006).
  42. Hart, R. P. & Folk, W. R. Structure and organization of a mammalian 5 S gene cluster. *J. Biol. Chem.* **257**, 11706–11711 (1982).
  43. Little, R. D. & Braaten, D. C. Genomic organization of human 5 S rDNA and sequence of one tandem repeat. *Genomics* **4**, 376–383 (1989).
  44. Long, E. O. & Dawid, I. B. Repeated genes in eukaryotes. *Annu. Rev. Biochem.* **49**, 727–764 (1980).
  45. Copenhaver, G. P. & Pikaard, C. S. RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of *Arabidopsis thaliana* adjoin the telomeres on chromosomes 2 and 4. *Plant J.* **9**, 259–272 (1996).
  46. Worton, R. G. *et al.* Human ribosomal RNA genes: Orientation of the tandem array and conservation of the 5' end. *Science (80-. ).* **239**, 64–68 (1988).
  47. Gonzalez, I. L. & Sylvester, J. E. Complete sequence of the 43-kb human ribosomal DNA repeat: Analysis of the intergenic spacer. *Genomics* vol. 27 320–328 (1995).
  48. Caburet, S. *et al.* Human ribosomal RNA gene arrays display a broad range of

- palindromic structures Human ribosomal RNA gene arrays display a broad range of palindromic structures. *Genome Res.* 1079–1085 (2005)  
doi:10.1101/gr.3970105.
49. Hori, Y., Shimamoto, A. & Kobayashi, T. The human ribosomal DNA array is composed of highly homogenized tandem clusters. *Genome Res.* **31**, 1971–1982 (2021).
  50. Stults, D. M., Killen, M. W., Pierce, H. H. & Pierce, A. J. Genomic architecture and inheritance of human ribosomal RNA gene clusters. *Genome Res.* **18**, 13–18 (2008).
  51. van Sluis, M., van Vuuren, C., Mangan, H. & McStay, B. NORs on human acrocentric chromosome p-arms are active by default and can associate with nucleoli independently of rDNA. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 10368–10377 (2020).
  52. Sakai, K. *et al.* Human ribosomal RNA gene cluster: identification of the proximal end containing a novel tandem repeat sequence. *Genomics* **26**, 521–526 (1995).
  53. McStay, B. & Grummt, I. The Epigenetics of rRNA Genes: From Molecular to Chromosome Biology. *Annu. Rev. Cell Dev. Biol.* **24**, 131–157 (2008).
  54. Floutsakou, I. *et al.* The shared genomic architecture of human nucleolar organizer regions. *Genome Res.* **23**, 2003–2012 (2013).
  55. Hernández-Hernández, A. *et al.* Changes of the nucleolus architecture in absence of the nuclear factor CTCF. *Cytogenet. Genome Res.* **136**, 89–96 (2012).
  56. St.Laurent, G., Wahlestedt, C. & Kapranov, P. The Landscape of long noncoding RNA classification. *Trends Genet.* **31**, 239–251 (2015).
  57. Haltiner, M. M., Smale, S. T. & Tjian, R. Two distinct promoter elements in the human rRNA gene identified by linker scanning mutagenesis. *Mol. Cell. Biol.* **6**, 227–235 (1986).
  58. Grummt, R. E. and I. Molecular Coevolution of Mammalian Ribosomal Gene Terminator Sequences and the Transcription Termination Factor TTF-I Author (

