

Uniwersytet im. Adama Mickiewicza w Poznaniu
Wydział Biologii
Instytut Biologii Molekularnej i Biotechnologii
Zakład Ekspresji Genów



Rozprawa doktorska

**Zależność biogenezy cząsteczki miR-21 od dojrzewania
jej prekursorów na poziomie alternatywnego splicingu
oraz poliadenylacji**

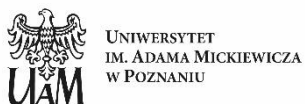
mgr Michał Sekrecki

Promotor: prof. dr hab. Krzysztof Sobczak

Poznań 2021

Finansowanie

Niniejsza rozprawa doktorska zrealizowana została ze wsparciem finansowym:



UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Pasport do przyszłości -Interdyscyplinarne studia doktoranckie na Wydziale Biologii UAM

POWR.03.02.00-00-I006/17



Specjalne stypendium naukowe z dotacji podmiotowej w ramach Krajowego Naukowego Ośrodka Wiodącego „Poznańskie Konsorcjum RNA”

Narodowego Centrum Nauki (NCN), w ramach projektu pt. **Mechanizmy prowadzące do zmian poziomu ekspresji i aktywności mikroRNA w dystrofii miotonicznej** na podstawie decyzji nr 2014/15/B/NZ2/02453



NARODOWE CENTRUM NAUKI

Podziękowania

Serdeczne podziękowania dla **prof. dr hab. Krzysztofa Sobczaka** ze te wszystkie lata współpracy, dzielenia się swoją wiedzą, uwagami oraz merytorycznym wsparciem podczas trwania całego doktoratu.

Narzeczonej Agacie za silną wiarę w moje dążenie do celu oraz nieocenione wsparcie w trudnych chwilach. Bez **Ciebie** na pewno ta praca by nie powstała.

Mojej Rodzinie, a zwłaszcza **Rodzicom**, którzy swoją ciężką pracą i dużym zaangażowaniem odpowiednio przygotowali mnie do dorosłego życia.

Koleżankom oraz Kolegom, tym obecnym oraz byłym, z Laboratorium prof. Sobczaka za przyjazną oraz twórczą atmosferę w laboratorium oraz poza nim.

W szczególny sposób chciałbym podziękować **dr Agnieszce Piaseckiej** za wspólną walkę na tym samym miRNA-froncie.

Wszystkim Koleżankom i Kolegom z Instytutu Biologii Molekularnej i Biotechnologii za tworzenie tak ciekawego oraz inspirującego miejsca pracy naukowej.

Pracownikom Instytutu Biologii Molekularnej i Biotechnologii za ich merytoryczną, techniczną oraz administracyjną pracę na rzecz całego Instytutu.

Pani **dr Karolinie Cerbin** za jej nieocenioną pracę oraz pomoc w załatwianiu wszystkich formalności.

Dedykacja

Moim Bliskim

Wykaz skrótów

2'-OMe-RNA-PS	analog RNA ze szkieletem tiofosforanowym (PS, ang. <i>phosphorothioate</i>) oraz podstawioną grupą metylowa przy atomie węgla 2' rybozy
3'ss	miejsce splicingu intronu po stronie 3' (ang. <i>3' splice site</i>)
5'ss	miejsce splicingu intronu po stronie 5' (ang. <i>5' splice site</i>)
AA	aminokwas (ang. <i>amino acid</i>)
AON	antysensowny oligonukleotyd (ang. <i>antisense oligonucleotide</i>)
APAS	alternatywne miejsce sygnału poliadenylacji (ang. <i>alternative polyadenylation site</i>)
CAGE	analiza ekspresji genów opierająca się na strukturze cap (ang. <i>Cap Analysis of Gene Expression</i>)
cDNA	komplementarny DNA (ang. <i>complementary DNA</i>)
CDS	sekwencja kodująca białko (ang. <i>coding sequence</i>)
CR-APA	alternatywna poliadenylacja wewnątrz sekwencji kodującej białko (ang. <i>coding region-APA</i>)
ENST	numer transkryptu wg. bazy Ensembl (ang. <i>Ensembl transcript</i>)
ER	retikulum endoplazmatyczne (ang. <i>endoplasmic reticulum</i>)
GFP	białko zielonej fluorescencji (ang. <i>green fluorescent protein</i>)
HeLa	ludzka, nabłonkowa linia komórkowa, wyizolowana z szyjki macicy
HSkM	ludzka mioblastyczna linia komórkowa (ang. <i>human skeletal myoblast</i>)
kpz	tysięcy par zasad
lncRNA	długie, niekodujące RNA (ang. <i>long noncoding RNA</i>)
mRNA	RNA informacyjne (ang. <i>messenger RNA</i>).
ncRNA	niekodujące RNA (ang. <i>noncoding RNA</i>)
nt	nukleotyd (ang. <i>nucleotide</i>)
pA	poliadenylowany (ang. <i>polyadenylated</i>)
PAS	miejsce sygnału poliadenylacji (ang. <i>polyadenylation signal</i>)
PAS dyst.	dystalne miejsce sygnału poliadenylacji (ang. <i>distal polyadenylation signal</i>)

PAS proks.	proksymalne miejsce sygnału poliadenylacji (ang. <i>proximal polyadenylation signal</i>)
piRNA	cząsteczki RNA związane z białkiem Piwi (ang. <i>Piwi interacting RNA</i>)
PR	punkt rozgałęzienia (ang. <i>branch point</i>)
PTBP	białko wiążące się do traktu polipirymidynowego (ang. <i>polypyrimidine tract binding protein</i>)
RACE	technika mapowania końców transkryptów cDNA (ang. <i>Rapid amplification of cDNA ends</i>)
RBP	białko wiążące się do RNA (ang. <i>RNA binding protein</i>)
RNAi	interferencja RNA (ang. <i>RNA interference</i>)
RNA-seq	sekwencjonowanie RNA (ang. <i>RNA-seq</i>)
siRNA	małe interferujące RNA (ang. <i>small interfering RNA</i>)
sncRNA	krótkie, niekodujące RNA (ang. <i>short noncoding RNA</i>)
snRNP	małe jądrowe nukleoproteiny (ang. <i>small nuclear ribonucleoproteins</i>)
TPM	normalizacja - transkrypty na million (ang. <i>Transcripts Per Million</i>)
TSS	miejsce startu transkrypcji (ang. <i>transcription start site</i>)
U2AF	czynnik pomocniczy U2 (ang. <i>U2 auxiliary factor</i>)
UCSC Genome Browser	internetowa przeglądarka genomowa (ang. <i>University of California Santa Cruz</i>)
UTR	region nie ulegający translacji (ang. <i>untranslated region</i>)
	3'UTR region nie ulegający translacji poniżej ORF
	5'UTR region nie ulegający translacji powyżej ORF
UTR-APA	alternatywna poliadenylacja wewnątrz sekwencji niekodującej białko (ang. <i>untranslated region-APA</i>)
VMP1	wakuolarne białko membranowe 1 (ang. <i>Vacuole membrane protein 1</i>)
WES	sekwencjonowanie całych eksonów (ang. <i>whole-exome sequencing</i>)
WGS	sekwencjonowanie całego genomu (ang. <i>whole-genome sequencing</i>)

Spis treści

Finansowanie	2
Podziękowania	3
Dedykacja	4
Wykaz skrótów	5
Spis treści.....	7
Streszczenie.....	10
Abstract	12
1. Wstęp	14
1.1. Świat RNA.....	14
1.1.1. Dojrzewanie cząsteczek prekursorowych RNA	16
1.1.1.1. Splicing RNA.....	16
Rola rybonukleoprotein (snRNP) w procesie splicingu pre-mRNA	18
1.1.1.2. Alternatywne splicing RNA (AS)	20
1.1.1.3. Poliadenylacja na końcu 3' RNA – dobudowanie ogona poli(A)	23
1.1.1.4. Alternatywna poliadenylacja (APA).....	24
1.1.1.5. Zależność pomiędzy splicingiem, a poliadenylacją	26
1.2. Długie niekodujące RNA (lncRNA).....	27
1.3. Krótkie niekodujące RNA (sncRNA).....	28
1.3.1. mikroRNA	29
1.3.1.1. Biogeneza miRNA - klasyczna droga powstawania miRNA	30
1.3.1.2. Mirtrony – alternatywna ścieżka biogenezy miRNA	31
1.3.1.3. Badania genomowe oraz identyfikacja kanonicznych cząsteczek miRNA.....	32
1.3.1.4. Genomowa lokalizacja miRNA.....	33
1.3.1.5. Mechanizm działania miRNA.....	35
1.3.1.6. Znaczenie cząsteczek miRNA w rozwoju oraz w chorobie	36
1.3.2. mikroRNA-21	37
1.3.2.1. Genomowa lokalizacja oraz regulacja biogenezy miR-21	37
1.3.2.2. Znaczenie miR-21 podczas rozwoju embrionalnego.....	38
1.3.2.3. miR-21 w kancerogenezie	39
2. Cel pracy.....	42
3. Materiały i metody.....	43
3.1. Konstrukty genetyczne <i>GFP_pri-miR-21</i>	43

3.2.	Hodowle bakteryjne	46
3.3.	Eksperymenty na liniach komórkowych.....	46
3.3.1.	Spis stosowanych linii komórkowych	46
3.3.2.	Hodowla oraz różnicowanie komórek.....	47
3.3.3.	Transfekcje komórek	47
3.3.4.	Spis stosowanych oligonukleotydów	47
3.3.5.	Izolacja RNA i odwrotna transkrypcja	49
3.3.6.	Odwrotna transkrypcja dla cząsteczek miRNA.....	49
3.4.	Analiza danych z portalu GTEx	50
3.5.	Analizy RT-PCR oraz RT-qPCR	50
3.5.1.	Lista starterów.....	51
3.6.	Western blot.....	53
3.7.	Frakcjonowanie komórek.....	54
4.	Wyniki.....	55
4.1.	Transkrypty <i>pri-miR-21</i> powstają niezależnie z własnych miejsc promotorowych, a terminacja transkrypcji zachodzi w alternatywnych miejscach poliadenylacji.....	55
4.1.1.	Gen <i>VMP1</i> oraz jego izoformy splicingowe.....	55
4.1.2.	Pierwotne prekursorzy <i>pri-miR-21</i>	59
4.1.3.	Poziom ekspresji <i>VMP1</i> , <i>IncVMP1</i> oraz <i>pri-miR-21</i> jest znacząco odmienny w różnych tkankach człowieka	61
4.1.4.	Komórki z wyciszeniem genu <i>DROSHA</i> stanowią model w badaniu krótko żyjących prekursorów <i>pri-miRNA</i>	64
4.1.5.	Zjawisko alternatywnej poliadenylacji prowadzi do powstania poliadenylowanych transkryptów <i>pri-miR-21</i>	65
4.1.6.	Poziom ekspresji <i>VMP1</i> , <i>pri-miR-21</i> oraz <i>miR-21</i> ulega istotnemu podwyższeniu podczas proliferacji mioblastów	68
4.1.7.	Brak zmian poziomu mRNA <i>VMP1</i> w komórkach z niedoborem <i>DROSHA</i>	70
4.1.8.	Podwyższona aktywność transkrypcyjna prowadzi do podwyższonego poziomu <i>miR-21</i> w komórkach raka piersi	72
4.2.	Alternatywna poliadenylacja ma bezpośredni wpływ na biogenezę <i>miR-21</i>	75
4.2.1.	Komórki z wyciszeniem <i>CPSF3</i> jako potencjalny model badania wpływu alternatywnej poliadenylacji <i>VMP1/pri-miR-21</i> na biogenezę <i>miR-21</i>	75
4.2.2.	Stworzenie konstruktów genetycznych do badań wpływu dojrzewania prekursora <i>pri-miR-21</i> na biogenezę <i>miR-21</i>	77
4.2.3.	Aktywność PAS proks. w RNA <i>VMP1/pri-miR-21</i> ma bezpośredni wpływ na biogenezę <i>miR-21</i>	79
4.3.	Zależność pomiędzy aktywnością czynników splicingowych działających w 3' ss ostatniego intronu, a zajściem procesu cięcia i poliadenylacji w <i>pri-miR-21</i>	82

4.3.1.	Blokowanie PAS proks. <i>VMP1/pri-miR-21</i> oraz jego delecja wpływają na zwiększoną retencję ostatniego intronu	82
4.3.2.	Izoformy <i>pri-miR-21</i> z i bez retencji intronu są prekursorami miR-21	84
4.3.3.	Obecność miejsca splicingu 3' (3'ss) ostatniego intronu <i>pri-miR-21</i> ma wpływ na biogenezę miR-21.....	85
5.	Dyskusja.....	88
6.	Wnioski.....	100
7.	Bibliografia	101
8.	Spis rycin i tabel.....	116

Streszczenie

Cząsteczki mikroRNA (miRNA) stanowią ważny element regulatorowy utrzymujący homeostazę w organizmach żywych. Te krótkie (20-22 nukleotydów), niekodujące RNA regulują poziom ekspresji niemal wszystkich genów kodujących białka, na drodze interferencji RNA. Na dzień dzisiejszy zidentyfikowanych jest blisko 2000 ludzkich miRNA. Jednym z lepiej poznanych jest miR-21. Cząsteczka ta od lat stanowi obiekt badań pod kątem jej roli w procesie rozwoju zwierząt, jak również w patogenezie chorób takich jak nowotwory. Pomimo dość dobrze opisanych szlaków w których zaangażowany jest miR-21, nadal słabo poznane są mechanizmy kontrolujące proces jego biogenezy. Alternatywny splicing (AS) oraz alternatywna poliadenylacja (APA) pierwotnych transkryptów są niezwykle ważnymi mechanizmami w regulacji ekspresji większości genów kodujących białka. Procesy te zachodzą podczas transkrypcji i są ze sobą ściśle związane i wzajemnie na siebie wpływają. Na dzień dzisiejszy niewiele prac badawczych skupiało się na udziale tych procesów w regulacji ekspresji genów miRNA.

Głównym celem niniejszej rozprawy doktorskiej było zbadanie zaangażowania AS oraz APA w proces biogenezy miR-21. Podczas pierwszego etapu badań wykazano, że pierwotny transkrypt *pri-miR-21* powstaje niezależnie od pre-mRNA swojego genu-gospodarza *VMP1* i może podlegać charakterystycznym dla pre-mRNA *VMP1* procesom składania. W pracy, dzięki analizie danych z sekwencjonowania RNA, określono dużą zmienność w poziomie ekspresji *pri-miR-21* w różnych ludzkich tkankach. Wyniki te wskazują na obecność precyzyjnych mechanizmów, bezpośrednio regulujących transkrypcję *pri-miR-21*, ale również sugerują zachodzenie wariantywnego splicingu tego prekursorowego RNA, co przypuszczalnie przekłada się na zróżnicowany poziom ich ekspresji w ludzkich tkankach.

Wykorzystując model komórkowy ze znacznie ograniczonym tempem przetwarzania transkryptów pri-miRNA scharakteryzowano rozmaite izoformy *pri-miR-21* w tym izoformy z retencją intronu 11. Dodatkowo poprzez zastosowanie blokera oligonukleotydowego wiążącego się do określonych sekwencji RNA lub poprzez mutację tego miejsca w konstrukcie genetycznym wykazano, że aktywność alternatywnych sygnałów poliadenylacji (APAS) występujących w 3'UTR *VMP1*, ma wpływ na efektywność biogenezy miR-21. Na podstawie przeprowadzonych analiz można wnioskować, że powstawanie kilku różnych izoform *pri-miR-21* odgrywać może ważną rolę w procesie biogenezy miR-21. Dodatkowo badania

wykazały bezpośredni wpływ PAS zlokalizowanego przed sekwencją kodującą miR-21 na ilość dojrzałej cząsteczki miR-21.

Uwzględniając obecność powszechnej izoformy *pri-miR-21* z retencją intronu, w ostatnim etapie pracy postanowiono sprawdzić zależność efektywności biogenezy miR-21 od splicingu *pri-miR-21*. Poprzez zablokowanie lub mutagenizację miejsca składania RNA zlokalizowanego po stronie 3' doprowadzono do zaburzenia splicingu *pri-miR-21* oraz zaobserwowano istotny wzrost poziomu tej cząsteczki RNA, co przełożyło się na zwiększenie liczby dojrzałych cząsteczek miR-21. Uzyskane wyniki dowodzą, że obecność izoformy z retencją intronu poprzez bezpośredni wpływ na aktywność miejsca 3' splicingu w 11 intronie *VMP1*, pozytywnie wpływa na biogenezę miR-21.

Uzyskane w pracy wyniki przyczynią się do lepszego zrozumienia mechanizmów, jakie zaangażowane mogą być w regulację poziomu cząsteczki miR-21. Opisywane w pracy zależności stanowią solidną podstawę do twierdzenia, że w globalne zmiany ilości cząsteczek miRNA zaangażowane są liczne mechanizmy po- i ko-transkrypcyjne. Taka wiedza może przyczynić się do wyjaśnienia mechanizmów prawidłowej regulacji ekspresji genów miRNA, ale również patomechanizmów wielu chorób oraz w dalszej kolejności do opracowania potencjalnych strategii terapeutycznych, w tym terapii celujących w dojrzewanie prekursorowych pri-miRNA lub biogenezę miRNA.

Słowa kluczowe: miR-21, *pri-miR-21*, alternatywna poliadenylacja, alternatywny splicing, *VMP1*, biogeneza miRNA

Abstract

MicroRNAs (miRNAs) are an important regulatory element maintaining homeostasis in living organisms. These short (20-22 nucleotides), non-coding RNAs regulate the expression level of almost all protein-coding genes through RNA interference mechanism. For now, nearly 2,000 human miRNAs have been identified. One of the well-known is the miR-21. This molecule has been an object of research for many years in terms of its role in the process of animal development, as well as in the pathogenesis of diseases such as cancer. Despite the fairly well-described pathways in which miR-21 is involved, the mechanisms controlling the process of its biogenesis are still poorly understood. Alternative splicing (AS) and alternative polyadenylation (APA) of primary transcripts are extremely important mechanisms in regulating the expression of most protein-coding genes. These processes take place during transcription and are closely related and influence each other. As of today, little research has focused on the involvement of these processes in regulating miRNA biogenesis.

The main purpose of the dissertation was to investigate the involvement of AS and APA in the miR-21 biogenesis process. During the first stage of the research, it was shown that the primary transcript of *pri-miR-21* is generated independently of the pre-mRNA of its host gene *VMP1* and may undergo splicing processes characteristic of *VMP1* pre-mRNA. Due to the analysis of RNA sequencing data, the study revealed a high variability in the expression level of *pri-miR-21* in various human tissues. These results indicate the presence of precise mechanisms directly regulating the transcription of *pri-miR-21*, but also suggest the occurrence of variant splicing of precursor RNA, which may have an impact on varying levels of expression in human tissues.

Using a cell model with a significantly reduced processing rate of pri-miRNA transcripts, various isoforms of *pri-miR-21* were characterized, including isoforms with intron 11 retention. Alternative polyadenylation signals (APAS) which are present in the 3'UTR of *VMP1* have an impact on the efficiency of miR-21 biogenesis. On the basis of the conducted analyses, it can be concluded that the formation of several different isoforms of *pri-miR-21* may play an important role in the process of miR-21 biogenesis. In addition, studies have shown a direct effect of PAS located upstream of the miR-21 coding sequence on the amount of the mature miR-21 molecule.

Taking into account the presence of the common *pri-miR-21* isoform with an intron retention, in the last stage of the work, it was decided to check the dependence of the effectiveness of miR-21 biogenesis on *pri-miR-21* splicing. By blocking or mutagenesis of the RNA splice site located on the 3' side (3'ss), *pri-miR-21* splicing was disturbed and a significant increase in the level of this RNA molecule was observed, which have an impact on increase in the number of mature miR-21 molecules. The obtained results prove that the presence of the intron retention isoform positively influences the biogenesis of miR-21 by directly affecting the activity of the 3' splicing site in intron 11 of *VMP1*.

The results obtained in this study will contribute to a better understanding of the mechanisms that may be involved in the regulation of the miR-21 biogenesis. The dependencies described in the paper provide a solid basis for the statement that numerous post- and co-transcriptional mechanisms are involved in global changes in the number of miRNA molecules. Such knowledge may help to elucidate the mechanisms of the correct regulation of miRNA gene expression, but also the pathomechanisms of many diseases, and subsequently to the development of potential therapeutic strategies, including therapies targeting the maturation of pri-miRNAs or miRNA biogenesis.

Key words: miR-21, *pri-miR-21*, alternative polyadenylation, alternative splicing, *VMP1*, miRNA biogenesis

1. Wstęp

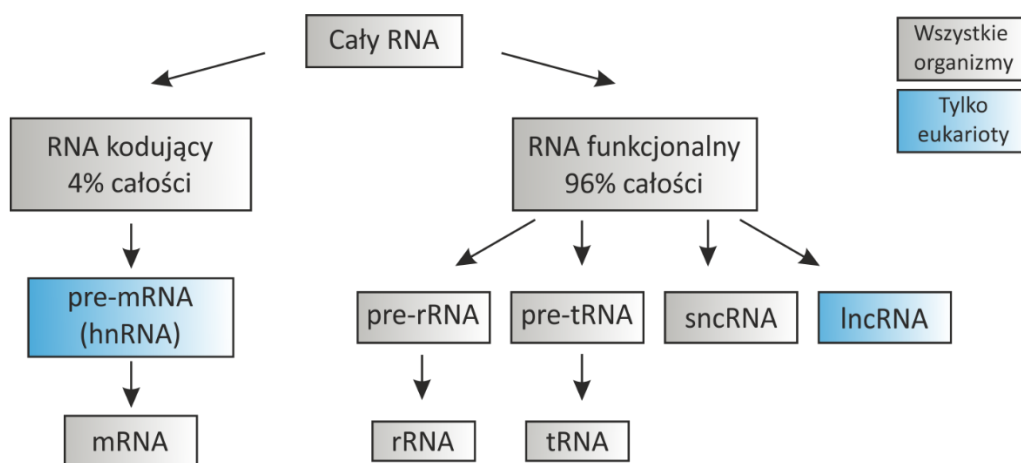
1.1. Świat RNA

Budowa i funkcjonowanie każdego żywego organizmu, a także wirusów opiera się na informacji genetycznej zapisanej w genomie. Przeważająca część genomów, w tym genomu człowieka, jest zbudowana z DNA (kwasu deoksyrybonukleinowego) a w przypadku niektórych wirusów z RNA (kwasu rybonukleinowego). Oba kwasy zbudowane są bardzo podobnie z dwoma głównymi różnicami. Pierwszą jest budowa cukru wchodzącego w skład danego kwasu nukleinowego, w przypadku RNA jest to ryboza, w przypadku DNA deoksyryboza. Drugą różnicą jest obecność w RNA uracylu, w miejscach gdzie w DNA znajdowała by się tymina. Grupę czterech nukleozydów tworzących łańcuch RNA stanowią: adenozyina (A), cytydyna (C), guanozyina (G) oraz urydyna (U). Ponadto poszczególne części nukleozydów mogą być modyfikowane, co zwiększa zmienność strukturalną i funkcjonalną cząsteczek RNA. W obu kwasach nukleinowych nukleozydy łączone są dzięki wiązaniom 3'-5' fosfodiesterowym. Jednakże stabilność cząsteczek RNA jest znacznie mniejsza, co wynika z obecności grupy hydroksylowej przy węglu 2' rybozy. RNA swoją długością rzadko przekraczają kilka, kilkanaście tysięcy nukleotydów. Pomimo tego, że w cząsteczkach RNA często występują wiązania wewnątrzcząsteczkowe, to łańcuchy RNA zazwyczaj mają formę jednoniciową. Co również charakterystyczne, cząsteczki RNA mogą przyjmować rozmaite struktury II- i III-rzędowe, w tym konformacje globularne. Za oddziaływania wewnątrzcząsteczkowe głównie odpowiedzialne są wiązania wodorowe tworzone przez komplementarne reszty nukleotydowe.

RNA powstaje na bazie matrycy kodującej ją sekwencji DNA podczas procesu zwanego transkrypcją. Przeprowadzana jest ona przez enzymy nazwane polimerazami RNA zależnymi od DNA, skrótowo nazywanymi polimerazami RNA. Transkrypcja w bakteriach oraz organizmach eukariotycznych przebiega w bardzo podobny sposób. Podstawą polimeryzacji RNA, wśród wszystkich typów organizmów, stanowi identyczna reakcja chemiczna. Każda z trzech eukariotycznych polimeraz RNA strukturalnie przypomina polimerazę bakterii *E. coli*. Oprócz podobieństw, transkrypcja u eukariontów i bakterii ma cechy specyficzne dla danej grupy organizmów. W przypadku bakterii nowopowstałe transkrypty, na bazie których w późniejszym etapie powstawać będą białka, nie ulegają żadnym procesom obróbki potranskrypcyjnej. Powstający transkrypt stanowi dojrzały mRNA, czyli RNA informacyjne

(ang. *messenger RNA*). W przypadku organizmów eukariotycznych, powstający transkrypt musi ulec szeregowi modyfikacji zanim posłuży jako matryca dla rybosomów biorących udział w syntezie białka. Wszystkie eukariotyczne mRNA posiadają strukturę tzw. czapeczki na swoim końcu 5', a w większości przypadków są także poliadenylowane (pA, ang. *polyadenylated*) na końcu 3'. Prekursorowe transkrypty (pre-mRNA) również podlegają bardzo często składaniu (ang. *splicing*), czyli wycinaniu intronów i łączeniu eksonów. Warto zauważyć, że wszystkie te procesy zachodzą jeszcze w trakcie trwania syntezy nowo powstającego RNA na matrycy DNA (1). Czapeczka na końcu 5' jest dodawana zaraz po tym, jak zainicjowana zostanie transkrypcja. Składanie i redagowanie pre-mRNA odbywa się w podczas działania polimerazy RNA. Natomiast poliadenylacja końca 3' transkryptu jest nieodłącznym elementem terminacji transkrypcji.

Ilość RNA w typowej bakterii wynosi między 0,05-0,10 pg, co odpowiada około 6% całkowitej masy. W przypadku komórek ssaków, które są znacznie większe od komórek bakteryjnych, ilość RNA jest większa i jest to 20-30 pg. Mimo, że jest to ilość większa to stanowi zaledwie 1% całkowitej masy komórki.



Rycina 1 Podział RNA występujący w komórkach. Schemat przedstawia ogólny podział RNA występujących we wszystkich organizmach z zaznaczeniem frakcji RNA, których występowanie odnotowano jedynie u organizmów eukariotycznych. Wartość 100% stanowi całą pulę RNA znajdującego się w komórce.

Podziałów RNA na typy, można przeprowadzić na wiele sposobów. Jednym z najpowszechniejszych jest podział ze względu na pełnione funkcje w komórce (**Ryc. 1**). Podstawowy podział to: RNA kodujący i RNA niekodujący. W obrębie RNA kodujących wyróżniamy jedynie mRNA. mRNA są dojrzałymi transkryptami genów, które w procesie translacji ulegają przepisaniu na białka. Zawartość mRNA w komórce wynosi ok. 4% z całości RNA znajdującego się w komórce. Cząsteczki te cechują się dosyć krótkim okresem półtrwania

i są szybko degradowane. W komórkach eukariotycznych znaczna część mRNA ulega degradacji w ciągu kilku godzin. Są jednak i takie, których okres półtrwania wynosi kilka minut. Mała stabilność komórkowa cząsteczek mRNA wskazuje na dużą dynamikę i zmienność całego transkryptomu w komórce. Drugim rodzajem opisywanych RNA są RNA niekodujące (ncRNA, ang. *noncoding RNA*). Cząsteczki te nie ulegają translacji na białka. Inną stosowaną nazwą dla tych cząsteczek jest RNA funkcjonalny, gdyż biorą aktywny udział w wielu istotnych procesach w komórce, innych niż kodowanie sekwencji białkowej. Dwie główne kategorie niekodujących RNA stanowią RNA rybosomowe (rRNA) oraz RNA transportujące (tRNA). rRNA zidentyfikować można we wszystkich organizmach i stanowią one najliczniejszą grupę RNA w komórce. rRNA są budulcem rybosomów, odpowiedzialnych za syntezę białek. tRNA, podobnie jak rRNA biorą udział w powstawaniu białek i ich rola polega na dostarczaniu aminokwasów do rybosomu w procesie syntezy białek. Poza dwoma opisanymi rodzajami niekodujących RNA, rozróżniamy wiele innych rodzajów cząsteczek RNA odgrywających znaczące funkcje w komórkach eukariotycznych. Cząsteczki te sklasyfikowane są jako krótkie niekodujące RNA (sncRNA, ang. *short noncoding RNA*), nieprzekraczające 200 nukleotydów długości, oraz długie niekodujące RNA (lncRNA, ang. *long noncoding RNA*), których długość szacowana jest na ponad 200 nukleotydów. Te klasy RNA również mogą powstawać jako cząsteczki prekursorowe i podlegać licznym procesom dojrzewania, w tym splicingowi i poliadenylacji.

1.1.1. Dojrzewanie cząsteczek prekursorowych RNA

1.1.1.1. Splicing RNA

Niemalże do końca lat 70. XX wieku naukowcy na całym świecie nie zdawali sobie sprawy z istnienia intronów. Dopiero w roku 1977, w trakcie sekwencjonowania DNA genów eukariotycznych, okazało się że geny te w swojej budowie zawierają tzw. sekwencje "wtrącone" (ang. *intervening sequences*). Sekwencje te w charakterystyczny sposób oddzielały fragmenty kodu genetycznego, które w dalszych etapach były odpowiedzialne za kodowanie sekwencji białkowej.

Składanie RNA (ang. *RNA splicing*) stanowi nieodłączny etap w dojrzewaniu prekursorowego RNA (pre-mRNA). Polega na usunięciu niekodujących informacji fragmentów, tzw. intronów. Proces splicingu bazuje na reakcjach cięcia i łączenia się eksonów, w wyniku czego powstaje dojrzała cząsteczka RNA kodująca białko - mRNA. Proces splicingu pre-mRNA stanowi podstawę różnorodności w powstawaniu białek w komórkach eukariotycznych.

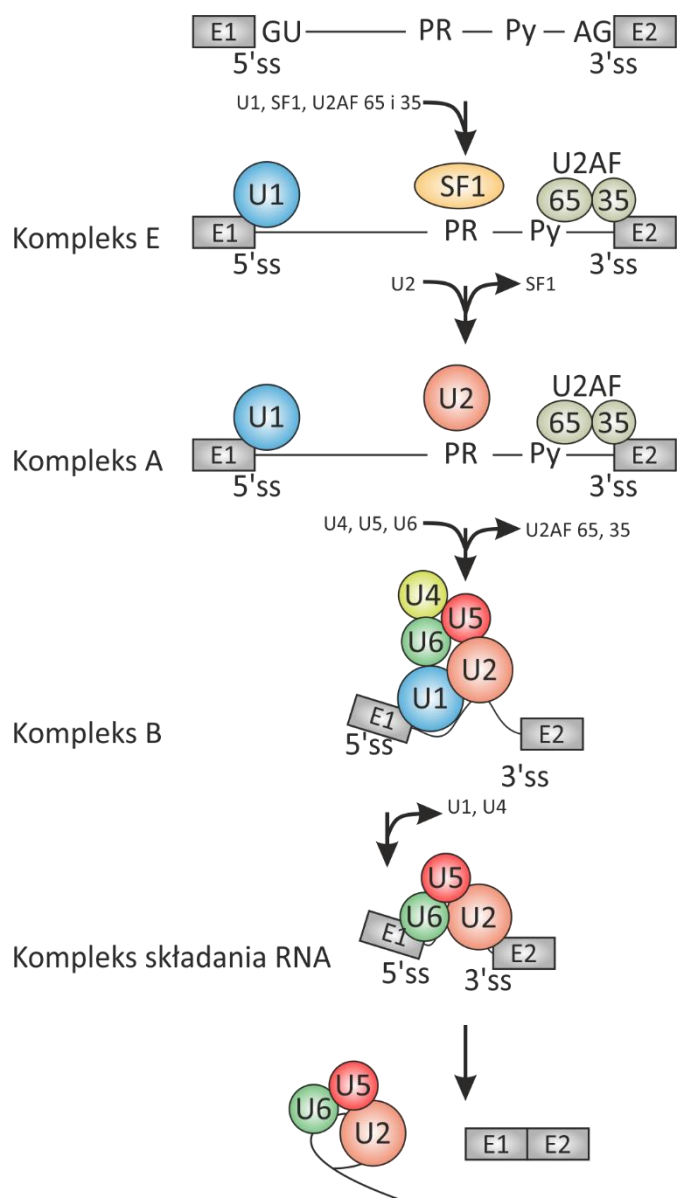
U człowieka blisko 95% genów podlega bowiem również alternatywnemu splicingowi (2). Charakterystyczną cechą pre-mRNA u eukariotów jest to, że może ono zawierać wiele, w niektórych przypadkach ponad 100 intronów, które stanowią znaczącą część całego prekursorowego transkryptu.