- s ): Raymond Evers and Ingrid Grummt. *Proc. Natl. Acad. Sci. United States Am.* , **92**, 5827–5831 (1995).
59. Akamatsu, Y. & Kobayashi, T. The Human RNA Polymerase I Transcription Terminator Complex Acts as a Replication Fork Barrier That Coordinates the Progress of Replication with rRNA Transcription Activity. *Mol. Cell. Biol.* **35**, 1871–1881 (2015).
  60. Dragon, F. *et al.* A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**, 967–970 (2002).
  61. Schneider, D. A. *et al.* Transcription Elongation by RNA Polymerase I Is Linked to Efficient rRNA Processing and Ribosome Assembly. *Mol. Cell* **26**, 217–229 (2007).
  62. Young, R. a. Rna polymerase II. *Young* 689–715 (1991).
  63. Brickey, W. J. & Greenleaf, A. L. Functional studies of the carboxy-terminal repeat domain of *Drosophila* RNA polymerase II in vivo. *Genetics* **140**, 599–613 (1995).
  64. Kuhn, C. D. *et al.* Functional Architecture of RNA Polymerase I. *Cell* **131**, 1260–1272 (2007).
  65. Werner, M., Thuriaux, P. & Soutourina, J. Structure-function analysis of RNA polymerases I and III. *Curr. Opin. Struct. Biol.* **19**, 740–745 (2009).
  66. Miller, G. *et al.* hRRN3 is essential in the SL1-mediated recruitment of RNA polymerase I to rRNA gene promoters. *EMBO J.* **20**, 1373–1382 (2001).
  67. Russell, J. & Zomerdijk, J. C. B. M. RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem. Sci.* **30**, 87–96 (2005).
  68. Ganley, A. R. D. & Kobayashi, T. Highly efficient concerted evolution in the ribosomal DNA repeats: Total rDNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Res.* **17**, 184–191 (2007).
  69. Goodfellow, S. J. & Zomerdijk, J. C. B. M. Basic Mechanisms in RNA Polymerase I Transcription of the Ribosomal RNA Genes. **61**, (2013).
  70. McStay, B. Nucleolar dominance: A model for rRNA gene silencing. *Genes Dev.* **20**, 1207–1214 (2006).

71. Dover, G. A. & Flavell, R. B. Molecular coevolution: DNA divergence and the maintenance of function. *Cell* **38**, 622–623 (1984).
72. Bell, S. P., Learned, R. M., Jantzen, H. M. & Tjian, R. Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. *Science* (80-. ). **241**, 1192–1198 (1990).
73. Gorski, J. J. *et al.* A novel TBP-associated factor of SL1 functions in RNA polymerase I transcription. *EMBO J.* **26**, 1560–1568 (2007).
74. Beckmann, H. Coactivator and Promoter-Selective Properties of RNA Polymerase I TAFs. *J. Clin. Endocrinol. Metab.* **10**, 1361–1362 (1995).
75. Sullivan, G. J. *et al.* Human acrocentric chromosomes with transcriptionally silent nucleolar organizer regions associate with nucleoli. *EMBO J.* **20**, 2867–2877 (2001).
76. Friedrich, J. K., Panov, K. I., Cabart, P., Russell, J. & Zomerdijk, J. C. B. M. TBP-TAF complex SL1 directs RNA polymerase I pre-initiation complex formation and stabilizes upstream binding factor at the rDNA promoter. *J. Biol. Chem.* **280**, 29551–29558 (2005).
77. Hempel, W. M., Cavanaugh, A. H., Hannan, R. D., Taylor, L. & Rothblum, L. I. The species-specific RNA polymerase I transcription factor SL-1 binds to upstream binding factor. *Mol. Cell. Biol.* **16**, 557–563 (1996).
78. Panova, T. B., Panov, K. I., Russell, J. & Zomerdijk, J. C. B. M. Casein Kinase 2 Associates with Initiation-Competent RNA Polymerase I and Has Multiple Roles in Ribosomal DNA Transcription. *Mol. Cell. Biol.* **26**, 5957–5968 (2006).
79. Yamamoto, R. T., Nogi, Y., Dodd, J. A. & Nomura, M. RRN3 gene of *Saccharomyces cerevisiae* encodes an essential RNA polymerase I transcription factor which interacts with the polymerase independently of DNA template. *EMBO J.* **15**, 3964–3973 (1996).
80. Panov, K. I. *et al.* RNA Polymerase I-Specific Subunit CAST/hPAF49 Has a Role in the Activation of Transcription by Upstream Binding Factor. *Mol. Cell. Biol.* **26**, 5436–5448 (2006).
81. Panov, K. I., Friedrich, J. K. & Zomerdijk, J. C. B. M. A Step Subsequent to

- Preinitiation Complex Assembly at the Ribosomal RNA Gene Promoter Is Rate Limiting for Human RNA Polymerase I-Dependent Transcription. *Mol. Cell. Biol.* **21**, 2641–2649 (2001).
82. Bierhoff, H., Dundr, M., Michels, A. A. & Grummt, I. Phosphorylation by Casein Kinase 2 Facilitates rRNA Gene Transcription by Promoting Dissociation of TIF-IA from Elongating RNA Polymerase I. *Mol. Cell. Biol.* **28**, 4988–4998 (2008).
  83. Laferté, A. *et al.* The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes Dev.* **20**, 2030–2040 (2006).
  84. Panov, K. I., Friedrich, J. K., Russell, J. & Zomerdijk, J. C. B. M. UBF activates RNA polymerase I transcription by stimulating promoter escape. *EMBO J.* **25**, 3310–3322 (2006).
  85. Dundr, M. *et al.* A kinetic framework for a mammalian RNA polymerase in vivo. *Science (80-. )*. **298**, 1623–1626 (2002).
  86. O’Sullivan, A. C., Sullivan, G. J. & McStay, B. UBF Binding In Vivo Is Not Restricted to Regulatory Sequences within the Vertebrate Ribosomal DNA Repeat. *Mol. Cell. Biol.* **22**, 657–658 (2002).
  87. Percipalle, P. *et al.* The chromatin remodelling complex WSTF-SNF2h interacts with nuclear myosin 1 and has a role in RNA polymerase I transcription. *EMBO Rep.* **7**, 525–530 (2006).
  88. Halkidou, K., Logan, I. R., Cook, S., Neal, D. E. & Robson, C. N. Putative involvement of the histone acetyltransferase Tip60 in ribosomal gene transcription. *Nucleic Acids Res.* **32**, 1654–1665 (2004).
  89. Rickards, B., Flint, S. J., Cole, M. D. & LeRoy, G. Nucleolin Is Required for RNA Polymerase I Transcription In Vivo. *Mol. Cell. Biol.* **27**, 937–948 (2007).
  90. Murano, K., Okuwaki, M., Hisaoka, M. & Nagata, K. Transcription Regulation of the rRNA Gene by a Multifunctional Nucleolar Protein, B23/Nucleophosmin, through Its Histone Chaperone Activity. *Mol. Cell. Biol.* **28**, 3114–3126 (2008).
  91. Birch, J. *et al.* FACT facilitates chromatin transcription by RNA polymerases I

- and III. *EMBO J. Eur. Mol. Biol. Organ.* 854–865 (2009).
92. McStay, B. & Reeder, R. H. A DNA-binding protein is required for termination of transcription by RNA polymerase I in *Xenopus laevis*. *Mol. Cell. Biol.* **10**, 2793–2800 (1990).
  93. Heix, J. & Grummt, I. Species specificity of transcription by RNA polymerase I. *Curr. Opin. Genet. Dev.* **5**, 652–656 (1995).
  94. Lang, W. H. & Reeder, R. H. The REB1 site is an essential component of a terminator for RNA polymerase I in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13**, 649–658 (1993).
  95. Merkl, P. *et al.* Binding of the Termination Factor Nsi1 to Its Cognate DNA Site Is Sufficient To Terminate RNA Polymerase I Transcription In Vitro and To Induce Termination In Vivo. *Mol. Cell. Biol.* **34**, 3817–3827 (2014).
  96. Reiter, A. *et al.* The Reb1-homologue Ydr026c/Nsi1 is required for efficient RNA polymerase I termination in yeast. *EMBO J.* **31**, 3480–3493 (2012).
  97. Grummt, I., Kuhn, A., Bartsch, I. & Rosenbauer, H. A transcription terminator located upstream of the mouse rDNA initiation site affects rRNA synthesis. *Cell* **47**, 901–911 (1986).
  98. Mcstay, B. & Reeder, R. H. A termination site for *Xenopus* RNA polymerase I also acts as an element of an adjacent promoter. *Cell* **47**, 913–920 (1986).
  99. Kawauchi, J., Mischo, H., Braglia, P., Rondon, A. & Proudfoot, N. J. Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes Dev.* **22**, 1082–1092 (2008).
  100. El Hage, A., Koper, M., Kufel, J. & Tollervey, D. Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. *Genes Dev.* **22**, 1069–1081 (2008).
  101. Braglia, P., Heindl, K., Schleiffer, A., Martinez, J. & Proudfoot, N. J. Role of the RNA/DNA kinase Grc3 in transcription termination by RNA polymerase I. *EMBO Rep.* **11**, 758–764 (2010).
  102. Németh, A., Guibert, S., Tiwari, V. K., Ohlsson, R. & Längst, G. Epigenetic regulation of TTF-I-mediated promoter-terminator interactions of rRNA genes.