Zaraz po tym jak odkryto istnienie intronów, naukowcy zaobserwowali zakonserwowane motywy, które jak się okazało w istotny sposób warunkują proces ich prawidłowego wycinania intronów z pre-mRNA. Na chwilę obecną opisanych zostało siedem odrębnych rodzajów intronów wśród organizmów eukariotycznych. W przeważającej liczbie intronów, w genach kodujących białka, pierwszymi dwoma nukleotydami są 5'-GU-3'. Natomiast na końcu intronu, znajduje się sekwencja 5'-AG-3'. Introny takie określa się intronami "GU-AG" i stanowią one jeden z najpopularniejszych typów. W miarę postępów badań nad budową intronu okazało się, że powtarzające się motywy na początku oraz na końcu stanowią jedynie część zakonserwowanych sekwencji. Po stronie 3' intronu, przed sekwencją 5'-AG-3', występuje trakt polipirymidynowy, czyli rejon bogaty w pirymidyny (nukleotyd U lub C). Dodatkowo w odległości około 10-40 nt powyżej traktu polipirymidynowego znajduje się sekwencja zwana sekwencją rozgałęzienia (ang. *branch point*). Dla kręgowców kanoniczny wariant tej sekwencji to 5'-CURAY-3', gdzie "R" to puryna, a "Y" pirymidyna.

W latach 1990. badania prowadzone na całym świecie wykazały, że wycinanie intronów z pre-mRNA jest procesem wieloetapowym. Pierwszym krokiem jest cięcie w miejscu splicingu 5' intronu (5'ss, ang. *5' splice site*). Następuje to poprzez reakcję transestryfikacji - zaangażowanie grupy hydroksylowej przyłączonej od węgla 2' nukleotydu adenozyнового w punkcie rozgałęzienia (PR, ang. *branch point*), znajdującego się wewnątrz intronu. Skutkiem takiego "ataku" jest przecięcie wiązania 3'-5' fosfodiesterowego po stronie 5' intronu i powstanie nowego wiązania 5'-2' fosfodiesterowego. Reakcja ta skutkuje utworzeniem się charakterystycznej struktury, tzw. lassa (ang. *lariat*), składającego się z intronu, oraz eksonu leżącego po jego 3' stronie. Druga reakcja transestryfikacji, przeprowadzona po 3' stronie intronu (3'ss, ang. *3' splice site*) uwalnia intron w postaci lassa, który zostaje przekształcony w liniową formę i ulega degradacji. W tym samym momencie uwolniony koniec 5' eksonu, który znajdował się poniżej intronu zostaje połączony z 3' końcem eksonu leżącego przed intronem.

Rola rybonukleoprotein (snRNP) w procesie splicingu pre-mRNA

Z punktu widzenia biochemicznego, dwie reakcje transestryfikacji jakie mają miejsce podczas wycinania intronu i składania eksonów, nie stanowią dla komórki dużego wyzwania. Jednakże, do przeprowadzenia tego na pozór prostego procesu, zaangażowana musi być liczna grupa białek, które ściśle nadzorują ten proces. Nierzadko zdarza się, że dwa eksony oddalone są od siebie o nawet kilkadziesiąt tysięcy nukleotydów. Taka sytuacja wymaga obecności całej maszynery białek, która będzie w stanie w odpowiedni sposób zbliżyć do siebie miejsca istotne w procesie splicingu. Pamiętać należy, że miejsca składania eksonów w genie są bardzo do siebie podobne. Gdyby nie zaangażowanie odpowiednich białek, dochodziłoby do połączenia się nieodpowiednich miejsc i tym samym do pominięcia niektórych eksonów. Innym znanym zaburzeniem podczas składania eksonów jest wybór miejsc krytycznych, czyli miejsc znajdujących się w intronie bądź eksonie, które swoją sekwencją przypominają właściwe miejsca składania pre-mRNA. Zadaniem białek kontrolujących proces splicingu jest rozpoznawanie takich miejsc i skuteczne ich pomijanie.



Rycina 2 Schemat splicingu RNA z zaznaczeniem zaangażowanych snRNP przeprowadzających cały proces. Zaznaczone na rycinie zostały najważniejsze białka bezpośrednio zaangażowane w proces splicingu. Kluczowym momentem jest przyłączenie się kompleksów U1 oraz U2, dzięki którym dochodzi do zbliżenia się fragmentów eksonowych. Końcowymi produktami składania są związane ze sobą eksony oraz wypięty intron, który przyjmuje formę lassa. W procesie splicingu uczestniczy znacznie więcej białek, które nie zostały przedstawione na schemacie.

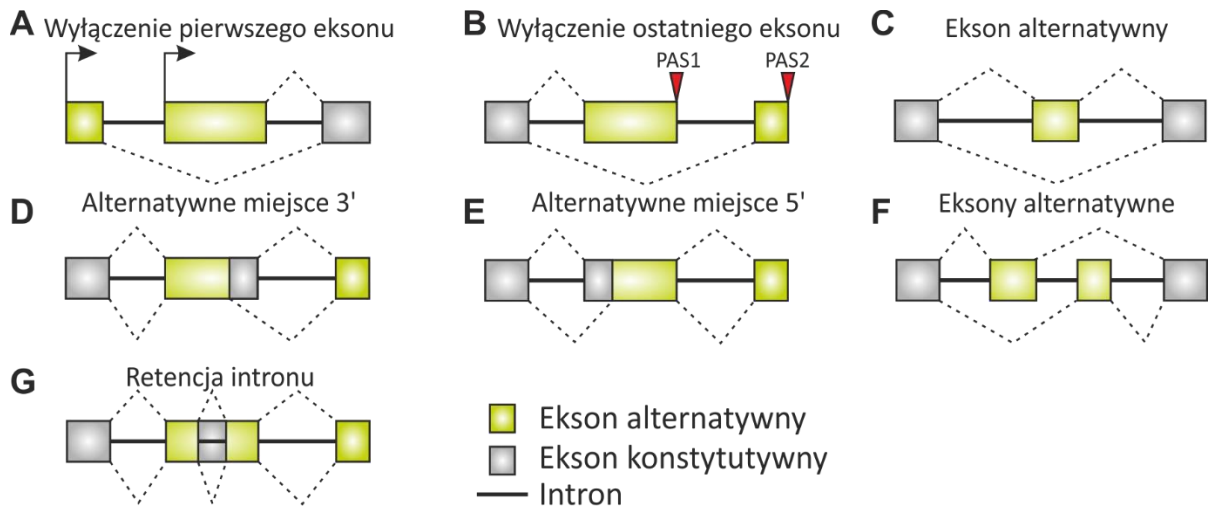
Reakcja składania RNA przeprowadzana jest przez kompleksy rybonukleoprotein (snRNP), którym towarzyszy ponad 100 innych białek związanych z procesem splicingu (3) (**Ryc. 2**). Kompleks ten nazywany jest *spliceosomem*. W jego podstawowym składzie zawartych jest 5 kompleksów U1, U2, U4, U5 oraz U6 snRNP. Głównym zadaniem wymienionych rybonukleoprotein jest rozpoznawanie charakterystycznych miejsc na końcach 5' oraz 3' intronów, traktów polipirymidynowych oraz miejsc rozgałęzienia w intronie. Aby doszło do utworzenia się spliceosomu, U1 snRNP przyłącza się do miejsca splicingu w części 5' intronu,

a w części 3' intronu wiąże się czynnik pomocniczy U2 (U2AF, ang. *U2 auxiliary factor*). U2AF jest heterodimerem składającym się z podjednostek U2AF65 oraz U2AF35. Obie rozpoznają miejsce składania na 3' końcu intronu wraz z traktem polipirymidynowym. W międzyczasie czynnik białkowy SF1 przyłącza się do miejsca rozgałęzienia wewnątrz intronu aby wspólnie z U1 oraz U2AF65 i U2AF35 utworzyć kompleks E. Kolejnym etapem jest przekształcenie kompleksu E w kompleks A, poprzez zamianę SF1 na U2 snRNP w miejscu rozgałęzienia. Duże powinowactwo U1 oraz U2 snRNP wymusza przyciągnięcie miejsca donorowego U1 w kierunku punktu rozgałęzienia. Po przyłączeniu się U4, U5 oraz U6 snRNP do intronu powstaje kompleks B. W tym momencie dochodzi do zbliżenia miejsca akceptorowego splicingu do miejsca donorowego i punktu rozgałęzień. Po uwolnieniu U1 i U4 snRNP tworzy się kompleks składania, a miejsca 5' oraz 3' wraz z punktem rozgałęzień znajdują się na tyle blisko siebie, że dochodzi do reakcji cięcia i łączenia RNA katalizowanych przez U2 i U6 snRNP. Po wycięciu intronu, produktem jest transkrypt zawierający połączone eksony.

1.1.1.2. Alternatywne splicing RNA (AS)

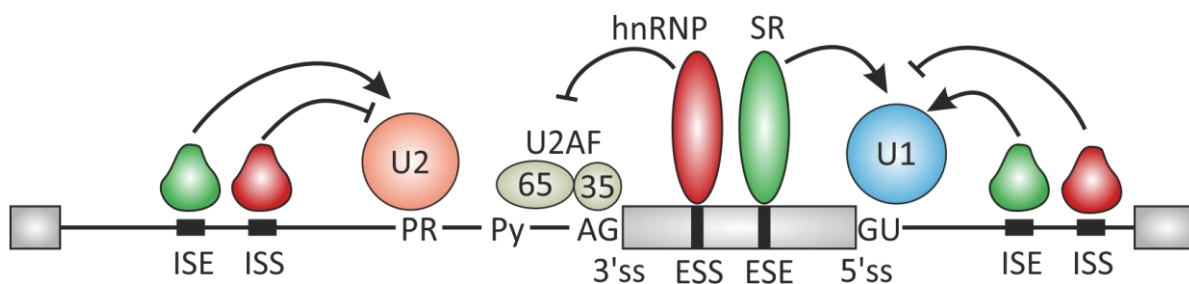
Na podstawie analiz masowego sekwencjonowania RNA (*RNA-seq*) zakłada się, że zjawisko alternatywnego składania RNA dotyczyć może blisko 95% ludzkich, wieloeksonowych genów (2)(37). Pomimo wspomnianego wcześniej faktu, że proces składania RNA z punktu widzenia biochemicznego jest prostą reakcją, to precyzja jaką wykazać musi się cała maszyna splicingowa jest zdumiewająca. Kluczowym momentem jest rozpoznanie właściwych miejsc splicingu po 5' oraz 3' stronie intronu, w szczególności w przypadku genów kodujących białko. Wystarczy przesunięcie miejsca złączenia się eksonów o kilka nukleotydów, a doprowadzić to może do zmiany całej ramki odczytu (ang. *frameshift*) i w konsekwencji najczęściej do szybkiej degradacji całego nowopowstałego transkryptu na drodze rozpadu NMD (ang. *nonsense-mediated decay*). Ponieważ w powstającym transkrypcie znajduje się wiele miejsc rozpoznawanych przez spliceosom, proces składania musi być poddany ścisłej kontroli przez wiele czynników promujących bądź hamujących zajście składania w konkretnym miejscu. Okazuje się, że na wybór wpięcia bądź wypięcia alternatywnego eksonu ma wpływ także polimeraza II RNA odpowiedzialna za proces transkrypcji, który odbywa się w tym samym czasie co proces splicingu (4). Dzięki zachodzącym procesom alternatywnego składania możliwości powstawania różnorodnych transkryptów stają się praktycznie nieograniczone. Wyodrębniono kilka typów zachodzących procesów, prowadzących do powstania

alternatywnych izoform mRNA. Przykładem jest wyłączenie pierwszego (AFE, ang. *exclusion of first exon*) oraz ostatniego eksonu (ALE, ang. *exclusion of last exon*). Wypinaniu z transkryptu mogą podlegać również eksony (alternatywne) znajdujące się pomiędzy konstytutywnymi eksonami. Innymi wariantami są: wybór alternatywnych miejsc 5' i 3' splicingu czy retencja intronu (5) (**Ryc. 3**).



Rycina 3 Schematyczne przedstawienie możliwych procesów alternatywnego składania RNA. Wyłączanie, bądź włączanie eksonów alternatywnych może odbywać się na kilka różnych sposobów. (**A,B**) Poprzez aktywacje alternatywnych miejsc startu transkrypcji, bądź włączanie ostatniego eksonu z aktywnym PAS. (**C,F**) Włączanie lub wyłączenie eksonów alternatywnych znajdujących się wewnątrz genu. (**D,E**) Pojawienie się dodatkowych miejsc 3' i 5' skutkuje powstaniem skróconej wersji alternatywnego eksonu. (**G**) Część intronów może zostać nie wycięta.

Główne znaczenie i wpływ na zachodzący proces alternatywnego splicingu mają konkurujące ze sobą miejsca splicingowe. O wyborze konkretnych miejsc decydują odpowiednie czynniki białkowe wiążące się w regionach *cis* RNA, a rozpoznające ułożone w intronach bądź eksonach miejsca SRE (ang. *splicing regulatory elements*). Poprzez interakcje z konkretnymi czynnikami działającymi w *trans*, miejsca wiązania się białek SRE modulują przebieg splicingu. Do czynników działających w *trans* zalicza się grupę białek SR (ang. *Ser/Arg-rich proteins*) oraz hnRNP (ang. *heterogeneous ribonucleoprotein*). W skład SRE wchodzi czynniki pełniące funkcję promującą zachodzenie splicingu, wiążące się w intronach (ISE, ang. *intron splicing enhancer*) lub w eksonach (ESE, ang. *exon splicing enhancer*). Oprócz istnienia wzmacniaczy, w rejonach intronowych przyłączają się również inhibitory splicingu (ISS, ang. *intron splicing silencer*). Podobnie jest w rejonach eksonowych (ESS, ang. *exon splicing silencer*) (6) (**Ryc. 4**).



Rycina 4 Białka regulujące proces składania RNA, w tym alternatywny splicing. Procesu składania RNA stanowi kluczowy element w dojrzewaniu prekursorów RNA, który musi podlegać istotnej regulacji. Do najpopularniejszych białek kontrolujących splicing zaliczane są białka hnRNP oraz SR, które łącząc się do odpowiednich regionów w eksonach oraz intronach mogą wpływać negatywnie lub pozytywnie na cały proces. ISE – intronowy wzmacniacz splicing; ISS – intronowy wyciszacz splicing; ESE- eksonowy wzmacniacz splicing; ESS – eksonowy inhibitor splicing; PR – punkt rozgałęzień; Py – trakt polipirymidynowy. Białka hnRNP zostały zaznaczone schematycznie jako czynniki inhibujące a SR jako aktywujące. Tak jest w większości przypadków, jednak często te klasy białek mogą pełnić przeciwstawne funkcje.

W przypadku zachodzenia alternatywnego splicing, podstawową rolę odgrywa definiowanie odpowiednich regionów w nowopowstającym RNA, które fragmenty będą stanowić introny, a które eksony. U ssaków, fragmenty eksonowe są istotnie krótsze niż intronowe. Przeważa definiowanie tych pierwszych i z dużo większą częstotliwością dochodzić może do usunięcia alternatywnego eksonu niż do retencji intronu (7).

Do głównych czynników mających bezpośredni wpływ na zajście alternatywnego splicing należą białka SR oraz hnRNP. Poprzez oddziaływanie z elementami ISE i ESE, aktywują składanie RNA w konkretnych lokalizacjach. Białka SR i hnRNP uczestniczą ponadto w formowaniu się spliceosomu. Przykładowo, promowanie łączenia się U1 snRNP determinuje wybór miejsca splicingowego 5', a działanie U2AF determinuje wybór miejsca splicingowego 3' (8, 9). Związane białka SR z miejscami ESE wzmacniają łączenie się U2 snRNP do miejsc rozgałęzień (10).

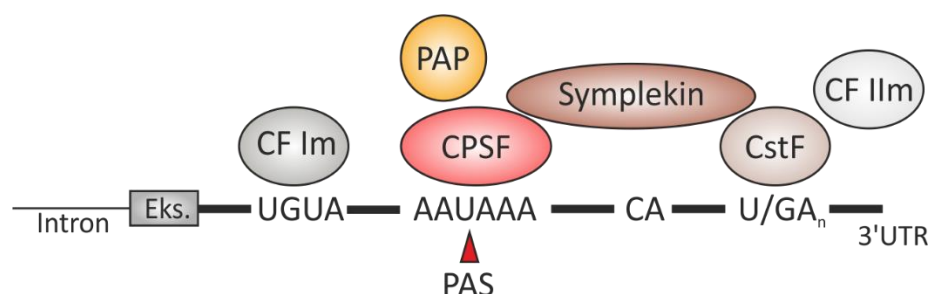
Białka hnRNP są dużą grupą regulatorowych białek wiążących się z RNA, których aktywność głównie wiązana jest z wyciszaniem splicing. Charakteryzują się one zróżnicowanym modelem działania. Dla przykładu, udowodniono, że białka hnRNP wchodzić mogą w interakcję z czynnikami regulującymi splicing, tj. SC35 i ASF/SF2 (8).

Jednym z lepiej poznanych przedstawicieli grupy białek hnRNP są białka PTBP (ang. *polypyrimidine tract binding protein*). Białko PTBP1, znane również pod nazwa hnRNP-I, łącząc się do fragmentów nowopowstającego, niedojrzałego transkryptu, bogatych w reszty C i U powoduje zahamowanie przebiegu splicing. Jednocześnie siła działania PTBP może być modulowana poprzez specyficzną lokalizację traktów polipirymidynowych. Pojawienie się

dotychczasowych miejsc wiązania dla PTB przy końcu 3' alternatywnie usuwanego eksonu, to wzmacnia ono wydajne wycięcie go ze składanego RNA (11).

1.1.1.3. Poliadenylacja na końcu 3' RNA – dobudowanie ogona poli(A)

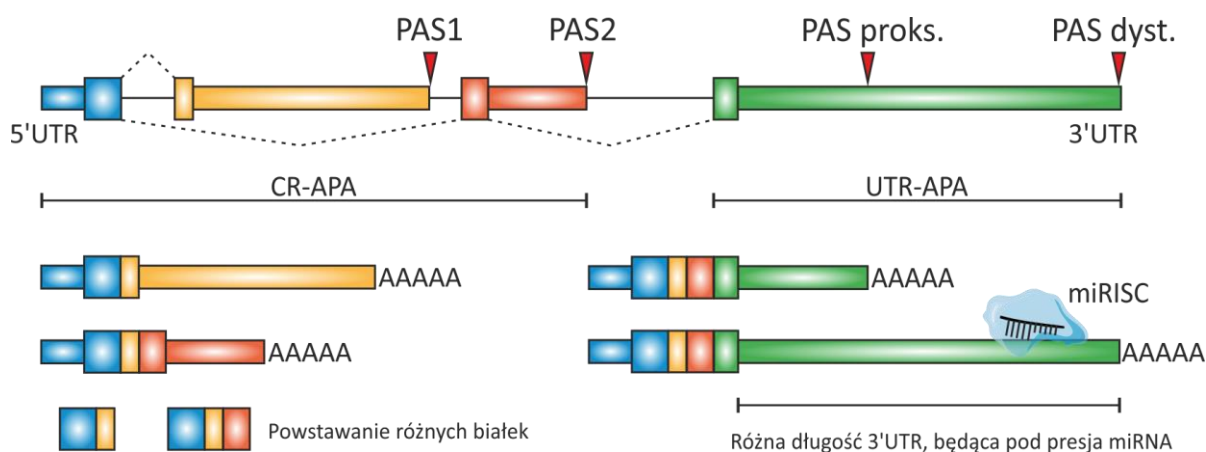
Obróbka końca 3' nowopowstałego transkryptu jest jednym z najważniejszych procesów zapewniających integralność cząsteczki RNA. Dodanie ogona poli(A) na końcu RNA jest zjawiskiem zachodzącym potranskrypcyjnie, dzięki któremu transport do cytoplazmy i proces translacji jest bardziej wydajny. Dodatkowo ogon poli(A) zabezpiecza nową cząsteczkę RNA przed degradacją (12). Zjawisko cięcia oraz poliadenylacji jest przeprowadzane przez duży kompleks białkowy, zawierający cztery główne subkompleksy – czynnik specyficzności cięcia i poliadenylacji (CPSF, ang. *Cleavage and polyadenylation specific factor*), czynnik stymulujący cięcie (CstF, ang. *Cleavage stimulation factor*), czynnik cięcia I (CF Im, ang. *Cleavage factor I*) oraz czynnik cięcia II (CF IIm, ang. *Cleavage factor II*) (13). Oprócz tego, w reakcji poliadenylacji końca 3' uczestniczą również polimerazy poliadenylowe (PAP, ang. *Poly(A) polymerase*), jądrowe białko wiążące się do poli(A) (PABPN, ang. *Nuclear poly(A) binding protein*) i białko Symplekin. Czynnik CPSF poprzez wiązanie się do sygnału poliadenylacji 5'-AAUAAA-3' (PAS, ang. *polyadenylation signal*) stanowi niezbędny element w procesie cięcia RNA (12). Miejsce PAS zlokalizowane jest w odległości 10-30 nukleotydów od miejsca, w którym zachodzi proces cięcia transkryptu. CstF rozpoznaje region bogaty w U bądź G i U. CstF swoją obecnością stabilizuje przyłączenie się CPSF do miejsca PAS (14). Czynnik cięcia CF Im wiąże się do motywów 5'-UGUA-3' znajdującymi się przed miejscem PAS. Dzięki zaangażowaniu wszystkich wyżej wymienionych czynników (**Ryc. 5**) możliwe jest endonukleolityczne cięcie RNA, i w następnej kolejności dobudowanie do wolnego końca 3' ogona poli(A) przez PAP.



Rycina 5 Kompleksy białkowe zaangażowane w proces poliadenylacji. Schematyczny obraz regionu niekodującego po stronie 3' (3'UTR) z zaznaczonymi sekwencjami do których wiążą się kompleksy białek zaangażowane w cały proces cięcia nowopowstałego RNA, a następnie poliadenylacji.

1.1.1.4. Alternatywna poliadenylacja (APA)

Zjawisko APA zostało po raz pierwszy odnotowane w roku 1980. Badanie opisywało geny kodujące immunoglobulinę M (*IgM*) (15) oraz reduktazę dihydrofolianowej (*DHFR*) (16). Od tamtego momentu, do roku 1997 naukowcom z całego świata udało się opisać jeszcze około 95 genów posiadających APA (17). Dzięki rozwojowi bioinformatyki oraz metod analiz całotranskryptomowych takich jak RNA-seq wiedza na temat procesów APA jest stale poszerzana i uzupełniana.



Rycina 6 Schematyczne przedstawienie możliwych procesów alternatywnej poliadenylacji. W sytuacji kiedy wybrane zostanie miejsce PAS wewnątrz sekwencji kodującej białko (CR-APA) prowadzić to może do powstania białka o zmienionej budowie. Aktywność PAS w 3'UTR skutkuje powstawaniem 3'UTR o różnej długości (UTR-APA), co prowadzi do wzmożonej regulacji konkretnej izoformy przez takie cząsteczki jak miRNA.

Niemal wszystkie mRNA powstające w komórkach eukariotycznych oraz wiele niekodujących RNA to cząsteczki poliadenylowane (18). Wiele genów na końcu 3' posiada kilka alternatywnych miejsc poliadenylacji (APAS, ang. *alternative polyadenylation sites*). Aktywność alternatywnych miejsc prowadzi do powstawania izoform, mogących różnić się długością 3'UTR (19). Uwzględniając lokalizację miejsc poliadenylacji, zjawisko APA można podzielić na dwie główne kategorie (**Ryc. 6**). W pierwszej z nich poliadenylacja będzie zachodzić wewnątrz sekwencji kodującej eksony bądź introny i poprzez swoją aktywność prowadzi do istotnego skrócenia transkryptu kodującego białko skrócone o koniec karboksylowy. Ten typ APA nazywany jest alternatywną poliadenylacją regionów kodujących (CR-APA, ang. *coding region-APA*) (20). W drugiej sytuacji alternatywne miejsca poliadenylacji zlokalizowane są w 3'UTR. Ich zmienna aktywność prowadzi do powstawania transkryptów o różnej długości 3'UTR (UTR-APA, ang. *untranslated region-APA*). Wyróżnia się PAS proksymalne (PAS proks.), które poprzez swoją aktywność istotnie skracają 3'UTR. Natomiast aktywność PAS dystalnych skutkuje wydłużeniem 3'UTR. UTR-APA nie prowadzi do zmian w

sekwencji kodującej białko. Zmiany w długości 3'UTR mogą mieć jednak istotny wpływ na stabilność mRNA, efektywność z jaką mRNA będzie przepisywane na białko i transport. Dłuższe 3'UTR podlegać mogą częściej regulacji przez cząsteczki miRNA (ang. *microRNA*), oraz białka wiążące się do RNA (RBP, ang. *RNA binding protein*).

Badania prowadzone w ciągu ostatnich dwóch dekad wykazały tysiące przypadków, w których izoformy APA ulegają różnej ekspresji w różnych warunkach komórkowych. Globalna analiza sekwencjonowania z wykorzystaniem specjalnie zaprojektowanych bibliotek cDNA, takich jak EST (ang. *expressed sequence tag*) wykazała zmienność w częstości występowania APA w różnych tkankach (21). W części analizowanych tkanek dominował konkretny wariant RNA wynikający z wyboru odpowiedniego miejsca poliadenylacji, który z kolei w innych typach komórek występował rzadko. Dla przykładu PAS dystalne są preferowane w tkankach neuronalnych, gdy w przypadku komórek krwi obserwowany jest odwrotny trend (21). Również w przypadku różnych etapów procesu spermatogenezy zaobserwowano zmiany w wyborze PAS.

W przypadku komórek nowotworowych zaobserwowano globalny trend występowania krótkich 3'UTR (22). Badanie zmian długości 3'UTR w nowotworach o różnym stopniu zaawansowania wykazało charakterystyczne zmiany w aktywności APA (23). Takie zjawisko wskazuje na bezpośrednie zaangażowanie procesu APA w zwiększony potencjał podziałów komórek nowotworowych oraz ich rozwój. Taka wiedza może służyć jako parametr prognostyczny oraz być wykorzystana jako ważny indykator stopnia zaawansowania choroby w diagnostyce nowotworów. Dzieje się już tak w przypadku badania komórek nowotworowych gruczołu piersiowego, gdzie okazało się, że różne typy komórek nowotworowych mogą wykazywać inne aktywności APA (24). W przypadku powszechnie stosowanych w badaniach modelowych komórek raka piersi MCF-7 dochodziło do skracania 3'UTR w porównaniu do komórek prawidłowych. Z kolei w innym modelu, komórkach MB231, wykazano obecność dłuższych 3'UTR w przypadku większości analizowanych mRNA.

Okazuje się, że procesowi APA podlegać mogą nie tylko pre-mRNA, ale również prekursorowe długie niekodujące RNA. Badanie oparte na analizie mysich lncRNA wykazało, że w przypadku co najmniej 66% znanych cząsteczek lncRNA zachodzi proces APA, jednak w odróżnieniu od APA w genach kodujących białko, w większości przypadków zachodził na początku transkryptów wywołując istotne skrócenie izoform lncRNA (25). Dla przykładu,

transkrypt NEAT1 (ang. *nuclear paraspeckle assembly transcript 1*) ma dwie izoformy – dłuższą (23000 nt) oraz istotnie skróconą (3700 nt). Długość skróconej izoformy wynika z aktywności PAS w początkowym rejonie transkryptu. W rezultacie występowanie różnych izoform NEAT1 wpływał na wydajność tworzenia się skupisk jądrowych (ang. *paraspeckles*) (26).

Regulacja aktywności APA zależy może w dużej mierze od działania RBP (27). Niektóre RBP hamują terminację transkrypcji w poszczególnych PAS poprzez blokowanie wiązania się podstawowych czynników związanych z procesem poliadenylacji, a niektóre zwiększają wybór odpowiednich PAS poprzez rekrutację kluczowych czynników dla terminacji i poliadenylacji. Oba kierunki działania mogą być regulowane przez to samo białko w różnych kontekstach, jak pokazano dla neuronalnego RBP NOVA. Jego wiązanie w pobliżu PAS inhibuje proces poliadenylacji, podczas gdy wiązanie w innym miejscu RNA promuje ten proces (28).

1.1.1.5. Zależność pomiędzy splicingiem, a poliadenylacją

Główne mechanizmy, tj. splicing oraz poliadenylacja, będące podstawą w dojrzewaniu transkryptów RNA, bardzo często są ze sobą ściśle powiązane i wzajemnie na siebie wpływają. Jednym z przykładów jest bezpośrednia zależność pomiędzy jednym z głównych czynników odpowiedzialnych za cięcie transkryptu – CPSF-160 – a białkiem U1A snRNP. Poprzez bezpośrednią interakcję białko-białko, U1A wzmacnia wiązanie się CPSF-160 do PAS i tym samym podnosi wydajność procesu cięcia i poliadenylacji (29). Inną wcześniej opisaną zależnością jest oddziaływanie pomiędzy kompleksem CPSF a U2 snRNP (30) oraz U2AF65 a CF Im (31). Okazuje się, że sekwencje intronowe zaangażowane w definiowanie ostatniego eksonu, a zlokalizowane w pobliżu PAS mogą istotnie wpłynąć na aktywność procesu poliadenylacji (32). Delecja miejsc splicingowych 5' nie wpływa na wydajność poliadenylacji, podczas gdy wprowadzenie mutacji 3' miejsc splicingowych może blokować ten proces, potwierdzając że aktywność miejsc 3'ss oraz czynników splicingowych wiążących się z tymi miejscami jest ważna dla stymulacji procesu poliadenylacji. Zależność ta może również działać w przeciwną stronę. Mutacja w proksymalnym PAS prowadzić może do istotnych zaburzeń związanych z wycięciem ostatniego intronu w genach wieloeksonowych (33).

1.2. Długie niekodujące RNA (lncRNA)

Klasyczne postrzeganie ekspresji genów w dużej mierze opierało się na znanym dogmacie, opisującym szlak DNA - mRNA - białko. Jednakże, na przestrzeni ostatnich dekad, bazując na wynikach analiz RNA-seq okazało się, że aż 75% ludzkiego genomu podlega transkrypcji. Dodatkowo ilość transkryptów niekodujących białka, tzw. długich niekodujących RNA (lncRNA), jest 4 razy większa niż transkryptów, z których to białka powstają (34). Udowodniono również, że lncRNA stanowią bardzo zróżnicowaną grupę cząsteczek, również pod względem ich powstawania. Sekwencje dla lncRNA mogą być zlokalizowane w wielu miejscach w genomie, tj. wewnątrz intronów genów kodujących pre-mRNA, w samych eksonach jak również w przestrzeniach międzygenowych. Tak duża różnorodność utwierdza w przekonaniu o modularnym modelu organizacji genów, w którym to możliwe jest przepisywanie sekwencji na wiele transkryptów sensowych, antysensowych, kodujących oraz niekodujących białko z jednego *locus*. Tak duża liczba zidentyfikowanych cząsteczek lncRNA rodzi zatem pytanie: skoro komórka zużywa tak dużo energii na biosyntezę tych cząsteczek, to w jaki sposób mogą być one przydatne dla jej funkcjonowania? Początkowo uznawano, że mnogość powstających lncRNA bierze się z mało specyficznego działania polimerazy II RNA (35). Na to mógłby dodatkowo wskazywać fakt bardzo niskiego poziomu zakonserwowania sekwencji lncRNA. Możliwym wytłumaczeniem jest to, że lncRNA w swojej sekwencji nie posiadają otwartej ramki odczytu (ORF, ang. *open reading frame*). Jednakże występują również takie cząsteczki lncRNA, które posiadają w swojej sekwencji krótkie zakonserwowane fragmenty, pełniące znaczącą rolę w procesach komórkowych. Przykładem może być gen *Xist*, który odpowiedzialny jest za wyciszenie ekspresji genów pochodzących z jednego z chromosomów X, w przypadku komórek żeńskich. W komórkach męskich pełni rolę regulatorową (36). Okazuje się jednak, że więcej lncRNA zaangażowanych jest w istotne procesy komórkowe, tj. różnicowanie, rozwój czy podziały komórkowe (37). Ich działanie opiera się na kontrolowaniu ekspresji genów poprzez wpływ na miejsca promotorowe (38), piętnowanie genomowe (ang. *genome imprinting*) (39) czy utrzymanie odpowiedniego uformowania chromatyny w jądrze komórkowym (40).