- EMBO J.* **27**, 1255–1265 (2008).
103. Németh, A. & Längst, G. Genome organization in and around the nucleolus. *Trends Genet.* **27**, 149–156 (2011).
  104. Shiue C-N, Berkson RG, W. A. cMyc\_induces\_changes\_in\_higher.PDF. (2009).
  105. Aprikian, P., Moorefield, B. & Reeder, R. H. New Model for the Yeast RNA Polymerase I Transcription Cycle. *Mol. Cell. Biol.* **21**, 4847–4855 (2001).
  106. Smith, S. D., O’Mahony, D. J., Kinsella, B. T. & Rothblum, L. I. Transcription from the rat 45S ribosomal DNA promoter does not require the factor UBF. *Gene Expr.* **3**, 229 (1993).
  107. Panov, K. I., Friedrich, J. K., Russell, J. & Zomerdijk, J. C. B. M. UBF activates RNA polymerase I transcription by stimulating promoter escape. *EMBO J.* **25**, 3310–3322 (2006).
  108. Moss, T. & Stefanovsky, V. Y. Promotion and regulation of ribosomal transcription in eukaryotes by RNA polymerase I. *Prog. Nucleic Acid Res. Mol. Biol.* **50**, 25–66 (1995).
  109. Stefanovsky, V. Y. DNA looping in the RNA polymerase I enhancesome is the result of non-cooperative in-phase bending by two UBF molecules. *Nucleic Acids Res.* **29**, 3241–3247 (2001).
  110. McStay, B., Frazier, M. W. & Reeder, R. H. xUBF contains a novel dimerization domain essential for RNA polymerase I transcription. *Genes Dev.* **5**, 1957–1968 (1991).
  111. Bazett-jones, D. P., Leblanc, B., Herfort, M. & Moss, T. Short-Range DNA Looping by the Xenopus HMG-Box Transcription Factor , xUBF Author ( s ): David P . Bazett-Jones , Benoît Leblanc , Manfred Herfort and Tom Moss Published by : American Association for the Advancement of Science Stable URL : <https://www.jst>. **264**, 1134–1137 (1994).
  112. Putnam, C. D., Copenhaver, G. P., Denton, M. L. & Pikaard, C. S. The RNA polymerase I transactivator upstream binding factor requires its dimerization domain and high-mobility-group (HMG) box 1 to bend, wrap, and positively

- supercoil enhancer DNA. *Mol. Cell. Biol.* **14**, 6476–6488 (1994).
113. Zentner, G. E., Saiakhova, A., Manaenkov, P., Adams, M. D. & Scacheri, P. C. Integrative genomic analysis of human ribosomal DNA. *Nucleic Acids Res.* **39**, 4949–4960 (2011).
  114. Sanij, E. *et al.* UBF levels determine the number of active ribosomal RNA genes in mammals. *J. Cell Biol.* **183**, 1259–1274 (2008).
  115. Diesch, J., Hannan, R. D. & Sanij, E. Genome wide mapping of UBF binding-sites in mouse and human cell lines. *Genomics Data* **3**, 103–105 (2015).
  116. Dhillon, A. S., Hamdane, N., Rothblum, L. I. & Richard, B. A novel role for the RNA polymerase I transcription factor UBTF in maintaining genome stability through the regulation of highly transcribed RNA Pol II genes Running. *Genome Res.* 201–212 (2014) doi:10.1101/gr.176115.114.Freely.
  117. Hamdane, N. Conditional Inactivation of Upstream Binding Factor Reveals Its Epigenetic Functions and the Existence of a Somatic Nucleolar Precursor Body. (2014).
  118. Scheer, U. & Rose, K. M. Localization of RNA polymerase I in interphase cells and mitotic chromosomes by light and electron microscopic immunocytochemistry. *Isotopenpraxis* **20**, 1431–1435 (1984).
  119. Roussel, P., André, C., Comai, L. & Hernandez-Verdun, D. The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. *J. Cell Biol.* **133**, 235–246 (1996).
  120. Jordan, P., Mannervik, M., Tora, L. & Carmo-Fonseca, M. In vivo evidence that TATA-binding protein/SL1 colocalizes with UBF and RNA polymerase I when rRNA synthesis is either active or inactive. *J. Cell Biol.* **133**, 225–234 (1996).
  121. Valdez, B. C. *et al.* The Treacher Collins Syndrome ( TCOF1 ) Gene Product Is Involved in Ribosomal DNA Gene Transcription by Interacting with Upstream Binding Factor Published by : National Academy of Sciences Linked references are available on JSTOR for this article : The Tr. (2004).
  122. Zillner, K. *et al.* Active human nucleolar organizer regions are interspersed with inactive rDNA repeats in normal and tumor cells. *Epigenomics* **7**, 363–378