1.3. Krótkie niekodujące RNA (sncRNA)

Do tej grupy zalicza się klasę małych jądrowych RNA (snRNA, *small nuclear RNA*), które tworzą cząsteczki RNA o długości nieprzekraczającej 200 nukleotydów. Zlokalizowane są głównie w jądrze komórkowym organizmów eukariotycznych. Do ich podstawowych funkcji należy regulacja ekspresji genów. Podczas tego procesu snRNA związane są z zestawem specyficznych białek, których kompleksy nazywane są jako małe jądrowe rybonukleoproteiny (snRNP, ang. *small nuclear ribonucleoprotein*). Podklasa snRNP nazywana jest małymi jąderkowymi RNA (snoRNA, ang. *Small nucleolar RNA*), które zlokalizowane są w jąderku i biorą udział w procesach dojrzewania takich cząsteczek jak rRNA, tRNA czy snRNA.

Kolejną klasę tworzą cząsteczki RNA związane z białkiem Piwi (piRNA, ang. *Piwi interacting RNA*). Stanowią liczną grupę krótkich niekodujących RNA występujących w komórkach eukariotycznych (41). Ze swoją długością wynoszącą 25-30 nukleotydów są nieznacznie większe niż inne regulatorowe krótkie RNA. Poprzez formowanie kompleksów RNA-białko, piRNA zaangażowane są w epigenetyczne jak i potranskrypcyjne wyciszenie retrotranspozonów oraz innych elementów genetycznych w komórkach biorących udział podczas procesu spermatogenezy (42).

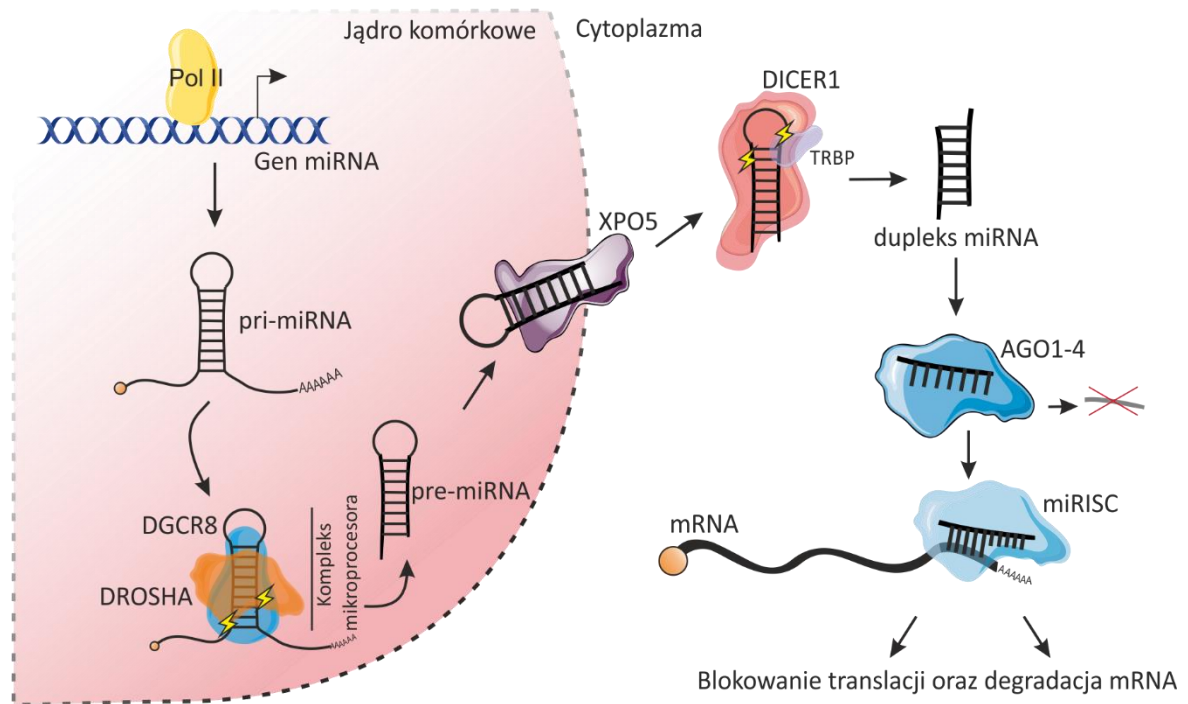
Krótkie interferujące RNA (siRNA, ang. *small interfering RNA*) stanowią kolejną grupę krótkich, dwuniciowych cząsteczek RNA, których długość wynosi od 20-25 nukleotydów. SiRNA odgrywają istotną rolę w szlaku interferencji RNA (RNAi, ang. *RNA interference*) (43). Poprzez w pełni komplementarne wiązanie się jednej z nici siRNA, będącej elementem kompleksu wyciszającego indukowanego przez RNA (RISC, ang. *RNA induced silencing complex*), do mRNA prowadzą do jej degradacji i tym samym regulują poziom ekspresji genów. Mechanizm działania siRNA po raz pierwszy został odkryty w komórkach roślinnych (44), po czym prawie natychmiast został on wykorzystany jako molekularne narzędzie do kontrolowania poziomu ekspresji konkretnych genów w innych organizmach (45). Działanie siRNA oparte jest na podobnych zasadach jak w przypadku trzeciego reprezentanta grupy krótkich niekodujących RNA – miRNA.

1.3.1. mikroRNA

W roku 1993, Victor Ambros wraz z Rosalind Lee oraz Rhondą Feinbaum, badając rozwój embrionalny *C. elegans*, odnotowali nowe zjawisko. Gen *lin-4*, pełniący istotną rolę podczas rozwoju badanego organizmu, okazał się nie być matrycą do produkcji białka, a lncRNA (46). Dodatkowo, z transkryptu tego powstawała para krótkich cząsteczek RNA. Jak się później okazało, posiadająca komplementarność do wielu miejsc w 3'UTR. Zależność poziomu mRNA genu *lin-14*, poprzez bezpośrednie działanie *lin-4* została opisana (47), a następnie funkcjonalnie potwierdzona poprzez wykazanie negatywnego wpływu *lin-4* na poziom białka *lin-14*, przy jednoczesnym braku obniżenia mRNA *lin-14* (48). Wyniki te po raz pierwszy zaprezentowały model działania krótkich, niekodujących cząsteczek RNA – mikroRNA (miRNA), które poprzez wiązanie się do regionu 3'UTR mRNA regulują ich poziom ekspresji, a w następstwie także ilość powstających na bazie tych mRNA białek.

Po prawie trzech dekadach badań wiemy, że *lin-4* okazał się być pierwszym reprezentantem bardzo licznej grupy RNA nazwanych miRNA (49). Jest to rodzina krótkich (21-25 nt), niekodujących cząsteczek RNA, których obecność została potwierdzona w wielu różnych organizmach. Wysoki stopień zakonserwowania ewolucyjnego tych cząsteczek wskazuje na ich istotną rolę w toku ewolucji (50). miRNA stanowią ważny element w prawidłowym funkcjonowaniu komórek roślinnych oraz zwierzęcych. Szacuje się, że ok 50% genów kodujących białka jest pod kontrolą tych krótkich cząsteczek regulatorowych. Badania funkcjonalne wskazują, że niemal każdy znany proces komórkowy zależny jest od miRNA, co przekłada się na rolę regulatora komórkowej homeostazy. Przeprowadzone badania na modelach zwierzęcych, a w szczególności na myszach, polegające na wyłączeniu aktywności poszczególnych miRNA wykazały wpływ na działanie wielu narządów. Były to między innymi takie organy jak: układ szkieletowy (51), mózg (52), oczy (53), neurony (54), serce (55), mięśnie szkieletowe (56), płuca (57), nerki (58), wątroba (59) czy trzustka (60). Mając na uwadze istotną rolę, jaką odgrywają cząsteczki miRNA, można przypuszczać że jakakolwiek nieprawidłowość w powstawaniu bądź działaniu miRNA, będzie miała konsekwencje w funkcjonowaniu komórki, następnie tkanki, a w rezultacie całego organizmu. Przykładem takiego zaburzenia jest zmieniony profil ekspresji miRNA i jego następstwa obserwowane w komórkach nowotworowych.

1.3.1.1. Biogeneza miRNA - klasyczna droga powstawania miRNA



Rycina 7 Schematyczne przedstawienie procesu biogenezy miRNA

miRNA powstają ze swoich pierwotnych transkryptów zwanych pri-miRNA (ang. *primary microRNA*). Transkrypty te tworzone są w wyniku działania polimerazy II RNA, jako takie opatrzone są czapeczką na końcu 5' (61), lecz nie zawsze poliadenylowane na końcu 3'. Może mieć to związek z aktywnością działania mikroprocesora zaangażowanego w dojrzewanie miRNA, który może wymuszać wcześniejszą terminację transkrypcji (62). Utworzona wewnątrz pri-miRNA krótka struktura typu spinki do włosów (ang. *hairpin structure*) rozpoznawana jest przez oraz białko rozpoznające podwójną nić RNA - DGCR8 (ang. *DiGeorge critical region 8*) oraz białko z rodziny RNaz III – DROSHA. Razem białka te tworzą kompleks mikroprocesora, którego produktem działania jest ok. 70 nukleotydowa cząsteczka pre-miRNA, ciągle tworząca strukturę typu spinki do włosów. W naturze występują tysiące cząsteczek RNA, które tworzą strukturę przypominającą spinkę do włosów. Jednak tylko część z nich jest rozpoznawana jako cząsteczka prekursorowa w pri-miRNA i przetwarzana dalej w w kolejnych etapach biogenezy miRNA. Jest to wynikiem ścisłej kontroli biogenezy miRNA przez mikroprocesor. Kompleks ten faworyzuje pojedynczą strukturę spinki, której długość wynosi około 35 par zasad, oflankowaną z obu stron jednoniciowymi sekwencjami RNA (63). Nowopowstałe pre-miRNA następnie zostaje przetransportowane z jądra do cytoplazmy przez Eksportynę 5 (XPO5) (64, 65). W cytoplazmie na skutek działania kolejnej RNazy typu III – Dicer,

z pre-miRNA wycinany jest dwuniciowy, 20-25 nukleotydowy dupleks miRNA/miRNA* (66)(8). W biologii stosuje się dodatkową nomenklaturę miRNA opierającą się o nici, z które wywodzi się dojrzała cząsteczką miRNA. Wyróżniamy w takim wypadku cząsteczkę miRNA z nici od strony 5' - miRNA-5p oraz strony 3' - miRNA-3p. Co charakterystyczne, u człowieka DICER współpracuje razem z innym białkiem wiążącym się do dwuniciowej formy RNA – TRBP (ang. *dsRBP trans-activation-responsive RNA-binding protein*). Jedna z nici w dupleksie, reprezentująca dojrzałą formę miRNA, rozpoznawana przez białko Argonaute (AGO), wprowadzana jest do kompleksu wyciszającego ekspresję genów przez aktywność miRNA (miRISC, ang. *miRNA-induced silencing complex*). Druga z nici – miRNA* – zostaje zdegradowana i w tym układzie nie odgrywa ważnej roli w procesie kontroli ekspresji genów (67). MiRNA będąc kluczowym elementem w kompleksie miRISC, mimo niepełnej komplementarności (w odróżnieniu do siRNA) do docelowego regionu w mRNA (68), naprowadza oraz powoduje związanie całego kompleksu i prowadzi w ten sposób do zablokowania translacji bądź do deadenylacji i degradacji związanej cząsteczki mRNA (**Ryc. 7**). Potencjalnie oba miRNA, powstające z dupleksu pre-miRNA (miRNA oraz miRNA*) stanowią mogą funkcjonalną cząsteczkę.

W ostatnich latach wykazano pewne wyjątki od reguły, w których to białka DROSHA, DICER1 oraz XPO5 stanowią niezbędny element podczas biogenezy miRNA. W przeprowadzonym w 2016 roku badaniu (69), po inaktywacji każdego z tych białek z osobna i sekwencjonowaniu całej puli krótkich RNA (ang. small RNA-seq) okazało się, że gdy usunięcie Droshy skutkuje istotnym spadkiem poziomu wszystkich miRNA, tak w przypadku zablokowania białka Dicer poziom niektórych z miRNA pozostaje niezmienny. Wynik ten świadczyć może o istnieniu alternatywnych dróg powstawania miRNA, niezależnych od kanonicznej ścieżki biogenezy. Pełen proces dojrzewania miRNA może okazać się dużo bardziej złożony niż postrzegano do tej pory.

1.3.1.2. Mirtrony – alternatywna ścieżka biogenezy miRNA

W 2007 roku udokumentowano istnienie nowej grupy miRNA, tzw. mirtrony (70). Istnienie tej puli, 14 miRNA u muszki owocowej oraz 4 miRNA u nicieni, stanowiło niezwykle odkrycie ze względu na fakt, że zlokalizowane są one w intronach genów kodujących białko i co najistotniejsze mogą powstawać bez udziału DROSHA. Sytuacja taka jest możliwa dzięki zachodzącemu procesowi składania eksonów przy jednoczesnym wycinaniu intronów z pre-

mRNA. W tym właśnie momencie z wyciętego intronu, tzw. lariaty formuje się pre-miRNA, który następnie jest przetwarzany w dalszych etapach biogenezy miRNA. W konsekwencji powstaje dojrzała cząsteczka. Istotnym elementem w tym systemie jest odpowiednia długość intronu, który po wycięciu z pre-mRNA swoją długością zbliżony jest do naturalnie występujących pre-miRNA.

1.3.1.3. Badania genomowe oraz identyfikacja kanonicznych cząsteczek miRNA

Pomimo, że cząsteczki miRNA odkryte zostały na początku lat 90. XX wieku, to dopiero po około 15 latach technika masowego sekwencjonowania krótkich RNA zrewolucjonizowała badania nad miRNA (71, 72). Należy jednak pamiętać, że nie każda odkryta w sekwencjonowaniu krótka cząsteczka jest cząsteczką miRNA. Wiele z wyników RNA-seq może stanowić, np. wchodzące w interakcje z białkami Piwi krótkie RNA (piRNA, ang. Piwi-interacting RNA) oraz endogennie występujące krótkie interferujące RNA (siRNA, ang. small interfering RNA) (72). W celu uniknięcia takich fałszywie pozytywnych odkryć nowych miRNA, postanowiono wprowadzić jasne i klarowne wytyczne co do definicji czy odkryta cząsteczka faktycznie jest miRNA. Przede wszystkim koniec 5' takiego RNA musi być zgodny z przypisaną anotacją. Dodatkowo sekwencja musi być ściśle komplementarna z miejscem w genomie, gdzie przewidywana jest charakterystyczna struktura spinki do włosów i być komplementarna z drugą nicią RNA, wspólnie tworząc dupleks RNA (73).

Stale rosnąca liczba nowo odkrywanych miRNA wymusiła stworzenie baz danych zbierających te wszystkie informacje. Jedną z najpopularniejszych baz jest miRBase (<http://microrna.sanger.ac.uk>) (74), obecnie zawierająca 1917 ludzkich prekursorów miRNA oraz 2656 dojrzałych cząsteczek miRNA (wersja V22) (75). Ze względu na ciągłe pojawianie się nowych cząsteczek, dużym wyzwaniem stała się weryfikacja tych doniesień.

W 2019 J. Alles wraz z współpracownikami przeprowadziła eksperyment, mający na celu potwierdzenie ile tak naprawdę zdeponowanych krótkich cząsteczek w najpopularniejszej bazie danych miRBase jest rzeczywiście cząsteczką miRNA. W badaniu przeanalizowano 28 866 dostępnych sekwencjonowań krótkich RNA, zawierających niemal 364 miliardy odczytów. Wyniki analiz bioinformatycznych zidentyfikowały 65% z 24 127 potencjalnych cząsteczek miRNA, jako wyniki fałszywie pozytywne. Ostatecznie wyniki grupy potwierdziły istnienie 2300 cząsteczek miRNA, z czego 1115 znajduje się w bazie danych miRBase V22 (76).

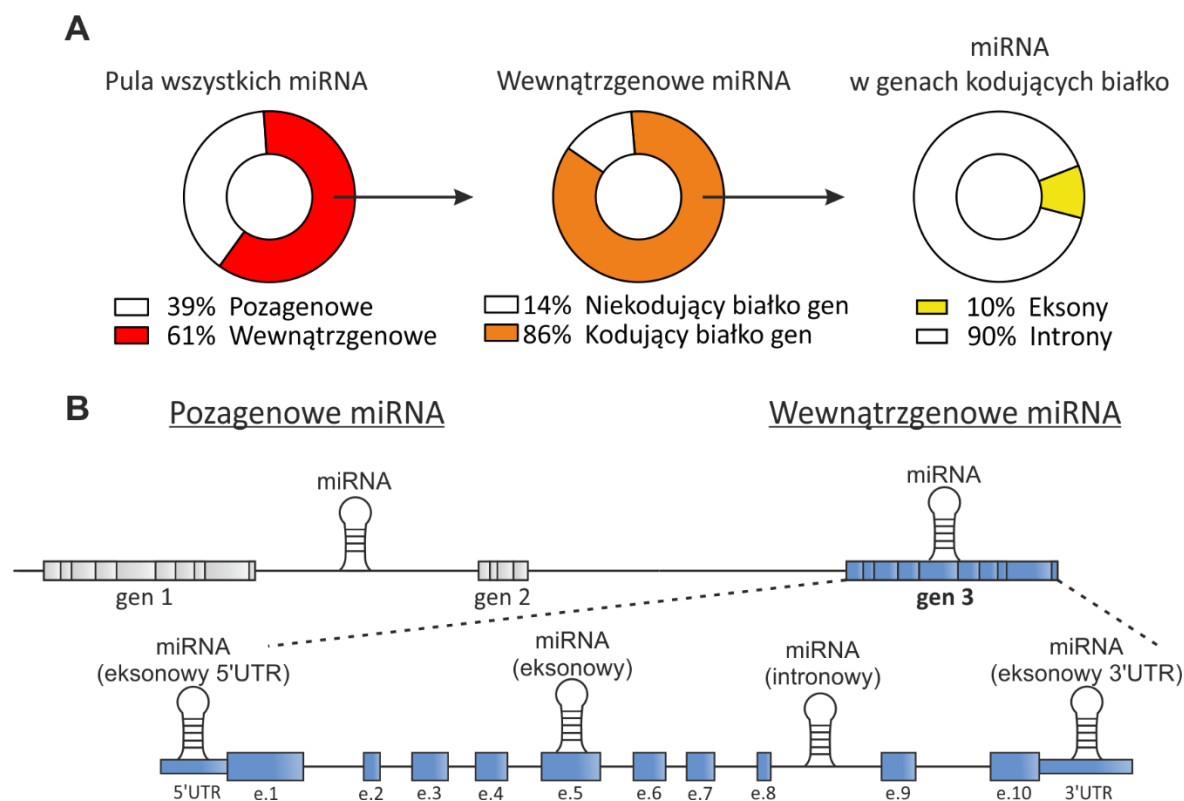
Oprócz bazy danych miRBase istnieje jeszcze kilka innych baz, zbierających informacje o istniejących miRNA. Jednym z takich zbiorów jest miRGenDB (<http://mirgenedb.org>) – baza danych zawierająca mniejszą liczbę zdeponowanych cząsteczek, ale z dużą pewnością co do jakości wyników (77). Inną bazą jest miRCarta (<https://mircarta.cs.uni-saarland.de>). Zbiór ten charakteryzuje się również zweryfikowanymi eksperymentalnie danymi oraz zestawieniem literaturowo-potwierdzonych cząsteczek miRNA (78).

1.3.1.4. Genomowa lokalizacja miRNA

Obecnie mechanizm biogenezy miRNA z ich prekursorów jest już dość dobrze opisany. Jednakże wpływ oraz kontrola mechanizmów transkrypcyjnych oraz potranskrypcyjnych na wydajność biogenezy miRNA nadal wymaga dalszej analizy i badań. Genomowa lokalizacja genów kodujących prekursorów miRNA jest główną determinantą, sugerującą jakie procesy będą zaangażowane, jak wydajnie biogeneza tych cząsteczek będzie przebiegać i jaki wpływ może mieć to na poziom ich ostatecznej ilości w komórce. Opierając się na genomowej lokalizacji, miRNA możemy ogólnie sklasyfikować jako “wewnątrzgenowe” (ang. *intragenic*) oraz “pozagenowe” (ang. *intergenic*). MiRNA wewnątrzgenowe zlokalizowane są najczęściej w intronach genów kodujących białko oraz lncRNA. Sekwencje kodujące miRNA zlokalizowane również mogą być w częściach eksonowych swoich genów-gospodarzy (ang. *host gene*). W dużej części przypadków podlegają tej samej regulacji transkrypcyjnej, co geny wewnątrz których się znajdują. Jednak badania wysokoprzepustowe wykazały, że miRNA wewnątrzgenowe, zlokalizowane w intronach swoich genów-gospodarzy, mogą powstawać jednocześnie jako niezależne jednostki transkrypcyjne, z wykorzystaniem własnych miejsc promotorowych (79, 80). Faktem jest, że około 20% miRNA znajdujących się w intronowych częściach genów-gospodarzy, poprzez sprzężenie zwrotne reguluje ich własny poziom ekspresji (81). Oprócz tego wyróżniamy miRNA znajdujące się na połączeniu intron-ekson oraz miRNA, których transkrypcja przebiega na nici antysensownej. miRNA pozagenowe, zlokalizowane pomiędzy zdefiniowanymi genami, posiadają swoje własne miejsca promotorowe i powstają również przy udziale polimerazy II RNA oraz w rzadkich przypadkach polimerazy III RNA. Z przeprowadzonych badań wynika, że miRNA pozagenowe mogą powstawać przy udziale dwóch dodatkowych mechanizmów (80). Pierwszy, nazwany “przeciekaniem” (ang. *readthrough*), bazuje na zjawisku, podczas którego transkrypcja genu przebiega jeszcze kilka tysięcy par zasad za miejscem PAS, w którym dochodzi do przecięcia

RNA i dobudowania ogona poli(A). Przykładem miRNA, którego biogeneza może zależeć od opisanego przecieku transkrypcji, jest miR-21 (82). Drugim z mechanizmów, który odpowiedzialny jest za powstawanie prekursorów miRNA jest mechanizm “rozbieżnej transkrypcji” (ang. *divergent transcription*). W czasie tego procesu, ze wspólnego miejsca promotorowego dwie polimerazy RNA przeprowadzają niezależnie od siebie transkrypcję genów nici sens oraz antysens (83).

Według miRBase (miRBase v.20, N=1,870; <http://www.mirbase.org/>) około 39% ludzkich miRNA stanowi grupę miRNA pozagenowych, a zatem w przybliżeniu 61% zdeponowanych w bazie danych miRNA stanowią cząsteczki wewnątrzgenowe. Wśród nich, około 86% zlokalizowanych jest w genach kodujących białko, a 84% z nich transkrybowana jest na tej samej nici kodującej, co pre-mRNA ich genu-gospodarza. Zdecydowana większość wewnątrzgenowych miRNA zlokalizowana jest w intronach – 90%. Pozostałe 10% przypisane zostało jako miRNA eksonowe, przy czym 80% eksonowych miRNA lokalizuje w niekodujących fragmentach genów (5' oraz 3' UTR, ang. *untranslated region*) (84) (**Ryc. 8**).



Rycina 8 Genomowy podział cząsteczek miRNA.

1.3.1.5. Mechanizm działania miRNA

Działanie miRNA rozpoczyna się od załadowania pojedynczej cząsteczki miRNA do kompleksu RISC. Będąc związany w kompleksie, miRNA nakierowuje RISC na mRNA poprzez komplementarne połączenie z niekodującym regionem w 3'UTR. Uznaje się, że fragment pomiędzy nukleotydami 2 i 8 w cząsteczce miRNA stanowi niezbędny element dla jego wiązania się do regulowanego miejsca w mRNA i nazywany jest miejscem źródłowym (ang. *seed region*) (68). Działanie miRNA można podzielić na dwa główne etapy. We wczesnej fazie działania regulatorowego miRNA następuje zablokowanie procesu translacji, przy braku zmian w stabilności docelowego mRNA. W kolejnych etapach dochodzi do degradacji takiego mRNA (85, 86).

Uwzględniając bardzo krótką sekwencję wymaganą do wiązania się miRNA z docelowym miejscem w 3'UTR regulowanego transkryptu, można przyjąć że cząsteczki miRNA odgrywają szczególną rolę w kształtowaniu profilu ekspresji większości powstających genów. Biorąc do analizy zidentyfikowane regiony 3'UTR w ludzkim transkryptomie okazuje się, że w przypadku każdej z 90 silnie zakonserwowanych ewolucyjnie rodzin miRNA, współdzielących region źródłowy, istnieje średnio ponad 300 miejsc w 3'UTR do których będą wiązać się miRNA

na zasadzie 7-8 nukleotydowego sparowania. Jeżeli uwzględni się wiązanie regionu 6-nukleotydowego, to liczba potencjalnie regulowanych regionów wzrasta średnio do 500 dla jednej rodziny miRNA (87). Do niektórych 3'UTR jedna cząsteczka miRNA może przyłączyć się w kilku miejscach, jak również jeden 3'UTR w swojej sekwencji może zawierać zakonserwowane ewolucyjnie miejsca wiązania dla kilku różnych rodzin miRNA.

Po związaniu się miRNA z docelowym miejscem w mRNA, białko Argonaut (AGO) rekrutuje reprezentanta z grupy białek z licznymi powtórzeniami glicynowo-tryptofanowymi (GW). W przypadku ssaków jest to białko TNRC6. Białka te mają duże powinowactwo do AGO oraz zawierają domenę wyciszającą. Domena ta posiada reszty tryptofanowe, niezbędne do wiązania się z rybonukleazami specyficznymi dla miejsc poli(A) - PARN2 lub dla represora katabolitów węglowych (CCR4). CCR4 jest negatywnym regulatorem transkrypcji - jego działanie polega na skracaniu ogonów poli(A) w mRNA. Dodatkowym czynnikiem wpływającym na obniżenie poziomu regulowanego RNA jest rekrutacja kolejnych białek wiążących się do ogona poli(A), które prowadzą do deadenyacji i w następstwie do degradacji cząsteczek mRNA. Skrócone, bądź pozbawione całkowicie ogonów poli(A) cząsteczki mRNA zostają pozbawione czapeczki na końcu 5', co skutkuje aktywacją egzonukleazy 1 (XRN2), która degradowuje mRNA z aktywnością 5'-3' (68).

1.3.1.6. Znaczenie cząsteczek miRNA w rozwoju oraz w chorobie

miRNA stanowią cząsteczki regulujące liczne procesy wewnątrzkomórkowe, a ich obecność jest niezbędna do prawidłowego rozwoju organizmu. Wykazano, że nokaut genu *Dicer* w mysim modelu prowadzi do śmierci organizmu w 7. dniu rozwoju embrionalnego (88). Oprócz rozwoju embrionalnego udowodniono, że miRNA kontrolują metabolizm, podziały komórkowe czy dojrzewanie komórek nerwowych (89). Do procesów, w których miRNA pełnią znaczącą funkcję regulatorową zaliczone są rozwój mózgu (90), różnicowania komórek mięśniowych (91) czy prawidłowe różnicowanie komórek macierzystych (92). Zaburzenia w ilości oraz funkcjonowaniu cząsteczek miRNA prowadzić może do licznych patologii, z pośród których dużą grupę stanowią różne etapy nowotworzenia. Ponad 50% miRNA zlokalizowanych jest we wrażliwych miejscach w genomie, w których obserwowane są częste delecje lub insercje. Potwierdzają to wyniki badań prób pochodzących od pacjentów onkologicznych (93). miRNA, w zależności od pełnionej roli, dzieli się na dwie kategorie: onkogeny (ang. *OncomiR*) oraz supresory nowotworzenia (94). OncomiR poprzez blokowanie działania czynników

supresorowych promują rozwój nowotworów. Przykładami oncomiR są miR-221/222 (95), miR-27a (96) czy miR-21 (97–99). W odróżnieniu od oncomiR, istnieje grupa miRNA których rolą jest blokowanie powstawania nowotworów, poprzez bezpośrednią supresję działania onkogenów. Przykładami takich miRNA są miR-145 (100), miRNA z rodziny let-7 (101) oraz miR-205 (102).

1.3.2. mikroRNA-21

miR-21 już niemal od dwóch dekad stanowi cel badań wielu naukowców na całym świecie (103). Ilość i powszechność występowania tej cząsteczki w komórkach różnego typu daje podstawy do twierdzenia, iż zaangażowany jest w istotne procesy komórkowe. miR-21 występuje niemalże w każdym typie komórek ssaczych i odgrywa szczególną rolę w takich procesach jak rozwój komórki, jej podziały czy apoptoza. miR-21 jako jeden z pierwszych był intensywnie badany i analizowany pod kątem procesów biogenezy i dojrzewania. Podwyższony poziom tego miRNA obserwowany jest w większości nowotworów, jak również chorób układu sercowo-naczyniowego. Przykładowo, badania przeprowadzone na myszy z podwyższoną ekspresją miR-21 wykazały, że ten miRNA bierze czynny udział w powstawaniu chłoniaka złośliwego (ang. *Malignant B-cell lymphoma*) (104). Pomimo intensywnych badań, pozostaje wciąż wiele niewiadomych dotyczących biogenezy miR-21, a przede wszystkim mechanizmów regulujących ten proces. Poznanie ich może stanowić możliwość do opracowania narzędzi pozwalających wpłynąć bezpośrednio na procesy w które zaangażowane jest miR-21.

1.3.2.1. Genomowa lokalizacja oraz regulacja biogenezy miR-21

Gen kodujący miR-21 zlokalizowany jest na chromosomie 17 (17q.23.1), w intronie 11. genu kodującego wakuolarnie białko membranowe VMP1 (ang. *Vacuole membrane protein 1*). Prowadzone badania wykazały, że w obrębie genu kodującego *pri-miR-21* znajdują się silnie zakonserwowane miejsca promotorowe (105). W 2008 roku została przeprowadzona dokładna analiza miejsc promotorowych dla transkrypcji prekursora miR-21 (*miPPR-21*) (106)(50). Miejsce to zawiera charakterystyczne sekwencje promotorowe dla polimerazy RNA II, tj. TATA, CCAAT oraz motywy bogate w GC. Dodatkowo zidentyfikowano miejsca wiązania dla czynników transkrypcyjnych, tj. białko aktywatorowe 1 (AP-1, ang. *Activator protein 1*), Ets/PU.1, białko alfa wiążące się do motywu CCAAT (C/EBP α , ang. *CCAAT/enhancer-binding protein alfa*), czynnik jądrowy 1 (NF1, ang. *Nuclear factor I*), czynnik odpowiedzi na surowicę

(SRF, ang. *Serum response factor*), białko p53 oraz przekaźnik sygnałów i aktywator transkrypcji 3 (STAT3, ang. *Signal transducer and activator of transcription 3*) (107). Potencjalne miejsca wiązania dla tak wielu czynników transkrypcyjnych, zaangażowanych w szlaki sygnałowe różnego typu stymulacji po raz kolejny dowodzą istotności miR-21 i jego zaangażowania w odpowiedzi na zmieniające się warunki w komórce.

Transkrypcyjna kontrola poziomu miR-21 wydaje się być swego rodzaju fenomenem. Dla większości badanych miRNA, których poziom ulega istotnej zmianie w nowotworach, nie zaobserwowano zmian w ilościach ich prekursorów. Taka obserwacja wskazuje na to, że w zmienny poziom miRNA w przypadku nowotworów, zaangażowany musi być mechanizm potranskrypcyjny. W opisanym analizie, w przypadku miR-21 odnotowano silną korelację między wzrostem poziomu transkryptu *pri-miR-21*, a ilością dojrzałej cząsteczki miRNA (108).

Badanie z 2007 roku wykazało z kolei, że poziom miR-21 może być także uzależniony od hipometylacji DNA, wywołanej np. w nowotworowych komórkach jajnika (OVCAR3) (109).

Coraz więcej dowodów wskazuje na to, że poziom miR-21 kontrolowany jest nie tylko na poziomie transkrypcyjnym, ale również potranskrypcyjnym. Dla przykładu, w komórkach mięśni gładkich, białka morfogenetyczne kości (BMP, ang. *Bone morphogenic proteins*) wspólnie z transformującym czynnikiem wzrostu beta (TGF- β , ang. *Transforming growth factor beta*) przyczyniają się do zwiększonego poziomu miR-21 poprzez rekrutację RNA helikazy p68 (DDX5) do *pri-mir-21*, która stabilizuje działanie białka DROSHA (110).

1.3.2.2. Znaczenie miR-21 podczas rozwoju embrionalnego

Tuż po zapłodnieniu, komórka embrionalna z powodu braku odpowiednich czynników transkrypcyjnych, zależna jest od RNA pochodzącego od matki. W momencie, gdy genom komórki embrionalnej jest w pełni aktywny, RNA matczyne ulega stopniowej degradacji. Dokładnie w tym momencie zaangażowanie cząsteczek miRNA staje się dominujące (111). W badaniu przeprowadzonym na modelu zwierzęcym Danio pręgowany (ang. *Zebrafish*), okazało się że miR-21 był obserwowany już we wczesnym stadium rozwoju embrionalnego. Badacze zaobserwowali, że w 12-tej godzinie od zapłodnienia poziom tego miRNA stanowił 40% wszystkich badanych cząsteczek miRNA (112). W innym doświadczeniu, pracując na modelu pstrąga tęczowego (ang. *Rainbow trout*), naukowcy zaobserwowali istotny przyrost ilości miR-21, jak również zwiększoną ekspresję genu *Stat3* (czynnika transkrypcyjnego

zaangażowanego w transkrypcję *pri-miR-21*), podczas rozwoju embrionalnego (113). Podobnie jak w analizie przeprowadzonej na Danio pręgowanym, również i w tym przypadku wysoki poziom miRNA odpowiedzialny był za degradację matczynego mRNA.

Represor neuronalny REST (wyciszający czynnik transkrypcyjny RE1, ang. RE1-silencing transcription factor) ulega wysokiej ekspresji w mysich komórkach macierzystych. Heterozygotyczna delecja tego czynnika, bądź jego wyciszenie przy pomocy siRNA prowadzi do zaburzeń w procesie samoodnowy tych komórek. Profilowanie poziomu miRNA wykazało, że białko Res wpływa negatywnie na poziom wielu miRNA, w tym miR-21. W takim przypadku, brak miR-21 wpływa na zaburzenia poziomu ekspresji czynników kluczowych podczas samoodnowy komórek macierzystych, którymi są *Oct4*, *Nanog* oraz *Sox2*. Nadekspresja miR-21 obniża możliwości samoodnowy komórek niemalże o 60% (114). Badanie te pokazują, że w prawidłowym funkcjonowaniu embrionalnych komórek macierzystych, istotną rolę odgrywa prawidłowy poziom miR-21.

Morfogeneza rozgałęzień (ang. *branching morphogenesis*) stanowi podstawowy proces w rozwoju organów, tj. gruczołów zewnątrzwydzielniczych, płuc czy nerek. Badanie mysich embrionalnych gruczołów podżuchwowych stanowi podstawę w badaniu rozwoju organów, w tym również morfogenezy rozgałęzień. Wykazano, że degradacja macierzy zewnątrzkomórkowej wywołana przez metaloproteinazy macierzy (MMP, ang. *matrix metalloproteinases*) prowadzi do wzmocnienia morfogenezy rozgałęzień, ponieważ macierz zewnątrzkomórkowa w mezenchymie ulega przebudowie podczas rozwoju ślinianki podżuchwowej. Białko RECK stanowi inhibitor dla MMP i jego inhibicja prowadzi do przyspieszenia zajścia morfogenezy. W 2011 roku Hayashi i wsp. wykazali, że miR-21 poprzez oddziaływanie na MMP wpływa pośrednio na proces morfogenezy rozgałęzień (115).

1.3.2.3. miR-21 w kancerogenezie

Naukowcy na całym świecie zaczęli uważniej przyglądać się udziałowi miRNA w chorobach nowotworowych, w momencie odkrycia globalnych zmian w poziomie ekspresji wielu reprezentantów tej grupy w materiale pochodzącym z guzów od pacjentów onkologicznych (116). Biorąc po uwagę, jak znaczące są zmiany obserwowane w poziomie przeważającej liczby miRNA w tkankach nowotworowo zmienionych, uznano, że muszą być one zaangażowane w procesy karcynogenezy. W przeprowadzonym wielkoskalowym badaniu, w którym uczestniczyło 540 badanych okazało się, że miR-21 jako jedyny ma podwyższony

poziom w większości badanych typów nowotworów takich jak, rak płuca, sutka, prostaty, jelita grubego czy trzustki. W kolejnych badaniach miR-21 uzyskał już miano onkogenego miRNA (ang. *oncomiR*) co potwierdził podwyższony poziom tej cząsteczki w przeważającej liczbie nowotworów (97–99).

Badania funkcjonalne wykazały, że rola miR-21 w genezie chorób nowotworowych opiera się na jego bezpośrednim zaangażowaniu w procesy proliferacji komórek, zahamowanie procesów apoptotycznych, zwiększenie potencjału inwazyjności komórek i tworzenia przerzutów (117–120) (**Tabela 1**).

Typ nowotworu	Poziom miR-21	Rola miR-21	Geny regulowane przez miR-21	Referencje
Rak żołądka	Podwyższony	Promuje podziały komórek oraz możliwość ich przerzutów	<i>PTEN, PDCD4, RECK</i>	(121, 122)
Rak jelita grubego	Podwyższony	Formowanie się guza, metastaza	<i>SMAD7, SMAD6, ITG64, PDCD4</i>	(97, 123)
Glejak	Podwyższony	Wzrost komórek, przyspieszone podziały komórkowe, różnicowanie komórek macierzystych w nowotworowe	<i>PTEN, RECK, FasL, PDCD4, Bcl2</i>	(124)
Rak piersi	Podwyższony	Zwiększona przeżywalność komórek, zmniejszone procesy apoptotyczne, oporność na leczenie	<i>PTEN, PDCD4, TPM1, RTN4, LZTFL1</i>	(125, 126)
Rak płuc	Podwyższony	Zmniejszone procesy apoptotyczne, przyspieszone podziały komórkowe, blokowanie aktywacji czynnika NF-κB	<i>PTEN, PDCD4, Smad7, Bcl2, EGFR, Cas8, TGFβ</i>	(127, 128)
Rak prostaty	Podwyższony	Zwiększona inwazyjność, przerzuty, zmniejszone procesy apoptotyczne, niezahamowany wzrost	<i>RECK, MARCKS, PDCD4, TPM1, PTEN</i>	(129)

Tabela 1 Znaczenie miR-21 w różnych typach nowotworów

2. Cel pracy

Cząsteczka regulatorowa miR-21 jest jedną z najpowszechniej występujących form miRNA w komórkach zwierzęcych. Odgrywa ważną rolę w rozwoju embrionalnym oraz uczestniczy w utrzymaniu homeostazy komórek w dojrzałym organizmie. Jej znacząco podwyższony poziom obserwowany jest w większości nowotworów. Mimo dosyć dobrze poznanej funkcji jaką pełni miR-21, nadal nie wiadomo jakie procesy molekularne mogą być bezpośrednio zaangażowane w regulację poziomu tego miRNA.

Zasadniczym celem niniejszej pracy było **określenie zależności pomiędzy potranskrypcyjnymi mechanizmami biorącymi udział w dojrzewaniu pierwotnych transkryptów *pri-miR-21*, takich jak alternatywny splicing (AS) oraz alternatywna poliadenylacja (APA), a efektywnością biogenezy miR-21**. Te potranskrypcyjne procesy są nieodłącznymi elementami dojrzewania niemal wszystkich prekursorowych RNA, w tym prekursorów miRNA. Potranskrypcyjne mechanizmy obróbki *pri-miR-21* mogą mieć istotne znaczenie w fizjologicznej regulacji poziomu miR-21 oraz w patogenezie wielu chorób.

Sekwencja kodująca miR-21 zlokalizowana jest w regionie 3'UTR genu wakuolarnego białka membranowego 1 (*VMP1*). W literaturze opisanych jest wiele zależności pomiędzy poziomem ekspresji genu-gospodarza, a biogenezą zakodowanych w nim miRNA. Dlatego **pierwszym celem pracy była ocena zależności transkrypcji genu *VMP1*, a biogenezą miR-21**.

Aktywność alternatywnych miejsc poliadenylacji (APAS) genu *VMP1* oraz *pri-miR-21* może mieć istotne znaczenie dla biogenezy miR-21. Dlatego też **drugim celem pracy doktorskiej było określenie znaczenia APAS w transkryptach powstających z *locus VMP1* w kształtowaniu ilości dojrzałej cząsteczki miR-21**.

Prekursor *pri-miR-21*, obejmuje obszar kilku eksonów *VMP1*, w związku z czym może podlegać procesowi składania RNA podobnie jak pre-mRNA jego genu-gospodarza. Uwzględniając istnienie zależności pomiędzy procesami splicingu oraz poliadenylacji, **postanowiono również zbadać potencjalny wpływ alternatywnego splicingu *pri-miR-21* na wybór APAS i tym samym wydajność biogenezy miR-21**.

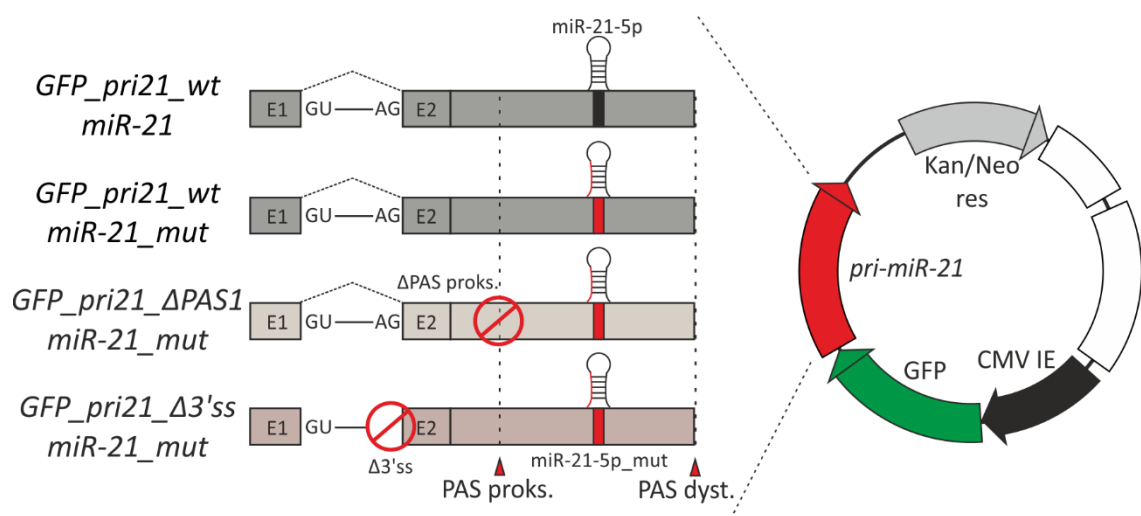
3. Materiały i metody

3.1. Konstrukty genetyczne *GFP_pri-miR-21*

Jako matrycy do uzyskania konstruktów *GFP_pri-miR-21* zastosowano szkielet wektora pEGFP-C1 (Addgene; 6084-1). Na potrzeby badań wektor ten najpierw został odpowiednio zmodyfikowany poprzez usunięcie miejsca SV40pA (miejsce sygnałowe dla terminacji transkrypcji). Do uzyskania wszystkich konstruktów, jak również do wprowadzenia zaplanowanych mutacji korzystano z zestawu NEBUiLder HiFi DNA Assembly Cloning Kit (*New England Biolabs*; E5520), który bazuje na metodzie klonowania Gibsona (ang. *Gibson assembly*). Metoda ta jest bardzo wydajną i efektywną techniką klonowania. Opiera się na aktywności egzonukleazy, która poprzez swoje działanie tworzy, tzw. "Lepkie końce" od 5' strony dwuniciowej części DNA, a następnie zawarta w mieszaninie ligaza łączy nowo powstałe lepkie końce. Podstawą prawidłowego działania tej metody jest takie zaprojektowanie wstawianej części, aby posiadała ona przynajmniej 15-20 komplementarnych nukleotydów do wcześniej zlinearyzowanego wektora. Linearyzacja wektora może przebiegać w różny sposób, tj. cięcie enzymami restrykcyjnymi bądź reakcja łańcuchowa polimerazy (PCR, ang. *polymerase chain reaction*) z wykorzystaniem odpowiednio wydajnego enzymu polimerazy (ang. *High fidelity polymerase*). Do linearyzacji wektora pEGFP-C1 używano CloneAmp HiFi PCR Premix (*Takara Bio*; 639298). Dzięki działaniu odpowiednio zaprojektowanym starterom możliwe było usunięcie danych fragment z wektora. Razem z usunięciem sekwencji SV40pA, do wektora wprowadzona została sekwencję *pri-miR-21* (4469 pz). Sekwencja została namnożona z genomowego DNA z komórek HeLa, z wykorzystaniem wysokowydajnej polimerazy Q5® High-Fidelity DNA Polymerase (*New England Biolabs*; M0491).

Lp.	Nazwa	Wielkość konstruktów (pz)
1	<i>GFP_pri-miR-21_wt</i>	8885
2	<i>GFP_pri-miR-21_wt_miR-21_mut</i>	8885
3	<i>GFP_pri-miR-21_Δ3ss_miR-21_mut</i>	8874
4	<i>GFP_pri-miR-21_ΔPAS proks.1_miR-21_mut</i>	8881

Tabela 2 Spis konstruktów używanych w pracy doktorskiej.



Rycina 9 Schematyczne przedstawienie konstruktów genetycznych *GFP_pri-miR-21* stosowanych w pracy doktorskiej.

Starter *forward* do namnożenia *pri-miR-21*

5'-ACGAGCTGTACAAGTGACAAGTCAGAGAGAGGGCG-3'

Starter *reverse* do namnożenia *pri-miR-21*

5'-TCAGTTATCTAGATCGACTAGAGGCTGACTTAGAAATTT-3'

Uzyskany wektor *GFP_pri-miR-21_wt* dodatkowo został zmodyfikowany w taki sposób, aby móc być w stanie zliczyć poziom egzogennej cząsteczki miR-21, przy zachowaniu niezmienionej struktury drugorzędowej RNA. W tym celu wprowadzono 3 mutacje punktowe w miejscu kodującym dojrzałą cząsteczkę miR-21-5p. Mutacje te uzyskano przy użyciu starterów:

Starter *forward* do wprowadzenia mutacji miR-21-5p w *GFP_pri-miR-21*

5'-ATCGGACTGCTGTTGGCTGTTGAATCTCATGGCAACACCA-3'

Starter *reverse* do wprowadzenia mutacji miR-21-5p w *GFP_pri-miR-21*

5'-CAACAGCAGTCCGATAAGCTACCCGACAAGGTGGTAC-3'

W celu uzyskania plazmidu z usuniętym miejscem poliadenylacji (AATAA), który znajduje się 198 pz przed sekwencją kodującą miR-21 (PAS proks.) ponownie zastosowano zestaw NEBuilder HiFi DNA Assembly Cloning Kit (*New England Biolabs; E5520*) z zaprojektowanymi starterami w taki sposób, aby usunąć dany fragment wektora. Reakcje przeprowadziłem na bazie wektora *GFP_pri-miR-21_miR-21_mut* ze zmienioną sekwencją kodującą miR-21-5p.

Starter *forward* do delecji fragmentu PAS proks. z wektora *GFP_pri-miR-21_miR-21_mut*

5'-GGGAGGAACAATTTTACTTTTTTCCTTTAGGAG-3'

Starter *reverse* do delecji fragmentu PAS proks z wektora *GFP_pri-miR-21_miR-21_mut*

5'-AAAATTGTTCTCCCAAGCAAAAACA-3'

W celu uzyskania plazmidu z usuniętym miejscem 3' splicingu (3'ss) (CTACAGGGAGA), ponownie użyto zestawu NEBuilder HiFi DNA Assembly Cloning Kit (*New England Biolabs; E5520*) z zaprojektowanymi starterami w taki sposób, aby usunąć dany fragment wektora. Reakcje przeprowadzono na bazie wektora *GFP_pri-miR-21_miR-21_mut* ze zmienioną sekwencją kodującą miR-21-5p.

Starter *forward* do przeprowadzenia delecji fragmentu 3' splicingowego z wektora *GFP_pri-miR-21_miR-21_mut*

5'-ATTTATTTAAACTGGTTGTCCTGGATGTTTGA-3'

Starter *reverse* do przeprowadzenia delecji fragmentu 3' splicingowego z wektora *GFP_pri-miR-21_miR-21_mut*

5'-CCAGTTTAAATAAATGCAGAACAATGCAATGA- 3'

Produkty PCR wizualizowane były w żelu agarozowym z bromkiem etydyny (Sigma; E1510). Następnie wycięte kawałki żelu agarozowego z produktem PCR oczyszczone zostały na kolumnach membranowych Gel-Out Concentrator (A&A Biotechnology;023-50). Do reakcji ligacji użyto 100ng, zlinearyzowanego plazmidu, który ligowany był z odpowiednią wstawką, w molarnym stosunku 1:2 (plazmid:wstawka). Proces ligacji trwał przez 45 minut, w temperaturze 50°C. Po tym czasie transformowałem komórki kompetentne *E. Coli DH5α* stosując protokół szoku cieplnego.

3.2. Hodowle bakteryjne

Rozmrożone na lodzie komórki kompetentne *E. Coli DH5α* zmieszano z połową wcześniej przygotowanej mieszaniny ligacyjnej. Następnie całość inkubowano na lodzie przez 30 min. Po tym czasie komórki zostały umieszczone w termobloku ustawionym na 42°C na 2 min, a następnie od razu przełożone komórki do lodu na kolejne dwie minuty. Po inkubacji na lodzie do komórek dodano podgrzaną do 37°C pożywkę LB (1% *Bacto Tryptone*, 0,5% *Yeast Extract*, 1% NaCl) i inkubowano całość w 37°C przez 2 h. Po inkubacji bakterie wysiano na szalki z LB z dodatkiem agaru oraz obecnością odpowiedniego antybiotyku (ampicylina 100 mg/ml; kanamycyna 50 mg/ml). Po całonocnej inkubacji w cieplarni nastawionej na temperaturę 37°C, wyselekcjonowano pojedynczą kolonię bakteryjną i przy użyciu PCR kolonijnego wstępnie weryfikowano jej zawartość (ang. *colony PCR*). Kolonie z pozytywnym wynikiem na zawartość konstruktów z insertem umieszczano w płynnej pożywce LB z dodatkiem antybiotyku i inkubowano kolejną noc w 37°C.

3.3. Eksperymenty na liniach komórkowych

3.3.1. Spis stosowanych linii komórkowych

Nazwa	Krótki opis	Organizm	Numer ATCC
HeLa	Rak szyjki macicy	<i>Homo sapiens</i>	(ATCC® CRM-CCL-2)
NIH3T3	Embrionalne komórki skóry	<i>Mus musculus</i>	(ATCC® CRL-1658)
HSKM	Pierwotne komórki mięśniowe	<i>Homo sapiens</i>	(ATCC® PCS-950-010)
MCF-12A	Komórki niezmienione gruczołu piersiowego	<i>Homo sapiens</i>	(ATCC® CRL-10782)
MCF-7	Komórki nowotworowe gruczołu piersiowego	<i>Homo sapiens</i>	(ATCC® HTB-22)

Tabela 3 Linie komórkowe używane w pracy doktorskiej

3.3.2. Hodowla oraz różnicowanie komórek

Do hodowli komórek HeLa, NIH3T3, Cos7, A549 używane było pełne medium hodowlane DMEM z L-glutaminą i 4,5g/l glukozą (ang. *Dulbecco's Modified Eagle's Medium*) z dodatkiem 10% bydlęcej surowicy płodowej (FBS, ang. Fetal Bovine Serum) oraz 1% antybiotykiem-antymykotykiem (Sigma). Komórki MCF-7 oraz MCF-12A hodowane były w pożywce EMEM (ang. *Eagle's Minimal Essential Medium*; Lonza) suplementowanej do 10% FBS, 1% roztworem modyfikowanej ludzkiej insuliny (*Insulin, human recombinant*), 1% antybiotykiem-antymykotykiem. Komórki HSkM hodowane były w pożywce HAM-F10, suplementowanej do 20% FBS, 1% antybiotykiem-antymykotykiem, 10 ng/ml EGF (ang. *Epidermal Growth Factor*), 0,4 µg/ml deksametazonem (ang. *Dexamethasone*).

Różnicowanie komórek HSkM wykonane zostało poprzez zmianę pożywki na różnicującą - DMEM z 4,5 g/l glukozy i suplementowanym do 2% końską surowicą, 1% antybiotykiem-antymykotykiem. Wszystkie hodowle komórkowe prowadzone były w inkubatorze, w standardowych warunkach (37°C i 5% CO₂).

3.3.3. Transfekcje komórek

Do transfekcji komórek siRNA wykorzystano Lipofektamine RNAiMAX (ThermoFisher; 13778100). Do transfekcji konstruktami genetycznymi oraz antysensownymi oligonukleotydami (AON) wykorzystano Lipofektamine 3000 (ThermoFisher; L3000150). Transfekcje wykonywane były zgodnie z protokołami producentów. Informacje dotyczące stężeń siRNA/AON oraz czasu traktowania komórek zamieszczone zostały w podpisach do odpowiednich rycin. W przypadku eksperymentów z konstruktami genetycznymi, do transfekcji używano ilości plazmidu wynoszącego 1-2 µg.

3.3.4. Spis stosowanych oligonukleotydów

Oznaczenia stosowane w zapisie mieszanych oligonukleotydów.

d DNA

r RNA

m 2'-O-metylo-RNA (2'OMe-RNA)

* szkielet tiofosforanowy (PS, ang. phosphorothioate)

P fosforan

Antysensowne oligonukleotydy (AON)

Nazwa	Sekwencja
<u>CTRL AON</u>	5'-mC*mC*mA*mG*mC*mA*mA*mG*mC*mA*mC*mA*mG*mG*mC *mA*mG*mC*mG*mG*mG-3' (Eurogentec)
<u>PAS proks. AON</u>	5'-mU*mU*mG*mU*mU*mU*mA*mU*mU*mU*mU*mC*mC* mU*mC*mC*mC*mA*mA*mG*mC-3' (Eurogentec)
<u>3'ss AON</u>	5'-mU*mG*mU*mA*mG*mA*mA*mA*mU*mA*mA*mA*mU* mG*mC*mA*mG*mA*mA*mC*mA-3' (Eurogentec)

Tabela 4 Antysensowne oligonukleotydy (AON) stosowane w pracy doktorskiej

Duplekсы siRNA

Nazwa	Sekwencja
<u>siCTRL S</u>	5'-P-rUrArArGrGrCrUrArUrGrArArGrArGrArUrArCdTdT-3' (Future Synthesis)
<u>siCTRL AS</u>	5'-P-rGrUrArUrCrUrCrUrUrCrArUrArGrCrCrUrUrAdTdT-3' (Future Synthesis)
<u>siDROSHA S</u>	5'-P-rCrGrArGrUrArGrGrCrUrUrCrGrUrGrArCrUrUdTdT-3' (Future Synthesis)
<u>siDROSHA AS</u>	5'-P- rArArGrUrCrArCrGrArArGrCrCrUrArCrUrCrGdTdT-3' (Future Synthesis)
<u>siCPSF3</u>	SMART Pool ON-TARGETplus siRNA, Human CPSF3 (51692) (Dharmacon nr kat. L-006365-01-0005) (GTGTCATTATAGCGGGAT, GTTCAATCCTCGGAATA, CGATGAATCCCTCCGAGAA, TGAGAAACGTGAAGAGCGA)

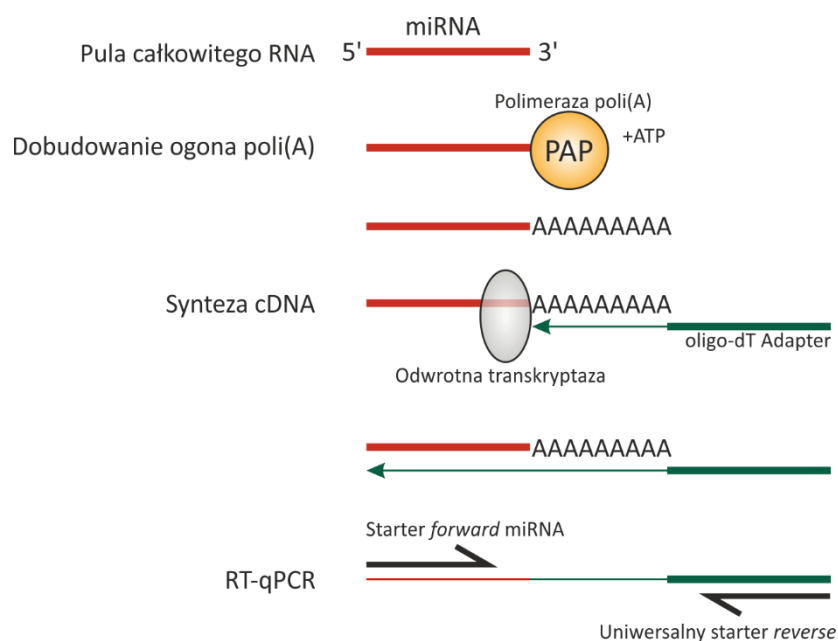
Tabela 5 Duplekсы siRNA stosowane w pracy doktorskiej

3.3.5. Izolacja RNA i odwrotna transkrypcja

Całkowity RNA z komórek izolowany był przy użyciu TRIzol™ Reagent (ThermoFisher; 15596018) oraz oczyszczany na minikolumnach (AABiotech; 043-100). Dodatkowo w trakcie oczyszczania na kolumnach, RNA było trawione DNazą. Do reakcji odwrotnej transkrypcji używane było 500-1000 ng całkowitego RNA, z wykorzystaniem zestawu TranScriba Kit (AABiotech; 4000-100). W reakcji odwrotnej transkrypcji używano starterów losowych heksametrów lub startera oligo-dT.

3.3.6. Odwrotna transkrypcja dla cząsteczek miRNA

Do uzyskania materiału, niezbędnego do analiz poziomu ilości miRNA, zastosowano zestaw qScript microRNA cDNA Synthesis Kit (Quantabio; 95107-100). W celu utrzymania cDNA z krótkich fragmentów RNA, w zestawie wykorzystano metodę dołączenia uniwersalnego adaptera na 3' końcu każdej cząsteczki miRNA, po uprzednim dobudowaniu ogona poli(A), przez działanie PAP. Do reakcji używano 1000ng całkowitego RNA, zawierającego również pulę krótkich RNA. Dzięki dobudowanemu adapterowi na końcu każdej z krótkich cząsteczek RNA, możliwa była analiza wielu różnych miRNA, z tej samej puli matrycy (**Ryc. 10**).



Rycina 10 Schemat procedury odwrotnej transkrypcji otrzymania cDNA z krótkich RNA jako matrycy

3.4. Analiza danych z portalu GTEX

Dane RNA-Seq użyte do analizy poziomów ekspresji genu VMP1 i jego transkryptów w tkankach ludzkich zostały pobrane z portalu The Genotype-Tissue Expression (GTEX) Project (wersja bazy danych phs000424.v8.p2 z 01.06.2020) (The Genotype-Tissue Expression (GTEX) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.). Dane RNA-Seq zostały pobrane w formie znormalizowanej w postaci transkryptów na milion (TPM, ang. *Transcript per Milion*).

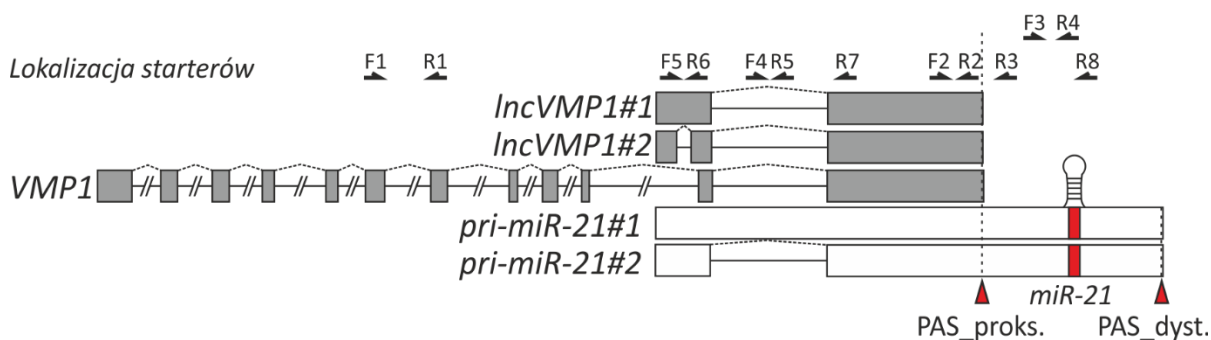
TPM oblicza się w następujący sposób:

1. Surowe wartości odczytów zostają podzielone przez długość transkryptu (wyrażoną w tysiącach par zasad (kbp)). Otrzymujemy odczyty na tysiąc par zasad (RPK, ang. *Reads per Kilobase*).
2. Wartości RPK w jednej próbce zostają zsumowane i podzielone przez 1,000,000. Otrzymujemy czynnik skalujący „na milion”.
3. Wartości RPK zostają podzielone przez czynnik skalujący „na milion”. Otrzymujemy TPM
4. Poziom ekspresji genu otrzymuje się przez dodanie wszystkich wartości TPM transkryptów pochodzących z jednego genu

3.5. Analizy RT-PCR oraz RT-qPCR

Wszystkie RT-PCR były przeprowadzane z użyciem GoTaq Flexi DNA Polymerase według zalecanej procedury (Promega; M3001). Produkty PCR rozdzielane były na żelu agarozowym (1-1,5%) z bromkiem etydy (Sigma; E1510). Do wizualizacji korzystano z systemu G:BOX (Syngene), a do analizy poziomu intensywności sygnału używano oprogramowania GeneTools (Syngene).

W przypadku analiz RT-qPCR, do reakcji używano Maxima SYBR Green/ROX qPCR Master Mix (ThermoFisher; K0223). Reakcje oraz analizy były przeprowadzane w aparacie QuantStudio™ 7 Flex Real-Time PCR System (ThermoFisher; 4485701).



Rycina 11 Schemat prezentujący lokalizacje starterów do badania poziomu zmian *VMP1/pri-miR-21*.