- (2015).
123. Heliot, L. *et al.* Electron tomography of metaphase nucleolar organizer regions: Evidence for a twisted-loop organization. *Mol. Biol. Cell* **8**, 2199–2216 (1997).
  124. McStay, B. & Grummt, I. The epigenetics of rRNA genes: From molecular to chromosome biology. *Annu. Rev. Cell Dev. Biol.* **24**, 131–157 (2008).
  125. Li, J., Santoro, R., Koberna, K. & Grummt, I. The chromatin remodeling complex NoRC controls replication timing of rRNA genes. *EMBO J.* **24**, 120–127 (2005).
  126. Strohner, R. *et al.* NoRC - A novel member of mammalian ISWI-containing chromatin remodeling machines. *EMBO J.* **20**, 4892–4900 (2001).
  127. Cate, J. H. *et al.* X-Ray Crystal Structures of 70S Ribosome Functional Complexes Published by : American Association for the Advancement of Science Stable URL : <http://www.jstor.com/stable/2899579> Linked references are available on JSTOR for this article : You may need to I. **285**, 2095–2104 (1999).
  128. Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S. & Ban, N. Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science (80-. )*. **334**, 941–948 (2011).
  129. Henras, A. K., Plisson-Chastang, C., O'Donohue, M. F., Chakraborty, A. & Gleizes, P. E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA* **6**, 225–242 (2015).
  130. Sloan, K. E. *et al.* Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J. Cell Biol.* **200**, 577–588 (2013).
  131. Tomecki, R., Sikorski, P. J. & Zakrzewska-Placzek, M. Comparison of preribosomal RNA processing pathways in yeast, plant and human cells – focus on coordinated action of endo- and exoribonucleases. *FEBS Lett.* **591**, 1801–1850 (2017).
  132. Ciganda, M. & Williams, N. Eukaryotic 5S rRNA biogenesis. *Wiley Interdiscip. Rev. RNA* **2**, 523–533 (2011).