3.5.1. Lista starterów

Nazwa	Symbol	Sekwencja 5' -> 3'
Analizy zmian <i>VMP1/pri-miR-21</i> (Ryc. 11)		
<i>mRNA_VMP1_F</i>	F1	TTTCCCGAACCACCCTATCC
<i>mRNA_VMP1_R</i>	R1	CAGACTCTGCATGTTCCAGC
<i>VMP1_PASproks._F</i>	F2	TGAAACTGGGCTGCTGCATA
<i>VMP1_PASproks._R1</i>	R2	ACACTTACAGAACGGCAAGA
<i>VMP1_PASproks._R2</i>	R3	GCAGACAGTCAGGCAGGATT
<i>pri-miR-21_F</i>	F3	GGGTTTCGATCTTAACAGGCC
<i>pri-miR-21_R</i>	R4	AATGTCAGACAGCCCATCGA
<i>VMP1/pri-miR-21_intron_F</i>	F4	CTCGAGGGAAGCACAGGTTT
<i>VMP1/pri-miR-21_intron_R</i>	R5	CGTGCCTGGAATCACAGCTA
<i>lncVMP1/pri-miR-21_F</i>	F5	TATCCAATCCCTGCCTTTTG
<i>lncVMP1/pri-miR-21_R</i>	R6	AATGGCTTCTGCAGAGATGG
<i>pri-miR-21_ret#1_intron_F</i>	F5	TATCCAATCCCTGCCTTTTG
<i>pri-miR-21_ret#1_intron_R</i>	R7	TTAAGGCAGAACCCATCCAC
<i>pri-miR-21_ret#2_intron_F</i>	F5	TATCCAATCCCTGCCTTTTG
<i>pri-miR-21_ret#2_intron_R</i>	R8	AAGGTGGTACAGCCATGGAG
Analizy poziomu ekspresji genów regulowanych przez <i>miR-21</i>		
<i>TGFBR2_F</i>	TGFBR2_F	CAAAGGTCCCATTTGCAGTT
<i>TGFBR2_F</i>	TGFBR2_F	GAGACCTTCCACCATCCAAA
<i>PTEN_F</i>	PTEN_F	ACCAGGACCAGAGGAAACCT

<i>PTEN_R</i>	PTEN_R	GCTAGCCTCTGGATTTGACG
<i>BTG2_F</i>	BTG2_F	AGTCTCCAGCTAGGCCCTTC
<i>BTG2_R</i>	BTG2_R	CACGCAACAGTTCTCTCCAA
<i>ICAM1_F</i>	ICAM1_F	AGTCCAGTACACGGTGAGGA
<i>ICAM1_R</i>	ICAM1_R	TGTGACCAGCCCAAGTTGTT
<i>NFIB_F</i>	NFIB_F	CACCAAGCAGCAAAAGACCC
<i>NFIB_R</i>	NFIB_R	CAGGTATTCCGGGATGGTGG
Poziom pri-miRNA		
<i>pri-miR-23b_F</i>	pri-miR-23b_F	GAAGCCCAGTGTGTGCAGA
<i>pri-miR-23b_R</i>	pri-miR-23b_R	GACCACGGTTTCTGGAGGAG
<i>pri-miR-27b_F</i>	pri-miR-27b_F	CTGAGGAAGATGCTCACCGG
<i>pri-miR-27b_R</i>	pri-miR-27b_R	CCAGCGGCTCCAACCTAACT
<i>pri-miR-24-1_F</i>	pri-miR-24-1_F	CTGAGGCGCTGCTTCTCC
<i>pri-miR-24-1_R</i>	pri-miR-24-1_R	CCTCGGGCACTTACAGACAC
<i>pri-miR-30d_F</i>	pri-miR-30d_F	TGTCTGTGAATAGCCGGTAGC
<i>pri-miR-30d_R</i>	pri-miR-30d_R	AGCTGAAGATGATGACTGGCA
Badanie poziomu miRNA		
<i>miR-21-5p</i>	miR-21-5p	TAGCTTATCAGACTGA
<i>miR-21-5p_mut</i>	miR-21-5p_mut	TAGCTTATCGGACTGC
<i>miR-16</i>	miR-16	GGATGACACGCAAATTCGTG
<i>miR-378a-3p</i>	miR-378a-3p	ACTGGACTTGGAGTCAGAAG
<i>miR-23b-3p</i>	miR-23b-3p	ATCACATTGCCAGGGATTACCAC
<i>miR-27b</i>	miR-27b	TTCACAGTGGCTAAGTTCTGC
<i>U6</i>	U6	GGATGACACGCAAATTCGTG
<i>Uniwersalny starter_R</i>	Universal_R	GCATAGACCTGAATGGCGGTA
Pozostałe startery		
<i>DROSHA_F</i>	DROSHA_F	GTACCCTCCGGGCTATTCTC
<i>DROSHA_R</i>	DROSHA_R	ACTGTGATCTCGGTGCCTGT
<i>GAPDH_F</i>	GAPDH_F	CACATCGCTCAGACACCATG
<i>GAPDH_R</i>	GAPDH_R	CTTGACGGTGCCATGGAATT
<i>MEF2C_F</i>	MEF2C_F	ATATGCAAGCAAATCTCCTCC

<i>MEF2C_R</i>	MEF2C_R	GTTGCTACGGAAAACCACTGGGGTA
<i>MYOG_F</i>	MYOG_F	AAGAGAAGCACCCCTGCTCAA
<i>MYOG_R</i>	MYOG_R	AGGTTGTGGGCATCTGTAGG
<i>CPSF3_F</i>	CPSF3_F	ACGACATGCTGTATACCGAGA
<i>CPSF3_R</i>	CPSF3_R	TTCAGCTGCCATTAAGTGCC
<i>GAPDH_PAS_F</i>	GAPDH_F1	CTGAATCTCCCCTCCTCACA
<i>GAPDH_PAS_R</i>	GAPDH_R1	TGCCCCAGACCCTAGAATAA
<i>GAPDH_PAS1kpz_F</i>	GAPDH_F2	TCCAGCCTAGGCAACAGAGT
<i>GAPDH_PAS1kpz_R</i>	GAPDH_R2	TGTGCACTTTGGTGTCACTG
<i>GFP_F</i>	GFP_F1	AGGACGACGGCAACTACAAG
<i>GFP_R</i>	GFP_F2	TTCTGCTTGTCGGCCATGAT

Tabela 6 Wykorzystywane w reakcjach RT-PCR oraz RT-qPCR startery

3.6. Western blot

Komórki były lizowane w buforze RIPA zawierającym 150 mM NaCl/ 50 mM Tris-HCl pH 8,0/ 1 mM EDTA/ 0,5% NP-40/ 0,5% TritonX-100/ 0,5% deoksycholany sodu/ 0,1% SDS suplementowanym koktajlem inhibitorów proteaz Halt™ Protease Inhibitor Cocktail (ThermoFisher; 78439). Lizat inkubowany był na lodzie przez 1h, a następnie minimum przez całą noc w -80°C. Po inkubacji próby były wirowane 10 000 RPM przez 10min w 4°C. Po przeniesieniu supernatantu, mierzone było stężenie białka przy użyciu zestawu Pierce™ BCA Protein Assay Kit (ThermoFisher; 23227), bazującego na reakcji kolorymetrii. Przed nałożeniem próbek na żel, poddawane były one denaturacji w 70°C przez 10min. Rozdział białek odbywał się na poliakrylamidowych, gradientowych (4-12%) żelach Bolt™ 4-12%, Bis-Tris (ThermoFisher; NW04122BOX). Po rozdzielaniu próbek były poddawane mokremu elektrotransferowi (1 h, 100 V) na membranę PVDF (Amersham™ Hybond® P, 45µm, Sigma, GE10600023). Do wykonania procedury western blot wykorzystywano system SNAP i.d.® 2.0 (Merck, C73105). Membrana była blokowana w 0,125% roztworze odtłuszczonego mleka (Sigma, 1153630500) rozpuszczonego w TBST (TBS/0,1% Tween-20) przez 20min. Następnie inkubacja z I-rzędowym króliczym przeciwciałem anti-DROSHA (1:500, CST, 3410S) trwała 1h. Po tym czasie membrana była płukana 5-krotnie w TBST. Inkubacja z kozim przeciwciałem II-rzędowym anty-króliczym skoniugowanym z peroksydazą chrzanową (HRP, ang. *horseradish*

peroxidase) (Sigma, A9169) trwała 13min. Następnie membranę płukano 5-krotnie buforem TBS. Do detekcji sygnału używano Immobilon Forte Western HRP (Sigma, WBLUF0500). Do obrazowania sygnału używano G:Box Chemi-XR5 (Syngene).

3.7. Frakcjonowanie komórek

Procedura frakcjonowania komórek na materiał pochodzący z jądra komórkowego oraz cytoplazmy bazuje na opublikowanym wcześniej protokole (130).

Komórki fibroblastyczne hodowane były na szalkach o średnicy 150 mm (powierzchnia hodowlana 145 cm²). Dzięki zastosowaniu tak dużej powierzchni hodowlanej, w każdej próbie znajdowało się ok. 10 milionów komórek przeliczonych za pomocą automatycznego licznika komórek Countness (Invitrogen). Zwirowany osad komórkowy po zebraniu był trzymany na lodzie. Komórki zostały rozpuszczone przez pipetowanie w 380 µl HLB (ang. *Hypotonic Lysis Buffer*): 10 mM Tris (pH 7,5); 10 mM NaCl, 3 mM MgCl₂; 0,3% NP-40 (v/v); 10% glicerol (v/v)) suplementowanym 100 U inhibitora RNaz (RNasin, Promega), mieszaninę inkubowano przez 30min na lodzie. Następnie mieszanina była wirowana w 1000xg, 4°C przez 3 min. Supernatant, w którym znajdowała się frakcja cytoplazmatyczna, przeniesiono do nowej probówki, dodano 1 ml RPS (ang. *RNA Precipitation Solution*): 0,5 ml 3 M octanu sodu (pH 5,5) z 9,5 ml etanolu) i inkubowano w -20°C przez całą noc. Osad, który jest połowicznie oczyszczona frakcją jądrową, był przemywany pięciokrotnie 1 ml HLB i wirowany w 300xg, 4°C przez 2 min. Do oczyszczonej frakcji jądrowej dodano 500 µl TRIzol™ Reagent (ThermoFisher; 15596018), a następnie izolowano całkowity RNA przy użyciu oczyszczania RNA na kolumnkach krzemionkowych (AABiotech). Po całonocnej inkubacji supernatant z frakcją cytoplazmatyczną wytrząsano przez 30 s po czym wirowano w 18 000 g, 4°C przez 15 min. Osad był oczyszczany przez wymieszanie z 70% etanolem i wirowanie w 18 000 g, 4°C przez 5 min. Osad był pozostawiony do samoistnego osuszenia, po czym dodano do niego 500 µl TRIzol™ Reagent (ThermoFisher; 15596018) i izolowano podobnie jak frakcję jądrową.

4. Wyniki

4.1. Transkrypty *pri-miR-21* powstają niezależnie z własnych miejsc promotorowych, a terminacja transkrypcji zachodzi w alternatywnych miejscach poliadenylacji

4.1.1. Gen *VMP1* oraz jego izoformy splicingowe

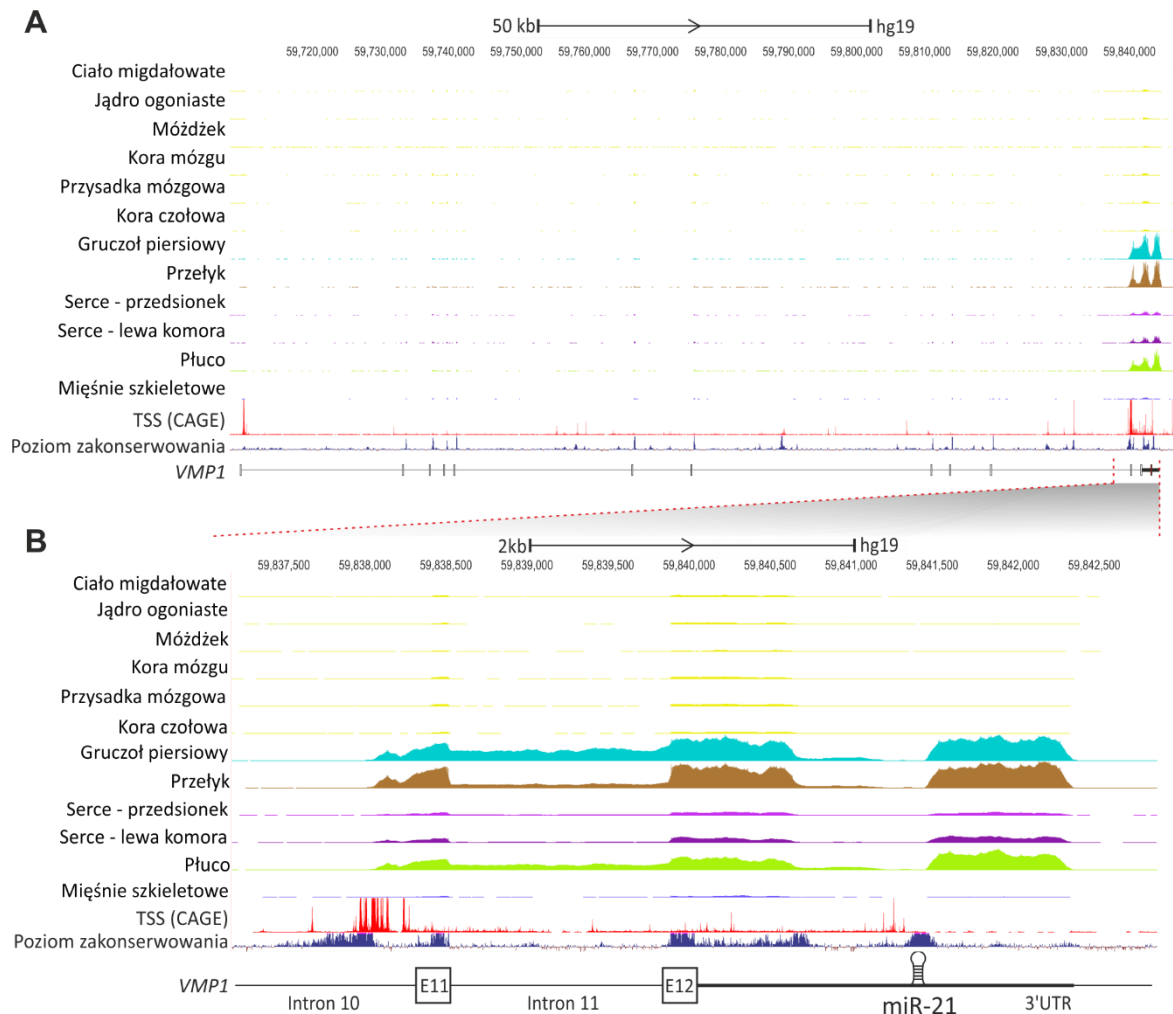
Gen kodujący białko VMP1 zlokalizowany jest na chromosomie 17 (17q23). Składa się z 12 eksonów oraz 11 intronów. Długość sekwencji mRNA VMP1 wynosi 1260 nukleotydów, co przekłada się na sekwencję aminokwasową o długości 419. W 3'UTR VMP1 zlokalizowana jest sekwencja kodująca cząsteczkę miR-21.

Rola biologiczna wakuolarnego białka membranowego 1 (VMP1, ang. Vacuole membrane protein 1) została opisana w 2002 r. jako białka uczestniczącego w rozwoju zapalenia trzustki poprzez aktywność modulacyjną w przekazywaniu sygnału przez błony oraz zaangażowanie w procesy autofagii (131). Co więcej, poziom ekspresji VMP1 może być istotnie podwyższony w sytuacji niedoboru substancji odżywczych a wzrost ten istotnie wpływa na formowanie się autofagosomów w komórkach zwierzęcych (132). VMP1 odgrywa również istotną rolę w przekazywaniu informacji dotyczących funkcji komórkowych pomiędzy retikulum endoplazmatycznym (ER, ang. endoplasmic reticulum), a innymi organellami komórkowymi takimi jak mitochondria, prekursorzy membranowe autofagii, aparat Golgiego czy endosomy. Podwyższony poziom ekspresji VMP1 obserwowany jest w nowotworowych komórkach gruczołu piersiowego i potencjalnie może służyć jako biomarker stopnia zaawansowania tej choroby (133).

Projekt GTEx (ang. The Genotype-Tissue Expression project) jest efektem współpracy międzynarodowego konsorcjum, którego założeniem było stworzenie szeroko dostępnej platformy zbierającej wyniki badań wysokoprzepustowych badań sekwencjonowania, tj. WGS (ang. whole-genome sequencing), WES (ang. whole-exome sequencing) czy sekwencjonowania RNA (ang. RNA-seq) przeprowadzonych na materiale z 54 różnych zdrowych tkanek od blisko 1000 osób (<https://www.gtexportal.org/home/>). Jedną z funkcjonalności platformy GTEx jest możliwość wizualizacji zebranych danych w przeglądarce internetowej tj. UCSC Genome Browser (ang. University of California, Santa Cruz).

Korzystając z przeglądarki genomowej UCSC Genome Browser przeanalizowano poziom ekspresji genu VMP1 (**Ryc. 12A**) z uwzględnieniem zestawu danych RNA-seq z

dwunastu różnych tkanek człowieka (ciało migdałowe, jądro ogoniaste, mózdzek, kora mózgowa, przysadka mózgowa, kora czołowa, gruczoł piersiowy, przełyk, przedsionek i lewa komora serca, płuco oraz mięśnie szkieletowe). Patrząc na dane surowe można zauważyć, że największa liczba odczytów z części tkanek (5 z 12) pochodzi z dwóch ostatnich eksonów oraz regionu nie ulegającego translacji poniżej sekwencji kodującej białko (3'UTR, ang. 3' untranslated region).



Rycina 12 Poziom ekspresji VMP1 i pri-miR-21 w ludzkich tkankach. (A) Poziom ekspresji genu VMP1 na podstawie wyników sekwencjonowania RNA (RNA-seq), pochodzący z 12 różnych tkanek ludzkich (Platforma GTEx, ang. *The Genotype-Tissue Expression*). Obecność miejsca startu transkrypcji (TSS; ang. *transcription start site*) zlokalizowanego przed pierwszym eksonem VMP1 oraz drugiego TSS pod koniec 10 intronu VMP1 sugeruje istnienie różnych izoform RNA powstających z genu VMP1. Eksperymentalnie potwierdzone miejsca TSS zostały określone na podstawie analizy struktury czapeczki na końcu 5' genów (TSS CAGE). Duża liczba odczytów znajdująca się na 3' końcu VMP1 wskazuje na silną aktywność drugiego miejsca promotorowego. Wysoki poziom zakonserwowania konkretnych regionów genu charakteryzuje eksony, potencjalne miejsca promotorowe oraz sekwencję prekursora pre-miRNA. **(B)** 25-krotne przybliżenie 3' końca VMP1 wyraźnie wskazuje na obecność TSS w 10 intronie VMP1, zlokalizowanego około 350 pz od eksonu 11 VMP1. Porównanie poziomu ekspresji pomiędzy tkankami wskazuje na dużą zmienność. Odczyty sekwencjonowania pochodzące z intronu 11 dowodzą możliwości istnienia transkryptów z retencją intronu 11. W 3'UTR, w odległości około 1100 pz, zlokalizowana jest sekwencja kodująca miR-21. Duży poziom zakonserwowania w intronie 10 VMP1 wskazuje na obecność istotnego

miejsca promotorowego. E11 – ekson 11; E12- ekson 12. Rycina została stworzona na podstawie danych zawartych w internetowej przeglądarce UCSC Genome (<https://genome-euro.ucsc.edu>) Przedstawiony na rycinie poziom ekspresji dla poszczególnych tkanek stanowi znormalizowaną wartość wyrażoną w TPM (ang. *transcripts per million*). Poziom zakonserwowania został wyliczony na podstawie przypasowania genomów pochodzących ze 100 reprezentantów kręgowców.

Dodatkowo analiza ekspresji genów opierająca się na strukturze cap (CAGE, ang. *Cap Analysis of Gene Expression*) wykazała istnienie silnego miejsca startu transkrypcji (TSS, ang. *Transcription start site*) w rejonie promotorowym genu *VMP1*, jak również w jego 10. intronie. Porównując poziom zakonserwowania sekwencji DNA u różnych zwierząt kręgowych można stwierdzić wysoki stopień podobieństwa regionów eksonowych (**Ryc. 12A**). Można również zauważyć nagromadzenie odczytów z analiz RNA-seq również poza eksonami *VMP1*, co świadczyć może o obecności dodatkowej jednostki transkrypcyjnej, której początek zlokalizowany jest kilkaset nukleotydów przed eksonem 11. genu *VMP1* (**Ryc. 12B**). W przygotowanym na Ryc. 12B zestawieniu wyników sekwencjonowania RNA-seq obrazujących tę dodatkową jednostkę transkrypcyjną uwidoczniona została jej specyficzność tkankowa. W różnych obszarach mózgu, w przedsionku serca oraz w mięśniach szkieletowych liczba odczytów jest istotnie mniejsza niż w pozostałych analizowanych tkankach. Kolejną interesującą obserwacją jest względnie duża liczba odczytów RNA-seq zmapowanych do 11. intronu *VMP1*, sugerującym powszechne występowanie izoform RNA z jego retencją.

Korzystając z przeglądarki UCSC można zlokalizować sekwencje kodujące miRNA. W przypadku genu *VMP1*, oddalona od końca 5' ostatniego eksonu o blisko 1350 pz, w regionie 3'UTR o bardzo silnym zakonserwowaniu ewolucyjnym, zlokalizowana jest sekwencja kodująca miR-21.

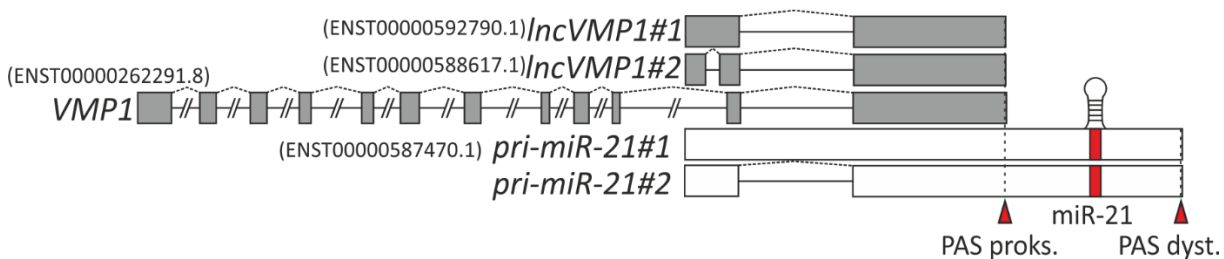
Celem szczegółowej charakterystyki *VMP1*, wykorzystano informacje zgromadzone w bazach danych Ensembl (www.ensembl.org), UCSC oraz danych RNA-seq zdeponowanych na platformie GTEx. Do tej pory określono 17 jego różnych izoform splicingowych (**Ryc. 13B**). Przy każdej izoformie umieszczony został indywidualny numer ENST (ang. *Ensembl transcript*) stanowiący podstawę identyfikacji konkretnej izoformy w bazie danych Ensembl. Analizując zestaw prezentowanych izoform należy mieć na uwadze, że duża część z przedstawionych izoform ma charakter wyłącznie predykcyjny, opierający się jedynie na bioinformatycznych algorytmach i automatycznej adnotacji transkryptów do genomu. Porównując poziom ekspresji poszczególnych izoform *VMP1* w panelu tkanek (**Ryc. 13A**), zauważyć można wyraźną dominację trzech pierwszych form transkryptów - ENST00000262291, ENST00000587470,

mózgowego. W przypadku tkanek ze śledziony, płuca, gruczołu piersiowego, nerki czy prostaty poziom ekspresji tych dwóch transkryptów istotnie przewyższa poziom ekspresji pre-mRNA *VMP1*. Według bioinformatycznej predykcji transkrypt ENST00000592790 nie podlega procesom składowania. W przypadku izoformy ENST00000587470 widać, że jej budowa przypomina transkrypt pre-mRNA, w którym występuje sekwencja intronu 11, który ulec może wycięciu na drodze splicingu.

Uwzględniając poczynione obserwacje dotyczące izoform splicingowych mRNA *VMP1* zdeponowanych w bazie Ensembl oraz odczytów RNA-seq pochodzące z platformy GTEx, transkrypt kodujący białko *VMP1* (ENST00000262291) oraz jedna z alternatywnych izoform - ENST00000592790 w swojej budowie nie zawierają sekwencji miR-21. Dodatkowo izoforma ENST00000592790 według bazy Ensembl stanowi długi, niekodujący RNA. Izoforma ENST00000587470 natomiast może stanowić sekwencję kodującą prekursor miR-21 i tym stanowić będzie obiekt badań w dalszej części rozprawy.

4.1.2. Pierwotne prekursorzy *pri-miR-21*

Pierwsze badania na temat biogenezy miR-21 wykazały, że pierwotny prekursor miR-21 (*pri-miR-21*) jest niezależnie powstającym transkryptem i jego poziom ekspresji nie zależy od poziomu swojego genu-gospodarza - *VMP1* (134). Autorom pracy udało się określić pełną długość *pri-miR-21* na około 3,4 tysięcy nukleotydów, którego początek zlokalizowany jest w intronie 11. *VMP1*. Dodatkowo autorzy wykazali, że powstający transkrypt *pri-miR-21* na swoim końcu 5' posiada strukturę czapeczki, oraz ulega procesowi poliadenylacji. Dodatkowo sugerowano, że transkrypt ten nie podlega procesom splicingu. W późniejszym czasie pojawiły się prace naukowe wskazujące, że początek transkrypcji *pri-miR-21* może mieć miejsce w intronie 10. *VMP1*, gdzie zlokalizowanych jest kilka regionów do których przyłączać się mogą takie czynniki transkrypcyjne jak AP-1, NFI, SRF, p53 oraz STAT3 (106, 107).



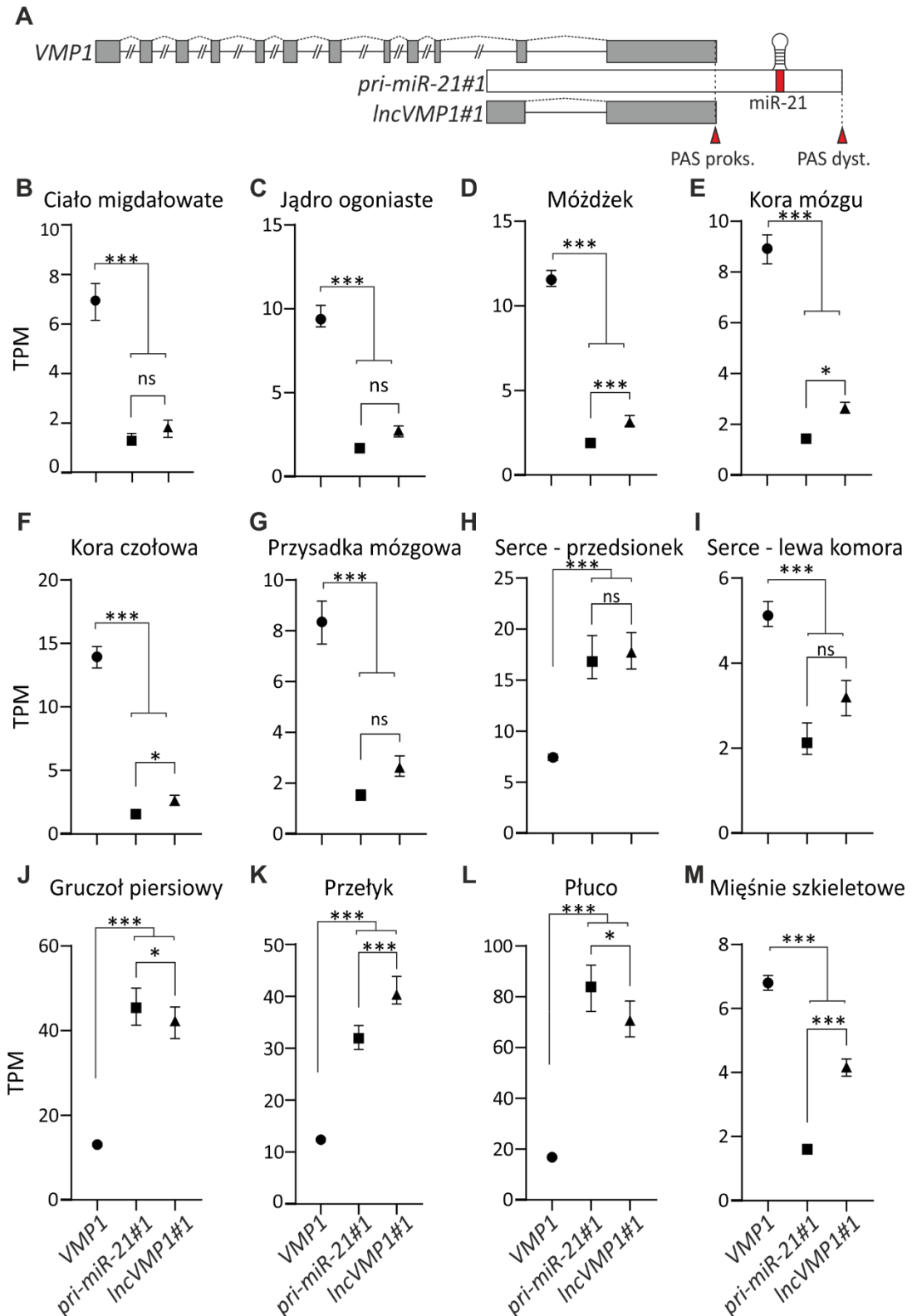
Rycina 14 Schemat izoform RNA pochodzących z *VMP1* włączając prekursorzy *pri-miR-21*. Graficzne przedstawienie najczęściej występujących izoform na podstawie wyników badania poziomu ekspresji

zdeponowanych na platformie GTEx. Czerwonymi trójkątami zaznaczone alternatywne miejsca poliadenylacji (APAS). Na 3' końcu, za proksymalnym miejscem poliadenylacji (PAS proks.), zlokalizowana jest sekwencja kodująca miR-21. Nazwą ENST w nawiasach zostały przyporządkowane jednostki transkrypcyjne, które w dalszej części pracy odwoływać będą się do odpowiedniej izoformy. Kolorem szarym zaznaczony jest transkrypt kodujący białko VMP1 oraz dwie prawdopodobnie niekodujące białka izoformy VMP1. Kolorem białym dwie izoformy *pri-miR-21*. Izoforma *pri-miR-21#1* różni się od izoformy *pri-miR-21#2* retencją intronu. Prostokąty symbolizują eksony, czarna linia stanowi introny. Przerwana linia łącząca eksony oznacza zająście splicingu.

Uwzględniając doniesienia literaturowe oraz analizy danych pochodzących z baz takich jak GTEx (**Ryc. 12 i Ryc. 13**) zaproponowano schemat reprezentujący różne formy transkryptów *pri-miR-21* na tle kodującego białko transkryptu VMP1 oraz długich, niekodujących cząsteczek RNA (*IncVMP1*), których transkrypcja kończy się przed sekwencją kodującą miR-21 i tym samym nie będących prekursorami dla tej cząsteczki miRNA (**Ryc. 14**). Zaznaczone na Ryc. 14 miejsca sygnału poliadenylacji (PAS, ang. *polyadenylation signal*) proksymalne oraz dystalne zostały umiejscowione w oparciu o dane uzyskane z bazy polyA_DB_3 (135). Uwzględniając długość analizowanych transkryptów oraz lokalizację sekwencji kodującej miR-21 zaproponowano podział izoform na prekursor miR-21 (***pri-miR-21***), zaznaczone białym kolorem, oraz izoformy które prawdopodobnie nie są zaangażowane w biogenezę miR-21 (***Inc-VMP1***), zaznaczone kolorem szarym. Opierając się na danych z RNA-seq oraz lokalizacjach PAS, miejsce cięcia i poliadenylacji transkryptu głównego VMP1 następuje w głównej mierze w proksymalnym PAS, zlokalizowanym około 200 pz od miejsca, w którym zakodowana jest sekwencja miR-21. Również opisywany w poprzednim rozdziale ulegający splicingowi transkrypt ENST00000592790 kończyć się może również na proksymalnym PAS. W dalszej części pracy będzie on nazywany *IncVMP1#1*. Natomiast izoforma ENST00000588617, różniąc się od transkryptu *IncVMP1#1* jedynie obecnością krótkiego intronu na samym początku nazywana będzie *IncVMP1#2*. Izoforma ENST00000587470 jako jedyna ze zdeponowanych w bazie GTEx kończy się na dystalnym PAS i tym samym stanowić może prekursor miR-21 (*pri-miR-21#1*). W 2018 roku A. Tscherner i wsp. (136), analizując materiał pochodzący z różnych tkanek bydłowych, opisali istnienie dodatkowej izoformy *pi-miR-21*, która od *pri-miR-21#1* różni się brakiem całego intronu. W niniejszej pracy transkrypt ten będzie określany jako *pri-miR-21#2*.

4.1.3. Poziom ekspresji *VMP1*, *IncVMP1* oraz *pri-miR-21* jest znacząco odmienny w różnych tkankach człowieka

Korzystając z dostępnych danych poziomu ekspresji genów w panelu tkankowym zdeponowanych w bazie danych GTEx, przeanalizowano poziom ekspresji trzech najpowszechniej występujących izoform *VMP1* – transkrypt *VMP1* kodujący białko, *pri-miR-21#1*, czyli pierwotny prekursor miR-21 z retencją intronu oraz *IncVMP1#1* kończącym się na PAS proks. (**Ryc. 15A**). W analizie wykorzystano dane pochodzące z dwunastu różnych tkanek. Wartości na wykresach prezentują względny poziom ekspresji bazujący na ilości odczytów sekwencjonowania przypisanych do badanego transkryptu (TPM, ang. *Transcripts Per Million*). Sześć spośród nich są to tkanki pochodzenia mózgowego (**Ryc. 15B-G**). We wszystkich tkankach mózgowych widać istotnie wyższy poziom transkryptu kodującego białko *VMP1* niż dwóch pozostałych, krótszych transkryptów. Wyniki RNA-seq dla dwóch obszarów mięśnia sercowego, czyli przedsionka (**Ryc. 15H**) oraz lewej komory serca (**Ryc. 15I**), ukazały dosyć ciekawą prawidłowość. Mianowicie poziom mRNA *VMP1* w przedsionku jest istotnie niższy niż poziom badanych niekodujących cząsteczek RNA. Natomiast w lewej komorze serca sytuacja bardzo przypomina rozkład obserwowany w tkankach mózgowych. Wyraźnie zróżnicowany poziom ekspresji w obrębie jednego narządu, jakim jest mięsień sercowy, wskazuje na obecność odmiennego systemu ściśle regulującego poziom poszczególnych transkryptów w obrębie jednego organu. Taka sytuacją może się bezpośrednio wiązać z rolą, jaką pełnić mogą powstające prekursorzy *pri-miR-21*. Analiza wyników RNA-seq dla gruczołu piersiowego ukazała znaczącą dominację ilościową prekursora *pri-miR-21* nad pełnym transkrypcem *VMP1* (**Ryc. 15J**). W materiale z przełyku, podobnie jak w przypadku gruczołu piersiowego, izoforma *pri-miR-21* oraz *IncVMP1* wykazują istotnie większy poziom ekspresji niż *VMP1* (**Ryc. 15K**).