133. Correll, C. C., Bartek, J. & Dundr, M. The Nucleolus: A Multiphase Condensate Balancing Ribosome Synthesis and Translational Capacity in Health, Aging and Ribosomopathies. *Cells* **8**, (2019).
134. Lafontaine, D. L. J., Riback, J. A., Bascetin, R., Brangwynne, C. P. & Lafontaine, Denis L J ; Riback, Joshua A ; Bascetin, Rümeyza ; Brangwynne, C. P. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* **22**, 165–182 (2021).
135. Mangan, H. & McStay, B. Human nucleoli comprise multiple constrained territories, tethered to individual chromosomes. *Genes Dev.* **35**, 483–488 (2021).
136. van Sluis, M., van Vuuren, C. & McStay, B. The Relationship Between Human Nucleolar Organizer Regions and Nucleoli, Probed by 3D-ImmunoFISH. *Methods Mol. Biol.* **1455**, 3–14 (2016).
137. Ochs, R. L., Lischwe, M. A., Spohn, W. H. & Busch, H. Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. *Biol. Cell* **54**, 123–133 (1985).
138. Thiry, M. & Lafontaine, D. L. J. Birth of a nucleolus: The evolution of nucleolar compartments. *Trends Cell Biol.* **15**, 194–199 (2005).
139. Anastassova-Kristeva. The nuclear cycle in man. *J. Hyg. (Lond).* **47**, 182–187 (1977).
140. Clemente-Blanco, A. *et al.* Cdc14 inhibits transcription by RNA polymerase I during anaphase. *Nature* **458**, 219–222 (2009).
141. Hernandez-Verdun, D. Assembly and disassembly of the nucleolus during the cell cycle. *Nucleus* **2**, 189–194 (2011).
142. Fomproix, N., Gébrane-Younès, J. & Hernandez-Verdun, D. Effects of anti-fibrillarin antibodies on building of functional nucleoli at the end of mitosis. *J. Cell Sci.* **111**, 359–372 (1998).
143. Theos, A. C. *et al.* Functions of adaptor protein (AP)-3 and AP-1 in tyrosinase sorting from endosomes to melanosomes. *Mol. Biol. Cell* **16**, 5356–5372 (2005).
144. Dimario, P. J. Cell and molecular biology of nucleolar assembly and

- disassembly. *Int. Rev. Cytol.* **239**, 99–178 (2004).
145. Cuylen, S. *et al.* Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. *Nature* **535**, 308–312 (2016).
  146. Bruno, S. & Darzynkiewicz, Z. Cell cycle dependent expression and stability of the nuclear protein detected by Ki-67 antibody in HL-60 cells. *Cell Prolif.* **25**, 31–40 (1992).
  147. Darzynkiewicz, Z. *et al.* Initiation and termination of DNA replication during S phase in relation to cyclins D1, E and A, p21WAF1, Cdt1 and the p12 subunit of DNA polymerase  $\delta$  revealed in individual cells by cytometry. *Oncotarget* **6**, 11735–11750 (2015).
  148. Booth, D. G. *et al.* Ki-67 is a PP1-interacting protein that organises the mitotic chromosome periphery. *Elife* **3**, 1–22 (2014).
  149. Prieto, J.-L. & McStay, B. Nucleolar biogenesis: the first small steps. *Biochem. Soc. Trans.* **33**, 1441–1443 (2005).
  150. Falahati, H. & Wieschaus, E. Independent active and thermodynamic processes govern the nucleolus assembly in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1335–1340 (2017).
  151. Goodpasture, C. & Bloom, S. E. Visualization of nucleolar organizer regions in mammalian chromosomes using silver staining. *Chromosoma* **53**, 37–50 (1975).
  152. Weisenberger, D. & Scheer, U. A possible mechanism for the inhibition of ribosomal RNA gene transcription during mitosis. *J. Cell Biol.* **129**, 561–575 (1995).
  153. Roussel, P., André, C., Comai, L. & Hernandez-Verdun, D. The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. *J. Cell Biol.* **133**, 235–246 (1996).
  154. Leung, A. K. L. *et al.* Quantitative kinetic analysis of nucleolar breakdown and reassembly during mitosis in live human cells. *J. Cell Biol.* **166**, 787–800 (2004).
  155. Chen, D. *et al.* Condensed mitotic chromatin is acc transcription factors and