Rycina 15 Poziom ekspresji VMP1, IncVMP1 oraz pri-miR-21 w wybranych tkankach ludzkich (A) Schemat przedstawiający analizowane transkrypty bazując na danych z portalu GTEx **(B-G)** Poziom ekspresji w tkankach pochodzenia mózgowego wyraźnie pokazuje dominację transkryptu VMP1 nad transkryptami IncVMP1 oraz pri-miR-21. **(H-I)** W materiale z różnych części mięśnia sercowego, widać wyraźną różnicę w profilu ekspresji VMP1. W przypadku tkanek gruczołu piersiowego **(J)**, przełyku **(K)** oraz płuca **(L)** istotnie zwiększona jest ilość cząsteczek RNA niekodujących białko. Profil ekspresji w mięśniach szkieletowych **(L)** wyraźnie przypomina profil z tkanek mózgowych z dominacją transkryptu VMP1. Na osiach Y wykresów przedstawiony jest znormalizowany poziom odczytów z sekwencjonowania RNA względem analizowanego genu (TPM, ang. *Transcripts Per Million*). Istotność

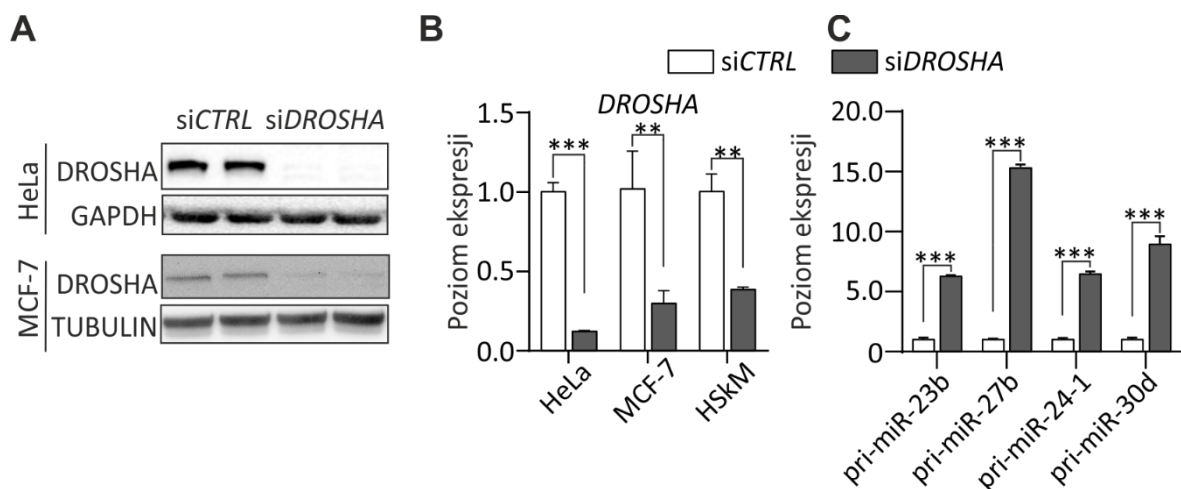
statystyczna obserwowanych różnic została określona za pomocą testu t-Studenta dla danych niesparowanych (* $P < 0,05$, *** $P < 0,001$).

Jednakże, w tym przypadku poziom ekspresji *lncVMP1* jest istotnie wyższy niż *pri-miR-21*. Podobnie jak w materiale z gruczołu piersiowego, profil ekspresji prezentuje się również i w tkance płucnej (**Ryc. 15L**). Dodatkowo należy zauważyć, że poziom izoformy *pri-miR-21* w płucach jest najwyższy wśród wszystkich analizowanych tkanek. W materiale pochodzącym z mięśni szkieletowych profil ekspresji bardzo przypomina rozkład jaki obserwowany był w przypadku tkanek mózgowych. W tym przypadku *VMP1* istotnie przewyższa poziomem ekspresji *lncVMP1* i *pri-miR-21*. Powyższe informacje stanowiły wskazówkę wyboru odpowiedniego modelu/modeli komórkowych do konkretnych eksperymentów przedstawionych w dalszej części pracy.

Analizując wyniki badania poziomu ekspresji wybranych izoform *VMP1* widać wyraźnie znaczną zmienność, również w dystrybucji poszczególnych transkryptów. Taki wynik sugerować może, że poziom ekspresji badanych izoform w różnych tkankach zależy może od pełnionych funkcji w komórkach pochodzących z różnych tkanek. Dodatkowo zmienny poziom transkryptów *lncVMP1* oraz *pri-miR-21* sugeruje, że w wybranej tkance proces transkrypcji, bądź dojrzewania danej izoformy może być różny. Pamiętać należy jednak, że analiza RNA-seq ukazuje jedynie aktualny poziom rozmaitych form RNA (z ograniczeniami wynikającymi również ze sposobu przygotowania bibliotek cDNA do sekwencjonowania) i nie odzwierciedla jedynie efektywności transkrypcji. Na poziom izoform RNA wpływają również inne czynniki, takie jak stabilność RNA, na które wpływają rozmaite nukleazy komórkowe w tym główna nukleaza mikroprocesora DROSHA. W tym miejscu rodzi się szereg pytań: (i) czy aktywność alternatywnych miejsc poliadenylacji w 3'UTR *VMP1* (rozdział 4.1.5. poniżej) oraz (ii) czy istnienie prekursora *pri-miRNA* z retencją intronu, może mieć wpływ na ilość powstających cząsteczek *miR-21*.

4.1.4. Komórki z wyciszeniem genu *DROSHA* stanowią model w badaniu krótko żyjących prekursorów pri-miRNA

Cząsteczki miRNA powstają z długich, sięgających nawet do 30,000 nt transkryptów pri-miRNA, które następnie podlegają obróbce przy udziale kompleksu mikroprocesora. Głównymi składowymi tego kompleksu jest białko jądrowe DGCR8 oraz związana z nim rybonukleaza *DROSHA*, która jako enzym należy do grupy RNaz III (137). Proces wycinania pre-miRNA z pierwotnych prekursorów pri-miRNA może odbywać się jeszcze w czasie biosyntezy RNA przez polimerazę RNA II (138). Okres półtrwania prekursorów miRNA jest relatywnie krótki i w zależności od analizowanej grupy pri-miRNA wynosi od 8 do 41 minut odpowiednio dla prekursorów metabolizowanych szybko i w średnim tempie (139). Aby analizować zmiany zachodzące w poziomie transkryptów pri-miRNA, jak również mieć pewność, że analizowana cząsteczka stanowi prekursor miRNA, konieczna jest ingerencja w działanie kompleksu mikroprocesora. W celu istotnego obniżenia poziomu ekspresji *DROSHA* na poziomie mRNA (Ryc. 16B), co w dalszej kolejności przekłada się na widoczne obniżenie poziomu białka (Ryc. 16A), w niniejszej pracy posłużono się reagentami interferencji RNA.



Rycina 16 Komórki z obniżonym poziomem białka *DROSHA* jako model badań zmian poziomu prekursorów miRNA. (A) Analiza *western blot* białek z komórek nabłonkowych raka szyki macicy (HeLa) oraz komórek nowotworowych piersi (MCF-7), potraktowanych kontrolnymi dupleksami siRNA (siCTRL, 50 nM) oraz dupleksem siRNA nacelowanymi na mRNA *DROSHA* (siDROSHA, 50 nM) pokazuje wyraźny ubytek ilości białka *DROSHA* w próbach traktowanych siDROSHA. Traktowanie dupleksami siRNA komórek HeLa oraz MCF-7 trwało 48 h. (B) Poziom mRNA *DROSHA* po zastosowaniu siDROSHA (50 nM) w trzech liniach komórkowych – HeLa, MCF-7 oraz linii ludzkich mioblastów (HSKM) jest istotnie niższy w odniesieniu do grupy traktowanej siCTRL (50 nM). Poziom ekspresji został normalizowany względem mRNA *GAPDH*. (C) Analiza poziomu prekursorów miRNA – pri-miR-23b, -27b, -24-1 i -30d po wyciszeniu ekspresji *DROSHA* w komórkach HeLa. W przypadku każdego analizowanego prekursora, jego poziom uległ istotnemu podwyższeniu po wyciszeniu *DROSHA*. Poziom ekspresji pri-miRNA normalizowany był względem mRNA *GAPDH*. Eksperyment został wykonany w trzech powtórzeniach biologicznych dla każdej linii komórkowej, a wskazane wartości stanowią średnią z odchyleniem standardowym.

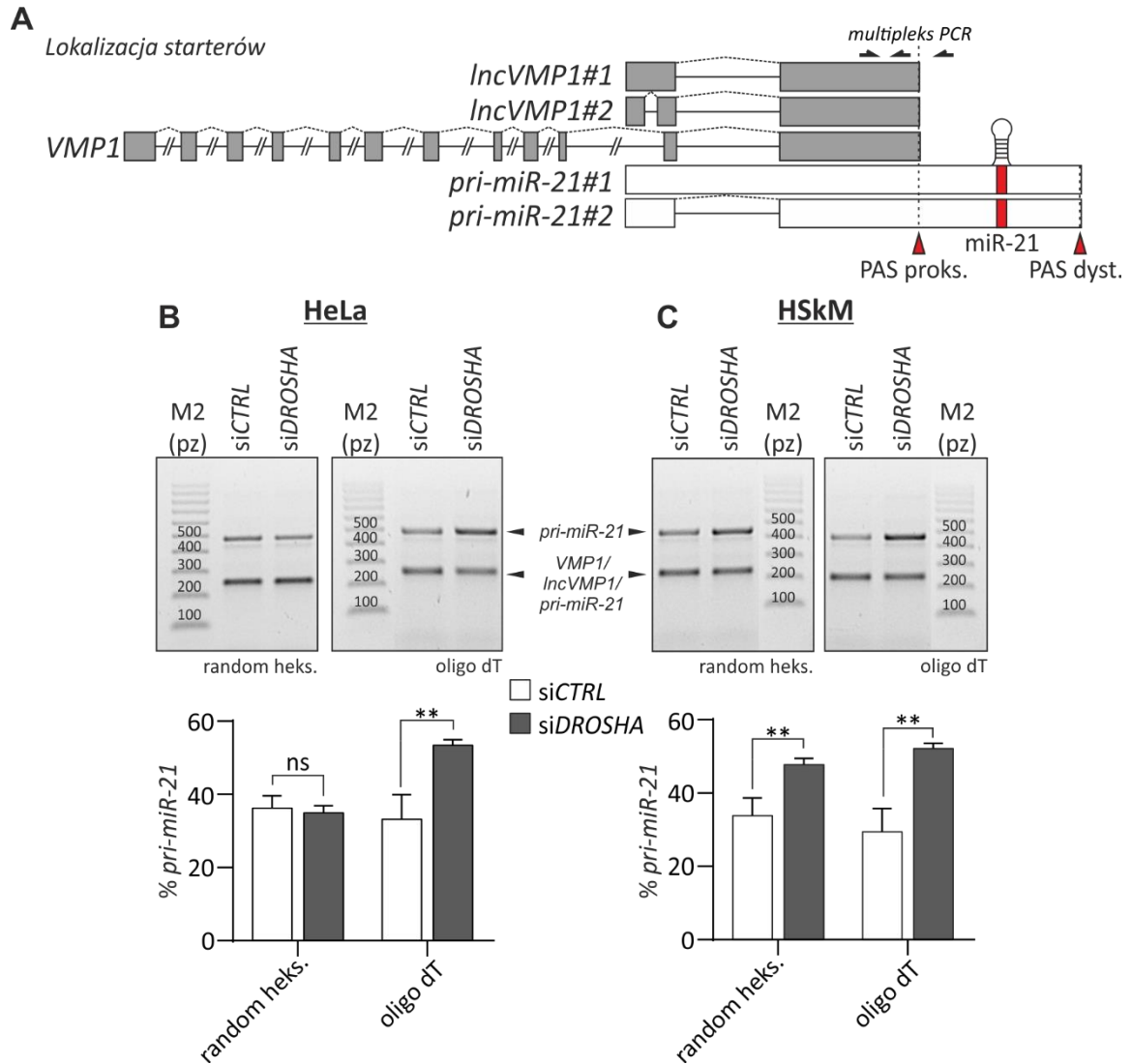
Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (** P< 0,01 *** P<0,001).

(RNAi, ang. *RNA interference*) poprzez działanie dupleksów siRNA (ang. *small interfering RNA*). Działaniu siRNA specyficznych dla DROSHA poddano trzy linie komórkowe HeLa, MCF-7 oraz HSkM, a po 48 h izolowano z nich całkowity RNA lub białko. Po tym czasie widać istotny przyrost ilości prekursorów pri-miRNA (**Ryc. 16C**) w odniesieniu do grupy kontrolnej, czyli komórek poddanych działaniu kontrolnego siRNA. Akumulacji transkryptów pri-miRNA towarzyszy jednoczesny spadek poziomu dojrzałych cząsteczek miRNA. W dalszych częściach pracy niejednokrotnie model ten posłuży do identyfikacji transkryptów będących prekursorami miRNA, jak również jako stan wyjściowy dla eksperymentów oceniających wpływ badanych substancji na poziom ekspresji wybranych pri-miRNA.

4.1.5. Zjawisko alternatywnej poliadenylacji prowadzi do powstania poliadenylowanych transkryptów *pri-miR-21*

Obecność silnego sygnału poliadenylacji (PAS), znajdującego się w odległości około 200 nt przed rozpoznawaną przez kompleks mikroprocesora strukturą *pre-miR-21*, może stanowić istotny element w procesie biogenezy miR-21. W celu sprawdzenia czy rzeczywiście proksymalne miejsce sygnału poliadenylacji (PAS proks.) jest zaangażowane w proces powstawania miR-21 zastosowano modele komórkowe HeLa oraz mioblastyczną linię komórkową (HSkM) z obniżonym poziomem białka DROSHA (zastosowanie duplesu siRNA - siDROSHA). Specjalnie zaprojektowana reakcja multipleks RT-PCR, z użyciem układu trójstarterowego (uniwersalny starter *forward*, oraz dwa startery *reverse*, gdzie jeden zlokalizowany jest przed PAS proks., a drugi za PAS proks.), umożliwiła jednoczesną analizę transkryptów kończących się na PAS proks. (długość produktu 250 pz) oraz izoform dłuższych, których poliadenylacja zachodziła w dystalnym miejscu sygnałowym (PAS dyst.) (długość produktu 480 pz) (**Ryc. 17A**). Reakcje PCR przeprowadzone były z zastosowaniem matrycy cDNA uzyskanej przy użyciu losowych heksamerów oraz oligo_dT. Wykorzystanie jako matrycy cDNA otrzymanego z zastosowaniem oligo_dT pozwala przypuszczać, że analizowane na żelu agarozowym produkty multipleks RT-PCR są transkryptami zakończonymi ogonem poli(A). Prezentowane na wykresach wyniki stanowią procentowy udział dłuższego produktu multipleks RT-PCR, zmierzony na podstawie intensywności sygnału z wybarwionego bromkiem etydyny DNA. Reakcja przeprowadzona w komórkach HeLa (**Ryc. 17B**), na matrycy cDNA

uzyskanej z losowych heksamerów, wykazała brak zmiany w procentowej ilości dłuższego produktu PCR. Analiza procentowego udziału dłuższego produktu w cDNA uzyskanym z zastosowaniem oligo_dT wykazała istotny wzrost poziomu prekursorów *pri-miR-21* po obniżeniu ilości białka DROSHA.



Rycina 17 Zaangażowanie proksymalnego miejsca poliadenylacji VMP1/pri-miR-21 w biogenezę miR-21. (A) Schemat przedstawiający izoformy VMP1 oraz *pri-miR-21*. Dodatkowo na schemacie zaznaczono lokalizację starterów użytych w multiplex RT-PCR. Obraz rozdziła elektroforetycznego na 1,5% żelu agarozowym produktów multiplex RT-PCR. Reakcja została zaprojektowana w celu wyodrębnienia transkryptów kończących się na proksymalnym PAS (wielkość produktu 250 pz) oraz dłuższych, zakończonych na dystalnym PAS (wielkość produktu 480 pz). Analiza przeprowadzona została w dwóch liniach komórkowych (B) komórki HeLa oraz (C) ludzkich mioblastach (HSkM). W obu przypadkach komórki były traktowane dupleksem siRNA kontrolnymi (siCTRL, 50 nM) oraz dupleksem siRNA skierowanym przeciwko DROSHA (siDROSHA, 50 nM). Traktowanie komórek trwało 48 h. cDNA, używany w reakcji, uzyskany został z zastosowaniem losowych heksamerów oraz oligo_dT. Na wykresach przedstawiono udział procentowy dłuższego produktu. Analiza w komórkach HeLa wykazała istotny wzrost procentowej ilości cząsteczek *pri-miR-21* (produkt 480 pz) w przypadku cDNA uzyskanego z oligo_dT, w próbie z obniżonym poziomie białkiem DROSHA. W komórkach HSkM w obu przypadkach ilość transkryptu *pri-miR-21* przeciekającego przez proksymalne PAS był istotnie podwyższony. Eksperyment został wykonany w trzech powtórzeniach biologicznych dla każdej linii komórkowej. Istotność

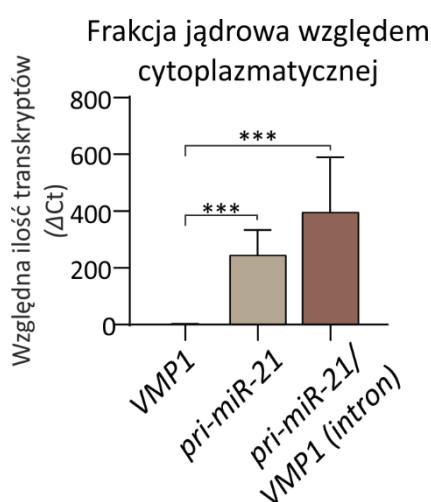
statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (** P < 0,01).

W komórkach HSkM (**Ryc. 17C**), w przypadku cDNA uzyskanego z zastosowaniem losowych heksamerów, jak również z oligo_dT, zaobserwowano istotny wzrost procentowego udziału dłuższej izoformy w próbie siDROSHA. Na podstawie przeprowadzonej analizy, można wnioskować że transkrypty dłuższe, kończące się na PAS dyst., są prekursorami *pri-miR-21*, ponieważ ich udział procentowy istotnie wzrasta po wyciszeniu DROSHA w porównaniu do komórek kontrolnych (siCTRL). Widoczny na obrazie żelu brak zmiany krótszego produktu PCR po zastosowaniu siDROSHA można skonkludować, że transkrypty kończące się w PAS proks. mogą nie stanowić matrycy dla mikroprocesora i tym samym nie być prekursorami dla miR-21. Jednakże należy pamiętać, że siła sygnału pochodzącego krótszego produktu reakcji jest wypadkową poziomu transkryptów kończących się na PAS proks., jak również izoform, których koniec określony jest poprzez działanie PAS dyst. Brak widocznej zmiany w ilości krótszego produktu wynikać może z istotnej przewagi w liczbie transkryptów nie będących prekursorami miR-21. Na podstawie uzyskanych wyników z dwóch linii komórkowych, w których poziom białka DROSHA został istotnie obniżony można przypuszczać, że proksymalny sygnał poliadenylacji transkryptów *VMP1* - jest bezpośrednio zaangażowany w biogenezę miR-21. Transkryptami *pri-miR-21* nazywać możemy jedynie takie cząsteczki RNA, które zakończone są w dystalnym miejscu poliadenylacji, ponieważ tylko ich ilość ulega istotnej akumulacji po wyciszeniu DROSHA.

W celu dodatkowej weryfikacji, czy badane transkrypty RNA, omijające PAS proks. *VMP1* stanowią pulę potencjalnych prekursorów miR-21, porównano poziom poszczególnych fragmentów RNA pochodzących z genu *VMP1* we frakcji jądrowej i cytoplazmatycznej. W eksperymencie przeprowadzono frakcjonowanie komórek fibroblastycznych (GM23963). Po analizie względnego poziomu transkryptów *VMP1*, *pri-miR-21* oraz intronowej części, która jest mieszaniną pre-mRNA *VMP1* oraz prekursorów *pri-miR-21* z retencją intronów, okazało się że ilość badanego *pri-miR-21* w jądrze komórkowym jest istotnie wyższa niż transkryptu *VMP1* (**Ryc. 18**). Taki wynik sugeruje, że długie transkrypty *pri-miR-21*, omijające PAS proks. mogą stanowić źródło dojrzałych cząsteczek miR-21, ponieważ dużo mniej wydajnie opuszczają jądro komórkowe. Zwiększona ilość transkryptów z retencją intronu może wynikać z występowania w jądrze komórkowym zarówno prekursora *pri-miR-21* jak i niedojrzałego

pre-mRNA *VMP1*. Brak *VMP1* we frakcji jądrowej świadczyć może, że dojrzały mRNA *VMP1* jest aktywnie transportowany do cytoplazmy, gdzie podlega translacji.

Na podstawie uzyskanych w tym rozdziale obserwacji oraz wyników można wnioskować, że aktywność alternatywnych miejsc sygnału poliadenylacji w 3'UTR *VMP1* prowadzi do powstania różnych transkryptów, w tym do prekursorów *pri-miR-21*. Poliadenylowane w proksymalnym PAS transkrypty *VMP1* bardzo efektywnie zostają transportowane do cytoplazmy. Natomiast poliadenylowane w dystalnym PAS, prekursor *pri-miR-21* kumulują w jądrze komórkowym, gdzie są przetwarzane przez kompleks mikroprocesora w procesie biogenezy miR-21.

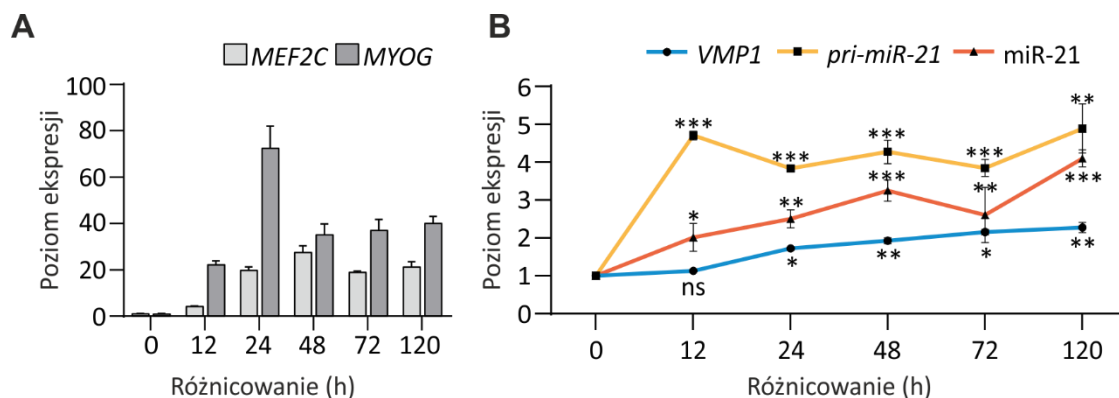


Rycina 18 Względny poziom transkryptów *VMP1* oraz *pri-miR-21* we frakcji jądrowej. Analiza względnej ilości transkryptów analizowana we frakcji jądrowej względem frakcji cytoplazmatycznej w linii komórkowej ludzkich fibroblastów (GM23963). Obserwowane jest istotne wzbogacenie *pri-miR-21* oraz mieszaniny *pri-miR-21* oraz pre-mRNA *VMP1* we frakcji jądrowej. Analizy zostały przeprowadzone w trzech powtórzeniach biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (***) $P < 0,001$.

4.1.6. Poziom ekspresji *VMP1*, *pri-miR-21* oraz miR-21 ulega istotnemu podwyższeniu podczas proliferacji mioblastów

Różnicowanie komórek mięśniowych stanowi bardzo złożony proces, podczas którego zaangażowanych jest wiele czynników genetycznych oraz środowiskowych (140). Bezpośrednie zaangażowanie cząsteczek miRNA w proces miogenezy zostało potwierdzone w wielu pracach badawczych (141, 142). Model badawczy, jakim są różnicujące komórki mięśniowe, stanowi niezwykle przydatne narzędzie do śledzenia globalnych zmian ekspresji genów na poziomie ich transkryptów (143). Dlatego w prezentowanej pracy postanowiono

wykorzystać ten model komórkowy do sprawdzenia, czy istnieje korelacja pomiędzy poziomem ekspresji *VMP1* lub *pri-miR-21* a ilością dojrzałej cząsteczki miR-21.



Rycina 19 Zmiany ekspresji *VMP1*, *pri-miR-21* oraz zmiany poziomu *miR-21* podczas różnicowania ludzkich mioblastów. (A) Potwierdzenie efektywnego zajścia procesu różnicowania na podstawie obserwowanych zmian w poziomie ekspresji markerów różnicowania *MEF2C* oraz *MYOG* w czasie różnicowania, analizowany w sześciu punktach czasowych (0 h, 12 h, 24 h, 48 h, 72 h oraz 120 h od rozpoczęcia różnicowania). Wykres przedstawia relatywny poziom ekspresji normalizowany względem uśrednionej ekspresji genów *GAPDH* i *RPS12*. (B) Względny poziom RNA *VMP1* i *pri-miR-21* oraz dojrzałej cząsteczki *miR-21* w czasie różnicowania. Dla *VMP1* i *pri-miR-21* został badany poziom znormalizowany do uśrednionego poziomu genów referencyjnych *GAPDH* i *RPS12*. Poziom dojrzałej cząsteczki *miR-21* znormalizowany został do *U6*. Wynik dla każdego punktu czasowego przedstawia średnią wartość z 3 powtórzeń biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta i odnosiła się do pierwszego punktu czasowego „0 h”, w którym ustalono wartość referencyjną 1 (* $P < 0,05$ ** $P < 0,01$ * $P < 0,001$).**

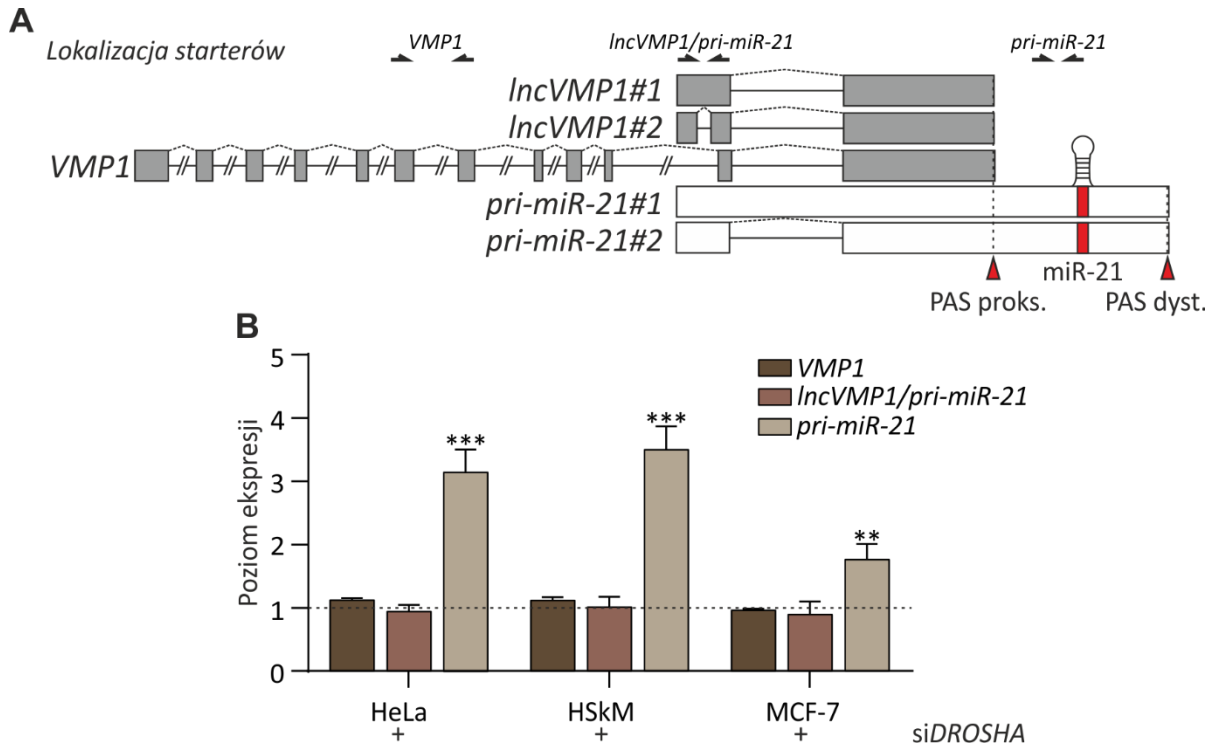
Jednym z markerów, świadczącym o prawidłowym zajściu różnicowania mioblastów, jest zmiana w profilu ekspresji genów czynników transkrypcyjnych *MEF2C* oraz *MYOG* (Ryc. 19A). W czasie różnicowania komórek mioblastycznych, trwającym 120 h zaobserwowano zmiany ilości mRNA *VMP1* oraz prekursora *pri-miR-21* (Ryc. 19B). Co więcej, w czasie trwania eksperymentu ilość dojrzałej cząsteczki miR-21 również wzrosła. Zmiany w poziomie badanych transkryptów oraz miR-21 obserwowane były w sześciu punktach czasowych różnicowania. Analizując tendencję zmian *pri-miR-21* oraz ilości miR-21 widać wyraźną asocjację w profilu zmian. Już po 12 h od indukcji różnicowania poziom badanych transkryptów uległ istotnemu wzrostowi, zarówno na poziomie prekursora jak i dojrzałej cząsteczki miRNA. W kolejnych punktach czasowych poziom *pri-miR-21* utrzymywał swój wzrost, aby osiągnąć maksimum w 120 h różnicowania. Podobne zmiany obserwowano w przypadku ilości dojrzałej cząsteczki miR-21. Analiza poziomu mRNA *VMP1* wykazała istotny wzrost dopiero po 24h, by od tego momentu już się znacząco nie zmieniać. Podobnie jak w

przypadku poziomu miR-21 oraz poziomu jej prekursora, poziom mRNA *VMP1* osiągnął największą wartość w 120 h.

Na podstawie otrzymanych wyników dla różnicowania ludzkich komórek mioblastycznych, można stwierdzić, że wzrost poziomu miR-21 jest mocno zależny od zmian w ilości jego prekursora. Wzrost poziomu prekursora znacznie wyprzedza wzrost poziomu miRNA. Obserwowana asocjacja zmian poziomu miR-21 i mRNA *VMP1* sugeruje istnienie potencjalnej relacji między transkrypcją samego *VMP1*, a biogenezą miR-21 znajdującym się w jego 3'UTR. Uzyskane wyniki pochodzące z wykorzystanego modelu badawczego nie stanowią zatem jasnej odpowiedzi na pytanie czy ekspresja genu *VMP1* jest bezpośrednio zaangażowany w biogenezę miR-21, tzn. czy istnieją cząsteczki mRNA *VMP1* z długim regionem 3'UTR kończącym się na PAS dyst., z których mogą powstawać pre-miR-21. Dlatego w następnych częściach pracy do zbadania zależności między ekspresją *VMP1*, a biogenezą miR-21 wykorzystano inne modele badawcze.

4.1.7. Brak zmian poziomu mRNA *VMP1* w komórkach z niedoborem DROSHA

W 2011 roku J. Ribas i wsp. opisali bezpośrednie zaangażowanie mRNA *VMP1* w biogenezę miR-21 (82). W swoich doświadczeniach zaobserwowali zależność ilości cząsteczek miR-21 od poziomu ekspresji *VMP1*. Według autorów, transkrypt *VMP1* poliadenylowany na dystalnym miejscu mógłby funkcjonować jako pierwotny prekursor miR-21. Taką cząsteczkę nazwali *VMP1-miR-21*. Po traktowaniu komórek związkiem PMA (ang. *Phorbol-12-myristate-13-acetate*), który powoduje wzmożoną aktywność czynnika transkrypcyjnego AP-1, zaobserwowali podobny kierunek zmian *VMP1* oraz *pri-miR-21*.