- chromatin. **168**, 41–54 (2020).
156. Muro, E. *et al.* The traffic of proteins between nucleolar organizer regions and prenucleolar bodies governs the assembly of the nucleolus at exit of mitosis. *Nucleus* **1**, 202–211 (2010).
  157. Verheggen, C. *et al.* The Ribosomal RNA Processing Machinery Is Recruited to the Nucleolar Domain before RNA Polymerase I during *Xenopus laevis* Development Stable URL : <https://www.jstor.org/stable/1619690>  
REFERENCES Linked references are available on JSTOR for this article : **149**, 293–305 (2000).
  158. Gonda, K. *et al.* Reversible disassembly of somatic nucleoli by the germ cell proteins FRGY2a and FRGY2b. *Nat. Cell Biol.* **5**, 205–210 (2003).
  159. Carron, C. *et al.* Post-mitotic dynamics of pre-nucleolar bodies is driven by pre-rRNA processing. *J. Cell Sci.* **125**, 4532–4542 (2012).
  160. Savino, T. M., Gébrane-Younès, J., De Mey, J., Sibarita, J. B. & Hernandez-Verdun, D. Nucleolar assembly of the rRNA processing machinery in living cells. *J. Cell Biol.* **153**, 1097–1110 (2001).
  161. Amenta, P. S. Fusion of nucleoli in cells cultured from the heart of *Triturus viridescens*. *Anat. Rec.* **139**, 155–165 (1961).
  162. Brangwynne, C. P., Mitchison, T. J. & Hyman, A. A. Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4334–4339 (2011).
  163. EHRENBERG, L. Influence of temperature on the nucleolus and its coacervate nature. *Hereditas* **32**, 407–418 (1946).
  164. Mangan, H., Gailín, M. & McStay, B. Integrating the genomic architecture of human nucleolar organizer regions with the biophysical properties of nucleoli. *FEBS J.* **284**, 3977–3985 (2017).
  165. Feric, M. *et al.* Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697 (2016).
  166. Van Sluis, M. & Mcstay, B. Supplementary Materials A localized nucleolar damage response involves recruitment of the homology- directed repair

- machinery, independent of cell cycle stage. 1151–1163 (2015)  
doi:10.1101/gad.260703.115.
167. van Sluis, M. & McStay, B. Nucleolar reorganization in response to rDNA damage. *Curr. Opin. Cell Biol.* **46**, 81–86 (2017).
  168. Mais, C., Wright, J. E., Prieto, J. L., Raggett, S. L. & McStay, B. UBF-binding site arrays form pseudo-NORs and sequester the RNA polymerase I transcription machinery. *Genes Dev.* **19**, 50–64 (2005).
  169. Mangan, H., Gailín, M. & McStay, B. Integrating the genomic architecture of human nucleolar organizer regions with the biophysical properties of nucleoli. *FEBS J.* **284**, 3977–3985 (2017).
  170. McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019).
  171. Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838 (2013).
  172. Cong, L. *et al.* Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., ... Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* (New York, N.Y.). *Science* **339**, 819–823 (2013).
  173. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).
  174. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* (80-. ). **337**, 816–821 (2012).
  175. Certo, M. T. *et al.* Tracking genome engineering outcome at individual DNA breakpoints. *Nat. Methods* **8**, 671–676 (2011).
  176. Chen, J. S. *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *bioRxiv* **550**, 407–410 (2017).
  177. Vakulskas, C. A. *et al.* *HHS Public Access*. vol. 24 (2019).

178. Jasin, M. & Haber, J. HHS Public Access Author manuscript DNA Repair (Amst). Author manuscript; available in PMC 2017 July 27. Published in final edited form as: DNA Repair (Amst). 2016 August ; 44: 6–16. doi:10.1016/j.dnarep.2016.05.001. The Democratization of Gene Editing: I. *DNA Repair (Amst)*. **176**, 139–148 (2016).
179. Bodnar, A. G. *et al.* Extension of life-span by introduction of telomerase into normal human cells. *Science (80-. )*. **279**, 349–352 (1998).
180. Jiang, X. R. *et al.* Telomerase expression in human somatic cells does not induce changes associated with a transformed phenotype. *Nat. Genet.* **21**, 111–1114 (1999).
181. Landry. The genomic and transcriptomic landscape of a hela cell line. *Cell Rep* **3**, (2015).
182. Haapaniemi, E., Botla, S., Persson, J., Schmierer, B. & Taipale, J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **24**, 927–930 (2018).
183. Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* **24**, 1216–1224 (2018).
184. Ryan, M. D., King, A. M. Q. & Thomas, G. P. Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *J. Gen. Virol.* **72**, 2727–2732 (1991).
185. Pelletier, J. & Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320–325 (1988).
186. Szymczak-Workman, A. L., Vignali, K. M. & Vignali, D. A. A. Design and construction of 2A peptide-linked multicistronic vectors. *Cold Spring Harb. Protoc.* **7**, 199–204 (2012).
187. Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, 1–8 (2011).
188. Wang, Y., Wang, F., Wang, R., Zhao, P. & Xia, Q. 2A self-cleaving peptide-

- based multi-gene expression system in the silkworm *Bombyx mori*. *Sci. Rep.* **5**, 1–10 (2015).
189. Chelly van Vuuren, P. thesis. A Refined Model of Nucleolar Formation Derived from the Analysis of UBF Variants and the Hierarchical Silencing of Human Nucleolar Organizer Regions Chelly van Vuuren Centre for Chromosome Biology , School of Natural Sciences , National University of Irel. (2017).
  190. Bachellerie, J. P., Cavallé, J. & Hüttenhofer, A. The expanding snoRNA world. *Biochimie* **84**, 775–790 (2002).
  191. Shubina, M. Y., Musinova, Y. R. & Sheval, E. V. Nucleolar Methyltransferase Fibrillarin: Evolution of Structure and Functions. *Biochemistry. (Mosc)*. **81**, 941–950 (2016).
  192. Roberts, B. *et al.* Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization. *Mol. Biol. Cell* **28**, 2854–2874 (2017).
  193. Hernandez-Verdun, D. The nucleolus today. *J. Cell Sci.* **99**, 465–471 (1991).
  194. Giepmans, B. N. G., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. The fluorescent toolbox for assessing protein location and function. *Science (80- )*. **312**, 217–224 (2006).
  195. Dundr, M. *et al.* A kinetic framework for a mammalian RNA polymerase in vivo. *Science (80- )*. **298**, 1623–1626 (2002).
  196. Savino, T. M., Gebrane-Younes, J., De Mey, J., Sivarita, J.-B. & Hernandez-Verdun, D. Nucleolar Assembly of the rRNA Processing Machinery in Living Cells. **153**, 1097–1110 (2001).
  197. Booth, D. G. *et al.* 3D-CLEM Reveals that a Major Portion of Mitotic Chromosomes Is Not Chromatin. *Mol. Cell* **64**, 790–802 (2016).
  198. Delepierre, M. *et al.* Reassessment of structural characteristics of the d(CGCG)<sub>2</sub>:actinomycin D complex from complete <sup>1</sup>H and <sup>31</sup>P NMR. *J. Biomol. Struct. Dyn.* **7**, 557–589 (1989).
  199. Wu, M. H. & Yung, B. Y. Cell cycle phase-dependent cytotoxicity of actinomycin D in HeLa cells. *Eur. J. Pharmacol.* **270**, 203–212 (1994).
  200. Berry, J., Brangwynne, C. P. & Haataja, M. Physical principles of intracellular

- organization via active and passive phase transitions. *Reports Prog. Phys.* **81**, (2018).
201. Feric, M. *et al.* Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697 (2016).
  202. Weber, S. C. & Brangwynne, C. P. Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr. Biol.* **25**, 641–646 (2015).
  203. Caragine, C. M., Haley, S. C. & Zidovska, A. Nucleolar dynamics and interactions with nucleoplasm in living cells. *Elife* **8**, 1–21 (2019).
  204. Lindström, M. S. NPM1/B23: A multifunctional chaperone in ribosome biogenesis and chromatin remodeling. *Biochem. Res. Int.* **2011**, (2011).
  205. Sen, O., Saurin, A. T. & Higgins, J. M. G. The live cell DNA stain SiR-Hoechst induces DNA damage responses and impairs cell cycle progression. *Sci. Rep.* **8**, 1–8 (2018).
  206. Caragine, C. M., Haley, S. C. & Zidovska, A. Surface Fluctuations and Coalescence of Nucleolar Droplets in the Human Cell Nucleus. *Phys. Rev. Lett.* **121**, 148101 (2018).
  207. Singh, S., Broeck, A. Vanden, Miller, L., Chaker-Margot, M. & Klinge, S. Nucleolar maturation of the human small subunit processome. *Science (80-. )*. **373**, (2021).
  208. Chambers, J. E. *et al.* An Optical Technique for Mapping Microviscosity Dynamics in Cellular Organelles. *ACS Nano* **12**, 4398–4407 (2018).
  209. Boersma, A. J., Zuhorn, I. S. & Poolman, B. A sensor for quantification of macromolecular crowding in living cells. *Nat. Methods* **12**, 227–229 (2015).