Rycina 20 Bezpośredni wpływ obniżenia poziomu białka DROSHA na ekspresję VMP1 i pri-miR-21. (A) Schemat przedstawiający izoformy mRNA VMP1 oraz pri-miR-21. Dodatkowo na schemacie zaznaczono lokalizację starterów użytych w reakcji RT-qPCR (B) Względny poziom ekspresji VMP1, IncVMP1 oraz pri-miR-21 w trzech liniach komórkowych - HeLa, HSkM oraz MCF-7 - traktowanych przez 48h siDROSHA (50 nM) w odniesieniu do komórek traktowanych siCTRL (50 nM). Brak istotnych zmian ekspresji VMP1 świadczy o braku, lub nieznacznym zaangażowaniu tego transkryptu w proces biogenezy miR-21. Brak istotnej zmiany w poziomie ekspresji puli dwóch klas transkryptów - IncVMP1 i pri-miR-21 - przy wyciszeniu DROSHA pokazuje, że udział cząsteczek IncVMP1, która nie podlega akumulacji po zadziałaniu siDROSHA, jest dominujący w tej puli. Statystycznie istotna akumulacja prekursorów pri-miR-21, sięgających za PAS proks. świadczy o bezpośredniej aktywności białka DROSHA w przetwarzaniu badanego transkryptu. Relatywny poziom ekspresji normalizowany był względem GAPDH. Dla każdej linii komórkowej analizę wykonano dla trzech powtórzeń biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (** P<0,01 *** P<0,001).

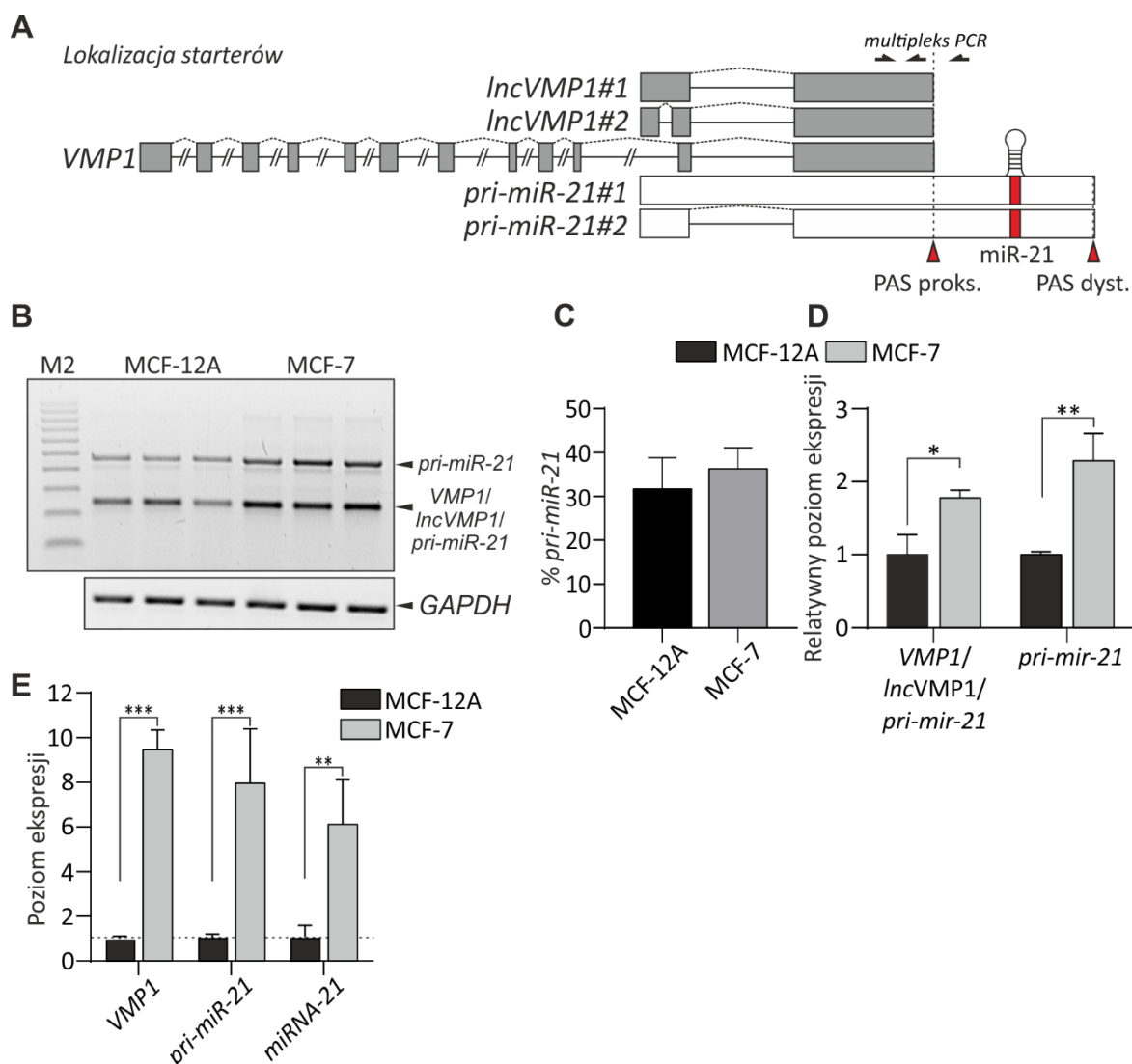
Aby sprawdzić czy istnieje bezpośrednia zależność pomiędzy transkrypcją genu VMP1 i pri-miR-21, a biogenezą miRNA wykorzystano opisane powyżej modele komórkowe z wyciszeniem DROSHA. Dodatkowo przeanalizowano poziom transkryptów, z których powstawać mogą również długie, niekodujące RNA (IncVMP1) kończące się na PAS proks. (Ryc. 20A). Na tej rycinie 20A zaznaczono miejsca, do których przyłączają się startery stosowane w testach PCR. Badania zostały przeprowadzone na trzech liniach komórkowych – HeLa, ludzkie mioblasty (HSkM) oraz komórki nowotworowe gruczołu piersiowego (MCF-7), które poddano działaniu dupleksu siRNA skierowanego przeciw mRNA DROSHA (siDROSHA). W żadnej z testowanych linii komórkowych z wyciszonym genem DROSHA nie zaobserwowano zmian w poziomie mRNA VMP1 w odniesieniu do komórek kontrolnych (siCTRL) (Ryc. 20B). Podobny wynik uzyskano dla poziomu puli RNA obejmującej zarówno IncVMP1 jak i

pri-miR-21 wykorzystując odpowiednią parę starterów. Ten test RT-qPCR pozwala odróżnić te dwie klasy RNA od mRNA *VMP1*, ponieważ starter *forward* zakotwiczony został w sekwencji kodującej intron 10. *VMP1*. Wynik ten może wskazywać na fakt, że bardzo duża ilość transkryptów, mających swój początek w miejscu promotorowym specyficznym dla cząsteczek *pri-miR-21*, podlega terminacji w PAS proks. i nie stanowi substratu dla kompleksu mikroprocesora. Innymi słowy główną pulę RNA powstających z charakterystycznego dla *pri-miR-21* miejsca promotorowego stanowią *IncVMP1*. Jedynie analiza poziomu długich izoform, przekraczających PAS proks. i zakończonych w dystalnym miejscu poliadenylacji odpowiada pozytywnie na wyciszenie *DROSHA*. Ten wynik po raz kolejny wskazuje na to, że prekursorami dla biogenezy pre-miR-21 są jedynie te transkrypty które omijają proksymalne miejsce poliadenylacji i zakończone są na miejscu dystalnym. Z drugiej strony brak zmian w poziomie mRNA *VMP1* w komórkach z obniżoną aktywnością mikroprocesora wskazuje na brak zaangażowania, lub nieznaczne zaangażowanie tego transkryptu w proces biogenezy miR-21. Duża zmienność w akumulacji prekursorów *pri-miR-21* w badanych liniach komórkowych wynikać może z różnego poziomu wyciszenia *DROSHA* (Ryc. 20B), bądź poziomu transkrypcji *pri-miR-21*, który różni się w badanych modelach komórkowych.

4.1.8. Podwyższona aktywność transkrypcyjna prowadzi do podwyższonego poziomu miR-21 w komórkach raka piersi

miR-21 jest jednym z lepiej poznanych i scharakteryzowanych reprezentantów grupy miRNA zaangażowanej w procesy nowotworowe, czyli tzw. *oncomiRs* (144). Poziom miR-21 jest istotnie podwyższony w dużej liczbie powszechnie badanych nowotworów. Jednym z typów nowotworu, w którym poziom miR-21 jest istotnie podwyższony, jest nowotwór piersi (145).

Linia komórkowa MCF-7 jest jedną z najbardziej znanych modelowych linii komórkowych raka piersi. Dzięki licznym badaniom z użyciem linii MCF-7 udało się dokonać wielu obserwacji przybliżających do lepszego zrozumienia mechanizmów kancerogenezy, a tym samym stworzenia odpowiednich terapii antynowotworowych (146). Linia komórkowa MCF-12A stanowiła powszechnie kontrolę w tych badaniach, będąc modelem reprezentującym prawidłowy stan komórkowy.



Rycina 21 Ekspresja VMP1 na poziomie jego mRNA i pri-miR-21 oraz dojrzałego miR-21 w komórkach raka piersi. (A) Schemat przedstawiający izoformy mRNA VMP1 oraz pri-miR-21. Dodatkowo na schemacie zaznaczono lokalizację starterów użytych w multiplex RT-PCR. (B) Rozdział elektroforetyczny produktów multiplex RT-PCR. Zaprojektowana odpowiednio reakcja wyodrębnia transkryptów kończących się na proksymalnym PAS (wielkość produktu 250 pz) oraz dłuższych, zakończonych na dystalnym PAS (wielkość produktu 480 pz). W analizie porównawczej wykorzystano niezmienniczo nowotworowo komórki gruczołu piersiowego (MCF-12A) oraz komórki raka piersi MCF-7 (C) Procentowy udział dłuższego produktu reprezentującego pri-miR-21 nie różnił się między badanymi liniami komórkowymi (D) Analiza relatywnego poziomu ekspresji mRNA VMP1 oraz pri-miR-21 w komórkach MCF-7 w porównaniu do komórek MCF-12A wykazała istotny wzrost ilości w przypadku obu transkryptów. (E) Ilościowa analiza RT-qPCR poziomu ekspresji wykazała istotny statystycznie wzrost mRNA VMP1, pri-miR-21, jak również dojrzałej cząsteczki miR-21 w komórkach MCF-7. Poziom VMP1 i pri-miR-21 normalizowany był względem GAPDH. Poziom miR-21 normalizowany został do U6. Analizy zostały przeprowadzone w trzech powtórzeniach biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (* $P < 0,05$ ** $P < 0,01$ *** $P < 0,001$).

Na poziom miRNA może wpływać wiele różnych czynników takich jak zwiększona transkrypcja ich prekursorów bądź zaburzenia lub modyfikacje w procesie ich dojrzewania, ale

również stabilność miRNA. W celu sprawdzenia jaki mechanizm może być zaangażowany w podwyższony poziom miR-21 w komórkach MCF-7 przeprowadzono szereg eksperymentów. Po pierwsze korzystając z systemu multipleks RT-PCR (**Ryc. 21A-B**) zbadano, czy zaburzenie terminacji transkrypcji w miejscu proksymalnym może prowadzić do wzrostu *pri-miR-21*. Okazało się, że procentowy udział dłuższej izoformy będącej prekursorem *pri-miR-21* nie różni się pomiędzy komórkami MCF-7 a MCF-12A (**Ryc. 21C**). Badanie względnego poziomu ekspresji zarówno transkryptów *pri-miR-21* (górny produkt multipleks RT-PCR) jak również transkryptów będących mieszaniną *VMP1/IncRNA/pri-miR-21* (krótszy produkt multipleks RT-PCR) pokazało podwyższony poziom obu klas RNA (**Ryc. 21D**). Dodatkowa analiza ilościowa poziomu ekspresji mRNA *VMP1* oraz *pri-miR-21* wykorzystująca reakcje RT-qPCR potwierdziła istotnie wyższy poziom obu klas RNA w komórkach MCF-7 w porównaniu do komórek MCF-12A (**Ryc. 21E**). W badanym modelu przeanalizowano również poziom dojrzałej cząsteczki miR-21, który zgodnie z literaturowymi doniesieniami był istotnie podwyższony w komórkach raka piersi.

Na podstawie uzyskanych wyników można stwierdzić, że w przypadku komórek nowotworowych raka piersi MCF-7, za istotnie podwyższony poziom miR-21 odpowiedzialny jest wzrost aktywności transkrypcyjnej. Wzmoczona biosynteza *pri-miR-21* skutkuje powstawaniem większej liczby miRNA. Wzrostowi *pri-miR-21* towarzyszy wzrost mRNA *VMP1*, co może sugerować zmiany epigenetyczne tego *locus* w komórkach nowotworowych. Brak również dowodu, aby zaburzony metabolizm *pri-miRNA* na drodze alternatywnej poliadnylacji był źródłem wzrostu badanego miRNA.

4.2. Alternatywna poliadenylacja ma bezpośredni wpływ na biogenezę miR-21

Koniec 3' większości transkryptów tworzonych przez polimerazę RNA II powstaje na skutek skoordynowanego działania wielu białek, w tym czynników specyficzności cięcia i poiadenylacji (CPSF, ang. *cleavage and polyadenylation specific factor*), czynnika stymulacji cięcia (CstF, ang. *cleavage stimulation factor*) oraz polimerazy poliadenylowej. Proces formowania końców 3' prawie zawsze odbywa się kilkanaście nukleotydów od ściśle określonego miejsca sygnałowego z zakonserwowaną sekwencją 5'-AAUAAA-3' (PAS) (147). Znaczna część transkryptów w swoim niekodującym regionie na końcu 3' (3'UTR) posiada wiele PAS, które poprzez działanie licznej grupy czynników determinują, w którym miejscu przecięty zostanie pre-mRNA i dołączony zostanie ogon poli(A) (19). Możliwość zajęcia poliadenylacji w alternatywnych miejscach wprowadza bardzo dużą zmienność w transkryptomie, która znacząco modyfikuje aktywność wielu genów.

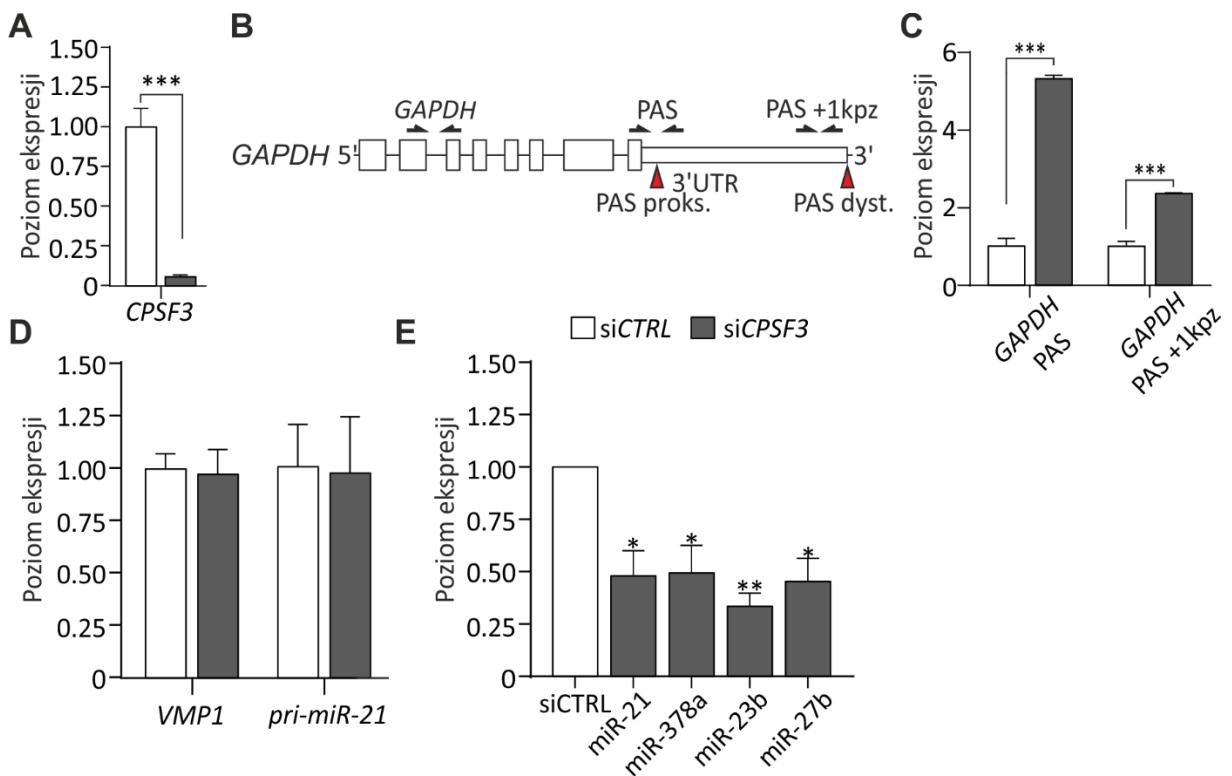
4.2.1. Komórki z wyciszeniem *CPSF3* jako potencjalny model badania wpływu alternatywnej poliadenylacji *VMP1/pri-miR-21* na biogenezę miR-21

Białko CPSF3 jest jednym z wielu czynników zaangażowanych w proces terminacji transkrypcji i wyboru PAS (148). Poprzez swoją aktywność endorybonukleazową, w połączeniu z pozostałymi czynnikami kompleksu CPSF, przeprowadza obróbkę końca 3' nowopowstałego transkryptu i swoim działaniem bezpośrednio determinuje końce 3' mRNA i innych klas RNA.

Celem tej części pracy było zbadanie, czy zmiana poziomu czynników zaangażowanych w kształtowanie końców 3' RNA może wpłynąć na zmianę poziomu miR-21. Aby sprawdzić, czy zmiana poziomu czynnika CPSF3, głównej nukleazy odpowiedzialnej za cięcie transkryptu RNA syntetyzowanego przez polimerazę RNA II, może bezpośrednio wpłynąć na wybór proksymalnego miejsca poliadenylacji w genie *VMP1*, przeprowadzono eksperyment wyciszenia ekspresji genu *CPSF3* w komórkach HeLa (**Ryc. 22A**). Jako eksperyment kontrolny dla określenia efektu niedoboru CPSF3, przeprowadzono analizę zmiany długości kontrolnego transkryptu *GAPDH* (**Ryc. 22B**). Istotny przyrost transkryptów przekraczających miejsce głównego PAS (PAS_proks.) w komórkach traktowanych siCPSF3 w odniesieniu do komórek kontrolnych (traktowanych siCTRL) wskazuje na znaczny wpływ traktowania siCPSF3 na zajęcia terminacji/poliadenylacji tego RNA w miejscu PAS proks. (**Ryc. 22C**). Następnie, w tym modelu komórkowym sprawdzono poziom *VMP1* oraz puli transkryptów *pri-miR-21*. Wydajne wyciszenie ekspresji *CPSF3* nie wpłynęło jednak w znaczący sposób ani na poziom ekspresji

VMP1, ani prekursorów *pri-miR-21* (**Ryc. 22D**). Ponadto analiza poziomu miR-21 oraz trzech innych reprezentantów grupy miRNA (miR-378a, miR-23b, miR-27b) wykazała nieoczekiwane istotne obniżenie poziomu wszystkich badanych cząsteczek miRNA (**Ryc. 22E**).

Brak zmian w poziomie *VMP1*, a zwłaszcza *pri-miR-21* wynikać może z faktu, że poziom wyciszenia *CPSF3* jest nadal niewystarczający do zmiany wyboru proksymalnego sygnału poliadenylacji w *VMP1/pri-miR-21*, co skutkuje brakiem wzrostu liczby transkryptów kończących się na dystalnych miejscach poliadenylacji. Innym wytłumaczeniem może być udział w kształtowaniu końca 3' badanego genu innej nukleazy niż *CPSF3*. Ponadto w komórkach poddanych działaniu *siCPSF3* zaobserwowano znaczne obniżenie poziomu wszystkich badanych miRNA. Efekt ten może być wywołany nie bezpośrednio przez zmianę efektywności poliadenylacji ich prekursorów na skutek niedoboru białka *CPSF3*, ale może też być wynikiem zaburzeń ekspresji wielu genów, w tym tych, które są zaangażowane w biogenezę lub stabilność miRNA. Dlatego też nie prowadzono dalszych badań z wykorzystaniem modeli komórkowych z globalnie modyfikowaną aktywnością terminacji transkrypcji/poliadenylacji RNA, a skorzystano z podejść celowanych bezpośrednio w dojrzewanie RNA *VMP1/pri-miR-21*.

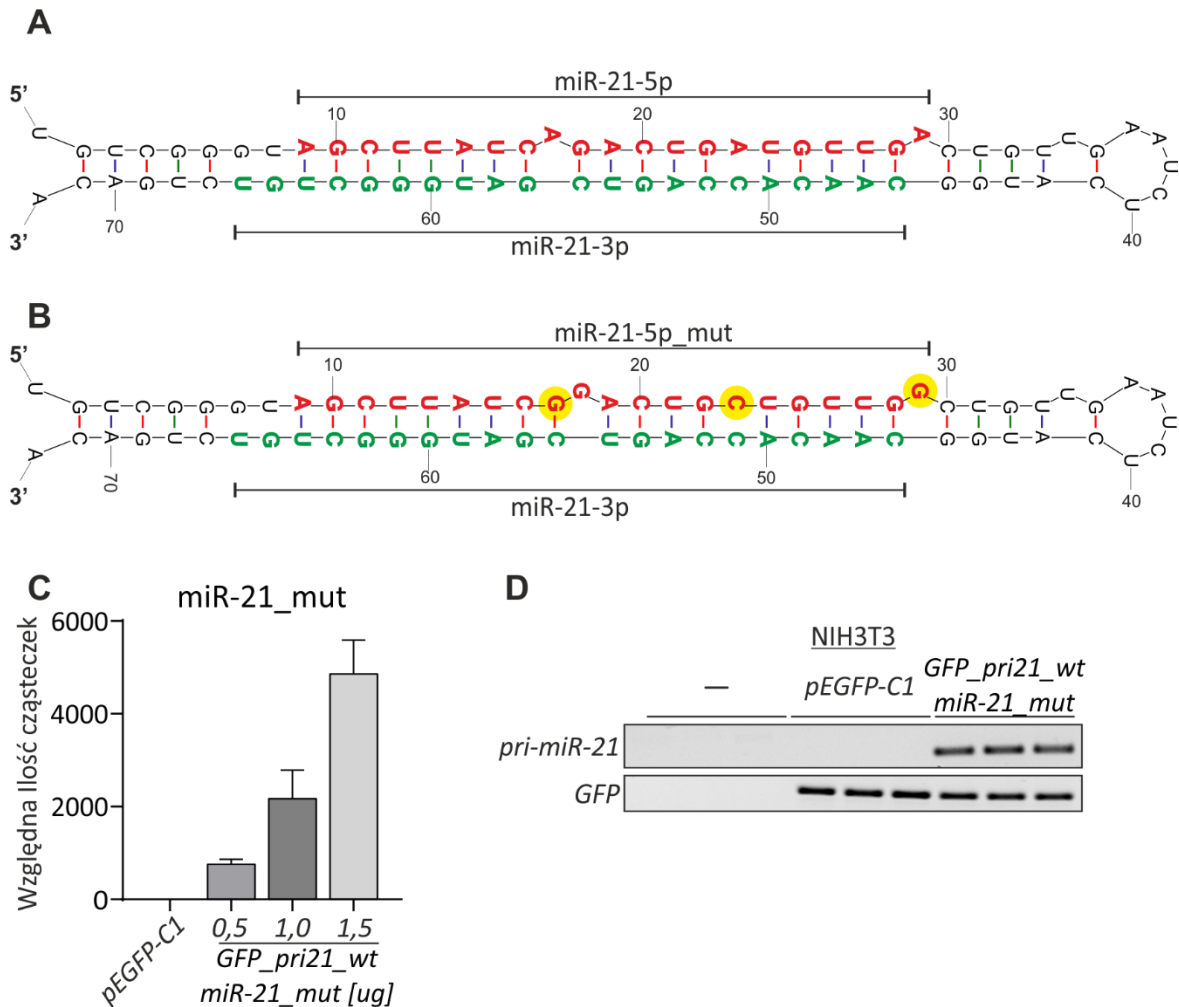


Rycina 22 Wpływ wyciszenia *CPSF3* na poziom ekspresji *VMP1*, *pri-miR-21* oraz cząsteczek miRNA. (A) Poziom ekspresji czynnika *CPSF3* w komórkach HeLa traktowanych dupleksami siRNA skierowanych przeciwko *CPSF3* (si*CPSF3*, 50 nM) w odniesieniu do traktowania komórek kontrolnym siRNA (siCTRL, 50 nM). **(B)** Transkrypt *GAPDH* z zaznaczonymi miejscami PAS w 3'UTR oraz parami starterów zastosowanymi do amplifikacji odpowiednich regionów pre-mRNA. **(C)** Istotny wzrost poziomu transkryptów *GAPDH* przekraczających PAS proks. w obu analizowanych regionach 3'UTR w materiale z si*CPSF3*. Poziom ekspresji badanych regionów 3'UTR *GAPDH* mierzony był względem mRNA *GAPDH*. **(D)** Wyciszenie ekspresji *CPSF3* nie wywołało efektu na poziom transkryptów *VMP1* oraz *pri-miR-21*. Poziom ekspresji liczony był względem *GAPDH*. **(E)** Relatywny poziom cząsteczek miR-21, 378a, 23b oraz 27b w komórkach HeLa z obniżonym poziomem *CPSF3*. Blisko o połowę spadkowi uległy wszystkie badane miRNA, w odniesieniu do próby kontrolnej. Poziom badanych miRNA normalizowany był do U6. Analizy zostały przeprowadzone w trzech powtórzeniach biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (* $P < 0,05$ ** $P < 0,01$ *** $P < 0,001$).

4.2.2. Stworzenie konstruktów genetycznych do badań wpływu dojrzewania prekursora *pri-miR-21* na biogenezę miR-21

Cząsteczka miR-21 jest jedną z najpowszechniej występujących miRNA w szerokim spektrum komórek i tkanek (149). Przez to analiza poziomu zmian miR-21 jest szczególnie utrudniona. Każda manipulacja w procesie dojrzewania jej prekursora może nie być możliwa do zaobserwowania właśnie przez istotne nagromadzenie transkryptów *pri-miR-21*, a zwłaszcza stabilnych, dojrzałych cząsteczek miR-21. Dlatego aby umożliwić obserwację zmian miR-21, stworzono konstrukt genetyczny zawierający pełną sekwencję kodującą ludzki pierwotny prekursor *pri-miR-21* (Ryc. 23A). Jako szkielet wykorzystano wektor pEGFP-C1 zawierający sekwencję kodującą białko zielonej fluorescencji (GFP, ang. *green fluorescent protein*). Aby móc obserwować zmiany zależne od PAS znajdujących się wewnątrz sekwencji kodującej *pri-miR-21* z wektora usunięto miejsce sygnałowe poliadenylacji SV40poli(A). Sekwencja kodująca cząsteczkę miR-21 została zmutowana w taki sposób, aby podczas analizy jej ilości w komórkach możliwym było monitorowanie jedynie zmiany egzogennej cząsteczki, zwanej dalej **miR-21_mut** (Ryc. 23B). Ponadto wprowadzone zmiany sekwencji nie powinny wpłynąć na biogenezę miRNA. Przeprowadzona analiza RT-qPCR wykazała, że ze zmienionej sekwencji kodującej pre-miR-21-mut powstaje dojrzała, cząsteczka miR-21_mut, której poziom można precyzyjnie mierzyć. Wzrastająca ilość wprowadzanego do komórek wektora *GFP_pri21_wt_miR-21_mut* była związana z wyraźnym wzrostem ilościowym miR-21-mut (Ryc. 23C). Przeprowadzony test specyficzności starterów wykazał, że w mysich komórkach NIH3T3 możliwa jest detekcja wyłącznie ludzkich sekwencji *pri-miR-21* pochodzących z wprowadzanego wektora – zarówno startery namnażające *GFP*, jak również *pri-miR-21* nie

prowadziły do amplifikacji mysich fragmentów RNA w nietransfekowanych komórkach NIH3T3 (Ryc. 23D).



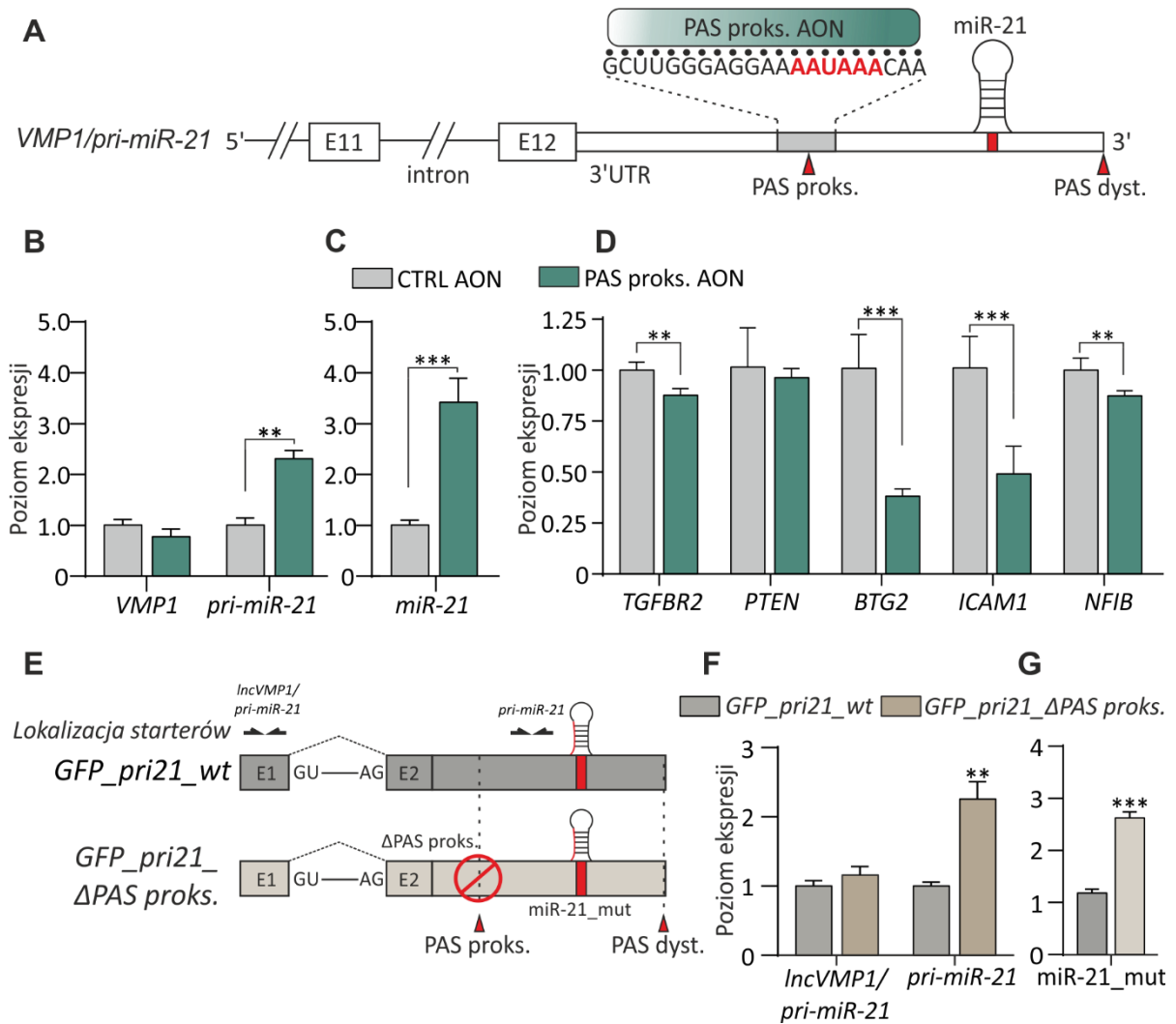
Rycina 23 Badanie konstruktu genetycznego *GFP_pri-miR-21*. **(A)** Fragment drugorzędowej struktury RNA ludzkiego *pri-miR-21* z przedłużoną sekwencją po stronie 5' i 3', która jest rozpoznawana przez mikroprocesor oraz **(B)** struktura pre-miR-21-mut z wprowadzonymi trzema mutacjami punktowymi w sekwencji miR-21-5p. Wprowadzone mutacje nie zaburzają struktury drugorzędowej RNA. Kolorem czerwonym zaznaczona jest sekwencja kodująca miR-21-5p, a kolorem zielonym miR-21-3p. Struktury RNA zostały stworzone za pomocą aplikacji mfold (<http://www.unafold.org/mfold.php>). **(C)** Analiza RT-qPCR mierząca względną ilość miR-21_mut w mysich płodowych komórkach skóry (NIH3T3). Wraz ze wzrostem ilości (0,5-1,0-1,5 ug) stosowanego wektora GFP_pri21_wt_miR-21_mut, wzrasta znacząco ilość wykrywanej cząsteczki miR-21_mut. Grupę kontrolną stanowiły komórki transfekowane wektorem pEGFP-C1 w ilości 1 ug. Poziom mierzonej ilości miR-21_mut normalizowany był do ilości endogennego miR-21 **(D)** Analiza RT-PCR badająca specyficzność użytych w eksperymencie starterów. W sytuacji braku wprowadzenia wektora do komórek NIH3T3, użyte startery na ludzki *pri-miR-21*, ani na GFP nie pozwalają na amplifikację DNA. Startery *pri-miR-21* jedynie namnażają produkt DNA w grupie komórek transfekowanych wektorem GFP_pri21_wt_miR-21_mut. Produkt DNA amplifikowanego fragmentu GFP powstaje również w grupie traktowanej niezmiennym wektorem pEGFP-C1. Analizy zostały przeprowadzone w trzech powtórzeniach biologicznych. Przed odwrotną transkrypcją całkowity DNA był traktowany DNazą w celu eliminacji DNA plazmidowego.

Uzyskany konstrukt genetyczny *GFP_pri21_wt_miR-21_mut* stanowi bardzo przydatne i wiarygodne narzędzie do dalszych badań procesów dojrzewania prekursora *pri-miR-21*.

Dzięki wprowadzeniu mutacji punktowych w sekwencji kodującej miR-21, możliwa jest dokładna analiza zmian ilości powstającego z wektora egzogenego miRNA. W dalszych częściach pracy uzyskany konstrukt oraz jego zmutowane warianty stanowiąc będą podstawę doświadczeń badających mechanizmy mogące wpływać na biogenezę miR-21.

4.2.3. Aktywność PAS proks. w RNA *VMP1/pri-miR-21* ma bezpośredni wpływ na biogenezę miR-21

Ze względu na braku konkludujących wyników dotyczących zmiany poziomu *pri-miR-21* i miR-21 uzyskanych w komórkach z wyciszonym *CPSF3*, postawiono bezpośrednio wpłynąć na aktywność proksymalnego sygnału poliadenylacji. W tym celu wykorzystano dwa podejścia badawcze. W pierwszym zastosowano antysensowe oligonukleotydy (AON, ang. *antisense oligonucleotide*) komplementarne do regionu RNA zawierającego sekwencję AAUAAA stanowiącą sygnał PAS proks. (**Ryc. 24A**). W przeprowadzonym eksperymencie zastosowano cząsteczki AON, będące zmodyfikowanymi cząsteczkami RNA posiadającymi szkielet tiofosforanowy (PS, ang. *phosphothioate*) oraz posiadające podstawioną grupę metylową przy atomie węgla 2' rybozy (2'OMe-RNA-PS). W drugim podejściu zastosowano genetyczną manipulację sekwencji badanego *pri-miRNA*. Przeanalizowano wpływ delekcji PAS proks. w konstrukcie genetycznym zawierającym całą sekwencję kodującą *pri-miR-21* (4,4 kpz).



Rycina 24 Wpływ blokowania i mutacji regionu PAS proks na biogenezę miR-21. (A) Schematyczna prezentacja transkryptu *VMP1* z dwoma ostatnimi eksonami (E11 - ekson 11, E12 - ekson 12) oraz 3'UTR. Zielony blok symbolizuje zastosowany w doświadczeniu 21-nukleotydzowy bloker RNA 2'OMe-PS (PAS proks. AON) komplementarnie wiążący się do miejsca sygnału poliadenylacji „AAUAAA” wraz z sąsiadującymi nukleotydami. Czerwonymi strzałkami oznaczone zostały miejsca poliadenylacji. Uformowana RNA spinka wskazuje miejsce kodujące *miR-21*. (B) Poziom ekspresji *VMP1*, *pri-miR-21* oraz (C) ilość dojrzałej cząsteczki *miR-21* po traktowaniu komórek HSKM blokerem (PAS proks. AON, 75 nM) przez 48 h w odniesieniu do grupy kontrolnej (CTRL AON, 75 nM). Działanie blokera nie wywołało efektu na ekspresję *VMP1*. Zaobserwowano istotny przyrost transkryptu *pri-miR-21* oraz zwiększoną ilość *miRNA-21*. Poziom ekspresji *VMP1* oraz *pri-miR-21* liczony był względem mRNA *GAPDH*. Poziom *miRNA-21* normalizowany był do poziomu *miR-16*. (D) Poziom ekspresji transkryptów będących pod kontrolą *miRNA-21*. Wśród przeanalizowanych pięciu transkryptów, cztery geny (*TGFBR2*, *BTG2*, *ICAM1* oraz *NFIB*) uległy obniżeniu poziomu ekspresji po zastosowaniu blokera PAS proks. Poziom *PTEN* nie uległ zmianie. Relatywny poziom ekspresji na poziomie mRNA badanych genów był mierzony względem mRNA *GAPDH*. (E) Schematyczna prezentacja fragmentu minigenu zawierającego sekwencję kodującą ludzki prekursor *miR-21* (4,4 kpz). Na schemacie zaznaczone zostały miejsca wiązania użytych w analizie starterów. Konstrukty *GFP_pri21_wt* oraz *GFP_pri21_ΔPAS proks.* zawiera mutację sekwencji kodującej cząsteczkę *miR-21*. Dodatkowo w konstrukcie *GFP_pri21_ΔPAS proks.* usunięte zostało PAS proks. Oba konstrukty wprowadzane były do komórek NIH/3T3. Analizując wyniki (F) widać brak zmian amplifikując fragment przed PAS proks. Amplifikacja transkryptów przeciekających PAS proks. wykazała istotny wzrost w odniesieniu do kontrolnego minigenu. Poziom ekspresji normalizowany był względem mRNA *GFP*. (G) Poziom *miR-21_mut* również był istotnie podwyższony w konstrukcie z *ΔPAS proks.* Zmiana *miR-21_mut* normalizowana była do poziomu endogennego *miR-21*. Eksperyment został przeprowadzony w 3 powtórzeniach biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (** P<0,01 *** P<0,001).

Na schemacie (**Ryc. 24A**) zaprezentowano dokładną sekwencję do której wiąże się oligonukleotydyd 2'OMe-RNA-PS. Wprowadzenie do komórek blokerów PAS proks. (PAS proks. AON) nie spowodowało zmian poziomu mRNA *VMP1*. Natomiast poziom transkryptu *pri-miR-21* uległ istotnemu podwyższeniu w komórkach traktowanych blokerem PAS proks. AON w porównaniu do komórek poddanych działaniu kontrolnego antysensownego oligonukleotydu (CTRL AON) (**Ryc. 24B**). Poziom dojrzalej cząsteczki miR-21, podobnie jak poziom *pri-miR-21* istotnie się zwiększył pod wpływem działania PAS proks. AON (**Ryc. 24C**). By wykazać czy zmiana jest istotna funkcjonalnie zbadano geny posiadające w swoich sekwencjach miejsca wiązania dla miR-21 (znane cząsteczki docelowe dla miR-21). Analiza poziomu ekspresji genów będących pod kontrolą miR-21 wykazała, że zastosowanie PAS proks. AON skutkuje istotnym obniżeniem ekspresji czterech z pięciu badanych transkryptów (*TGFBR2*, *BTG2*, *ICAM1*, *NFIB*) (**Ryc. 24D**).

W celu potwierdzenia wyniku stosowania AON maskującego PAS proksymalny *VMP1* oraz *pri-miR-21* przeprowadzono analizę z wykorzystaniem konstruktów genetycznych opisanego dokładnie w rozdziale 2.2. Konstrukt GFP_21_ΔPAS proks. zawiera zmutowaną sekwencję miR-21 w celu precyzyjnego pomiaru poziomu tego miRNA (**Ryc. 23A-B**). Wprowadzenie mutacji wynikało to z tego, że w większości ssaczych linii komórkowych poziom miR-21 jest wysoki lub bardzo wysoki. Uwzględniając dosyć dużą odmienność sekwencji kodujących *pri-miR-21* u człowieka oraz u myszy, konstrukty genetyczne zawierający ludzki wariant tej sekwencji (**Ryc. 24E**) były wprowadzane do embrionalnych fibroblastów pochodzących z myszy (NIH3T3). Dzięki temu, stosując odpowiednio zaprojektowane startery możliwe było badanie poziomu zmian jedynie egzogennie powstającego transkryptu *pri-miR-21*. Badanie regionu zlokalizowanego na początku *pri-miR-21* w mutacji ΔPAS proks. nie wykazało zmian w odniesieniu do wariantu typu dzikiego minigenu. Analiza obszaru RNA zlokalizowanego za zmutowanym miejscem PAS wykazała istotny przyrost dłuższych izoform *pri-miR-21* w sytuacji, gdy PAS proks. został usunięty (**Ryc. 24F**). Dodatkowo mutacja PAS proks. spowodowała istotny wzrost poziomu miR-21_21_mut powstającego ze zmutowanego konstruktów GFP_21_ΔPAS proks. (**Ryc. 24G**).

Przeprowadzone badanie z wykorzystaniem blokerów AON (PAS proks. AON) oraz opierające się na konstrukcie genetycznym z usuniętym PAS proks. (GFP_21_ΔPAS proks.) wykazały, że proksymalny sygnał poliadenylacji znajdujący się przed sekwencją kodującą

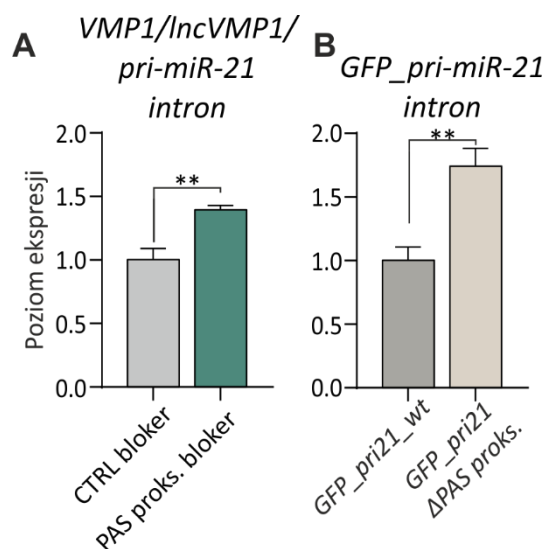
miR-21 może w bezpośredni sposób wpływać na ilość miR-21. Jego zablokowanie bądź usunięcie z konstruktów doprowadziły do istotnego przyrostu dłuższej izoformy RNA i tym samym do zwiększenia ilości dojrzałej cząsteczki miR-21.

4.3. Zależność pomiędzy aktywnością czynników splicingowych działających w 3' ss ostatniego intronu, a zajściem procesu cięcia i poliadenylacji w *pri-miR-21*

Proces dojrzewania prekursorów RNA składa się z kilku etapów. Należą do nich dobudowanie czapeczki na 5' końcu, składanie pre-mRNA oraz poliadenylacja końca 3'. Do początku lat 90. wszystkie te etapy były badane jako całkowicie niezależne procesy, nie wpływające bezpośrednio na siebie. Jednakże w późniejszych latach okazało się, że wszystkie zdarzenia odbywające się podczas dojrzewania prekursorów RNA, a zachodzące w większości podczas transkrypcji, są ze sobą ściśle powiązane (150). Jedną z takich powiązań dla RNA syntetyzowanych przez polimerazę RNA II jest związek splicingu i dojrzewania prekursorów na końcu 3', a dokładniej cięcia i poliadenylacji.

4.3.1. Blokowanie PAS proks. *VMP1/pri-miR-21* oraz jego delecja wpływają na zwiększoną retencję ostatniego intronu

Jednymi z pierwszych badań, które wykazały że proces zachodzący w miejscu sygnałowym poliadenylacji ma związek z wydajnością procesu składania dwóch ostatnich eksonów w pre-mRNA, była praca Niwa i Berget z 1991 (33). W swoim doświadczeniu autorzy wykazali, że mutacja w konsensusowym miejscu poliadenylacji 5'-AAUAAA-3' prowadzi do retencji ostatniego intronu w badanym transkrypcie. Na podstawie uzyskanych przez autorów pracy wyników można wnioskować, że aktywność poliadenylacyjna w pobliskim PAS wywiera pozytywny wpływ na wycinanie ostatniego intronu podczas dojrzewania prekursorów RNA. Te obserwacje zostały potwierdzone w wielu innych badaniach, w tym z zastosowaniem masowego sekwencjonowania RNA (dokładnie opisane w pracy przeglądowej (150).

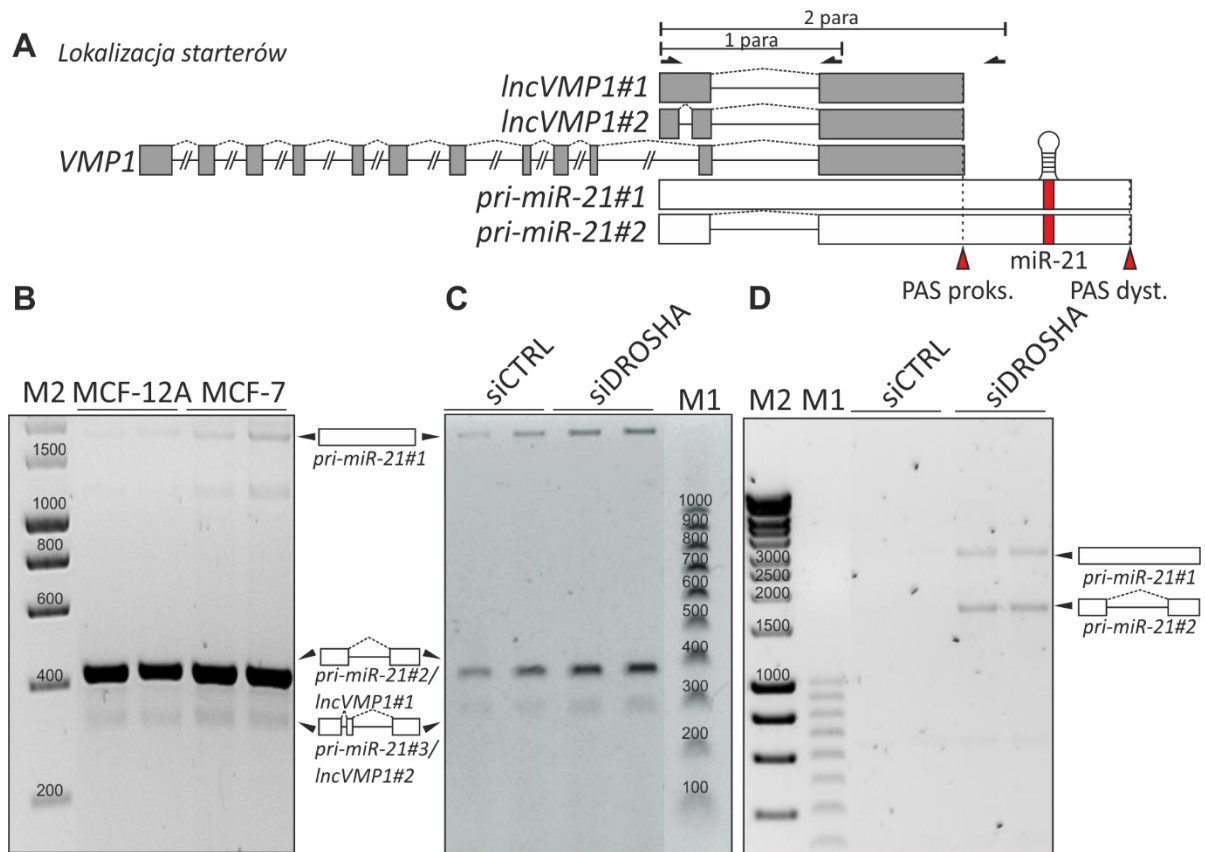


Rycina 25 Retencja ostatniego intronu pre-mRNA *VMP1* oraz *pri-miR-21* po zablokowaniu i mutacji PAS proks. (A) Badanie poziomu retencji intronu 11 *VMP1* w komórkach HSKM traktowanych PAS proks. AON wykazało istotny wzrost poziomu izoform RNA z pozostającym intronem (B) Analiza zmiany ilości izoformy z retencją intronu w konstrukcie genetycznym ze zmutowanym PAS proks. (*GFP_pri21_ΔPAS proks.*) wykazała istotny przyrost badanego transkryptu w odniesieniu do kontrolnego konstruktu (*GFP_pri21_wt*). W przypadku analizy zmian po zastosowaniu blokera PAS proks., genem referencyjnym był poziom *GAPDH*. Jako czynnika normalizującego w badaniu zmian w konstrukcie użyto poziomu ekspresji mRNA *GFP*. Eksperymenty zostały przeprowadzone w 3 powtórzeniach biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (** $P < 0,01$).

Aby sprawdzić, czy w przypadku dojrzewania mRNA *VMP1* oraz prekursora *pri-miR-21* również może dochodzić do interakcji pomiędzy omawianymi procesami wykorzystano materiał z wcześniej opisanego eksperymentu maskowania PAS proks. jak również mutowania proksymalnego sygnału poliadenylacji w konstrukcie genetycznym zawierającym prekursor *pri-miR-21*. Badanie poziomu retencji intronu 11 w mRNA *VMP1* i *pri-miR-21* w komórkach traktowanych oligonukleotydem maskującym PAS proks. wykazało istotny wzrost izoform RNA zawierających intron (Ryc. 25A). Podobnie było w przypadku badania poziomu retencji intronu w konstrukcie genetycznym ze zmutowanym PAS proks. w porównaniu z niezmiennym minigenem typu dzikiego (Ryc. 25B).

Przeprowadzona analiza zmian poziomu transkryptów z retencją intronu wykazała możliwą zależność pomiędzy procesami splicingu oraz cięcia i poliadenylacji w przypadku dojrzewania *pri-miR-21*. Sugeruje to, że w przypadku biogenezy miR-21 już na poziomie dojrzewania prekursora *pri-miR-21* dochodzi do regulacji aktywności proksymalnego PAS poprzez zachodzenie splicingu w ostatnim intronie *VMP1*.

4.3.2. Izoformy *pri-miR-21* z i bez retencji intronu są prekursorami miR-21



Rycina 26 Obecność izoformy *pri-miR-21*/*LncVMP1* z retencją intronu w różnych modelach komórkowych (A) Schemat przedstawiający izoformy *VMP1*, *LncVMP1* oraz *pri-miR-21*. Dodatkowo na schemacie zaznaczono lokalizację starterów użytych w reakcjach RT-PCR. **(B-C)** Obraz rozdziału elektroforetycznego produktów RT-PCR z użyciem starterów amplifikujących zarówno transkrypty z retencją intronu, jak również z wyciętym intronem (1 para). Reakcja wykonana na materiale z komórek MCF-12A i MCF-7 **(B)**, jak również modelu z wyciszeniem *DROSHA* (*siDROSHA*) **(C)** uwidoczniała trzy różne składowane transkrypty *LncVMP1/pri-miR-21*. **(D)** Zmiana startera *reverse*, wiążącego się do sekwencji zlokalizowanej za PAS proks. skutkowało pojawieniem się produktów PCR jedynie cDNA uzyskanego z komórek traktowanych *siDROSHA* (2 para). Długość produktów odpowiada prekursorom *pri-miR-21* z retencją intronu 11. oraz z wyciętym intronem. Ze względu na znaczne różnice w długości amplifikowanych fragmentów cDNA uzyskane wyniki nie mają charakteru ilościowego pozwalającego na oszacowanie udziału izoform z retencją intronu.

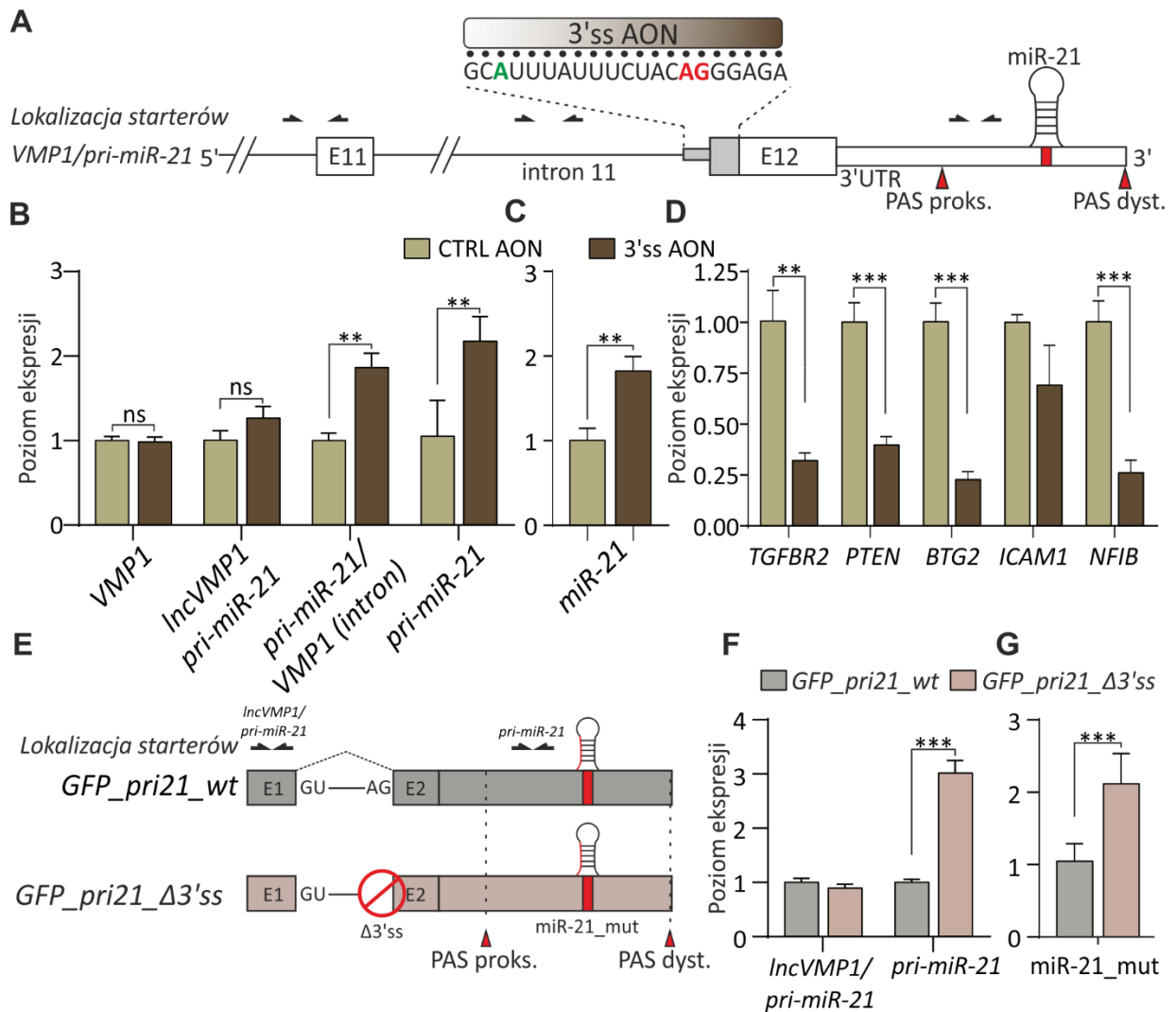
Dane przedstawione w pierwszej części pracy, w rozdziale 4.1.3., pochodzące z analiz sekwencjonowania RNA z różnych ludzkich tkanek, wykazały obecność niekodujących transkryptów *VMP1*, których poziom ekspresji był relatywnie wysoki w prawie każdej badanej tkance **(Ryc. 15B)**. Korzystając z modelu komórkowego z istotnie obniżoną ilością białka *DROSHA*, jak również porównując komórki raka piersi MCF-7 do komórek nowotworowo niezmiennych MCF-12A stwierdzono obecność izoform *pri-miR-21*, które różniły się od siebie obecnością sekwencji intronowej **(Ryc. 26B-C)**. Analiza RT-PCR, wykorzystująca startery amplifikujące izoformy mające swój początek w 10. Intronie *VMP1* oraz kończące się za miejscem PAS proks. (starter *reverse* wiąże się za PAS proks.) wykazała, że względna ilość

obydwu izoform jest podobna w przypadku analizy komórek traktowanych siDROSHA (Ryc. 26D). W komórkach kontrolnych (siCTRL) transkrypty te są niewykrywalne tą metodą prawdopodobnie dlatego, że prekursor pri-miRNA mają bardzo krótki okres półtrwania i prawie natychmiast są procesowane przez kompleks mikroprocesora, a jednocześnie ze względu na duży rozmiar amplifikowanego fragmentu cDNA.

Przedstawione wyniki wskazują na to, że obecność dwóch istotnie różnych izoform *pri-miR-21* może mieć znaczenie podczas procesu biogenezy miR-21. Biorąc pod uwagę powyżej opisane wyniki, sugerujące związek pomiędzy splicingiem, a poliadenylacją (Ryc. 26) można założyć, że obecność retencji intronu może wpływać na wydajną biogenezę miR-21.

4.3.3. Obecność miejsca splicingu 3' (3'ss) ostatniego intronu *pri-miR-21* ma wpływ na biogenezę miR-21

Aby sprawdzić, czy retencja intronu w izoformie *pri-miR-21* ma wpływ na ilość powstającego miR-21 przeprowadzono dwie serie eksperymentalne. W pierwszej zastosowano antysensowne oligonukleotydy maskujące miejsce splicingowe 3' (3'ss) - połączenie intron11/ekson12 (3'ss AON) genu *VMP1* (Ryc. 27A). W założeniu eksperymentu, aktywność AON poprzez zablokowanie miejsca splicingu miała doprowadzić do powstania większego udziału izoform *pri-miR-21* z retencją intronu. Dodatkowo zamaskowanie miejsca dla wiązania się U2AF, czynnika odpowiedzialnego bezpośrednio za prawidłowe wycinanie intronów i składanie eksonów, mogło doprowadzić do zaburzeń zależności pomiędzy procesem splicingu ostatniego intronu, a poliadenylacji w PAS proks. Po traktowaniu komórek 3'ss-AON zaobserwowano brak istotnych zmian w poziomie mRNA *VMP1* oraz regionu kodującego transkrypty *lncVMP1*. Z kolei poziom izoform zawierających intron *VMP1/pri-miR-21* był istotnie podwyższony w porównaniu do komórek traktowanych kontrolnym AON (CTRL-AON) (Ryc. 27B). Wynik ten świadczy o częściowym zaburzeniu procesu składania *VMP1/pri-miR-21*.



Rycina 27 Wpływ zastosowania blokowania miejsca 3'ss w ostatnim intronie *VMP1/pri-miR-21* oraz mutacji 3'ss w konstrukcie genetycznym *GFP_pri21_Δ3'ss* na biogenezę miR-21. **(A)** Schematyczna prezentacja transkryptu *VMP1* z dwoma ostatnimi eksonami (E11-ekson 11, E12-ekson 12) oraz 3'UTR. Brązowy blok symbolizuje zastosowany w doświadczeniu 21-nukleotydzowy bloker 2'OMe-PS (3'ss AON) komplementarnie wiążący się do połączenia intronu 11. i eksonu 12. Czerwonymi strzałkami oznaczone zostały miejsca poliadenylacji. Dodatkowo zaznaczone zostało miejsce sekwencji pre-miR-21. **(B)** Poziom ekspresji puli izoform *VMP1*, *lncVMP1*, zawierających intron 11, mRNA *VMP1* i *pri-miR-21* oraz **(C)** ilość dojrzałej cząsteczki miR-21 po traktowaniu komórek HeLa blokerem (3'ss AON, 75 nM) przez 72 h w odniesieniu do komórek kontrolnych (CTRL AON, 75 nM). Działanie blokera nie wywołało efektu na poziom mRNA *VMP1* jak również na poziom *lncVMP1*. Zaobserwowano jednak istotny wzrost retencji intronu 11 *VMP1*, *pri-miR-21* oraz zwiększoną ilość miR-21. Poziom ekspresji badanych transkryptów normalizowany był do mRNA *GAPDH*. Poziom miR-21 normalizowany był do poziomu miR-16. **(D)** Zmiany poziomu transkryptów będących pod kontrolą miR-21. Wśród przeanalizowanych pięciu mRNA, dla czterech (*TGFBR2*, *PTEN*, *BTG2*, oraz *NFIB*) zaobserwowano obniżenie poziomu po zastosowaniu 3'ss AON. Poziom mRNA *ICAM1* nie uległ zmianie. Relatywny poziom ekspresji badanych genów był normalizowany do *GAPDH*. **(E)** Schematyczna prezentacja wstawki minigenu zawierającego sekwencję kodującą ludzki prekursor miR-21 (4,4 kpz). Konstrukty *GFP_pri21_wt* oraz *GFP_pri21_Δ3'ss*. zawiera mutację sekwencji kodującej miR-21. Dodatkowo w konstrukcie *GFP_pri21_Δ3'ss*. usunięte zostało 15 nukleotydów obejmujących miejsce styku intron/ekson po 3' stronie intronu. Oba konstrukty wprowadzane były do komórek mysich NIH/3T3. Analizując wyniki widać brak zmian stosując sondy qPCR specyficzne dla fragmentu RNA przed PAS proks. Ilościowa RT-PCR specyficzna dla transkryptów przeciekających PAS proks. wykazała istotny wzrost ilości transkryptów zawierających tę sekwencję w odniesieniu do kontrolnego minigenu. Poziom ekspresji normalizowany był względem mRNA *GFP*. Poziom *miR-21_mut* również był istotnie podwyższony w konstrukcie z $\Delta 3'ss$. Zmiana *miR-21_mut* normalizowana była do poziomu endogennego miR-21. Eksperyment został

przeprowadzony w 3 powtórzeniach biologicznych. Istotność statystyczna obserwowanych zmian została określona za pomocą testu t-Studenta dla danych niesparowanych (** P< 0,01 *** P<0,001).

Wynik ten świadczy o częściowym zaburzeniu procesu składania *VMP1/pri-miR-21*. Podobny efekt zaobserwowano w przypadku badania ilości *pri-miR-21* mierzonych za PAS_proks. W tym przypadku zaobserwowano istotny wzrost poziomu tego RNA po działaniu 3'ss AON (**Ryc. 27B**). Ilość dojrzałej cząsteczki miR-21, podobnie jak ilość jej prekursora, uległa również istotnemu podwyższeniu (**Ryc. 27C**). Analiza zmian poziomu ekspresji genów będących pod kontrolą działania miR-21 wykazała statystycznie istotne obniżenie w przypadku czterech z pięciu badanych mRNA - *TGFBR2*, *PTEN*, *BTG2* oraz *NFIB*. (**Ryc. 27D**).

W drugim podejściu eksperymentalnym zastosowano konstrukty genetyczne, zawierające sekwencję kodującą ludzki *pri-miR-21* (**Ryc. 27E**), a opisane szczegółowo w rozdziale 4.2.2. Badania te wykazały że mutacja miejsca 3'ss, wiążącego U2AF a znajdującego się na połączeniu intron 11/ekson 12 prowadzi do istotnego przyrostu transkryptów *pri-miR-21* mierzonych za PAS_proks (**Ryc. 27F**). Zmianie nie uległ region przed PAS_proks., co wskazuje na częściową aktywność PAS_proks. w badanym modelu. Wraz ze wzrostem *pri-miR-21* powstającego z konstruktów genetycznych, w badaniu obserwowano również zwiększoną ilość cząsteczek *miR-21mut* po wprowadzeniu mutacji w 3'ss (**Ryc. 27G**).

Badania wpływu zmian splicingowych wywołanych blokowaniem miejsca 3'ss oraz mutacją tego miejsca w konstrukcie genetycznym wykazały istotne zmiany w ilości prekursora *pri-miR-21* oraz zwiększoną liczbę cząsteczek miR-21 w obu modelach badawczych. Taka obserwacja może sugerować że aktywność splicingowa w 3'ss *pri-miR-21*, a konkretnie wiązanie się w tym miejscu odpowiednich czynników splicingowych, może mieć związek z aktywnością PAS_proks. Obecność izoformy *pri-miR-21* z retencją intronu może mieć silny związek z powstawaniem dojrzałej cząsteczki miR-21.

5. Dyskusja

Znane od ponad 20 lat, krótkie niekodujące cząsteczki miRNA stanowią obiekt intensywnych badań naukowców. Po wpisaniu frazy „*microrna*” w wyszukiwarce w bazie PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) otrzymujemy 112,791 rekordów. Wynik ten pokazuje, jak dużą uwagę skupiają na sobie te małe cząsteczki RNA. Na dzień dzisiejszy, najpopularniejsza przeglądarka deponująca informację o miRNA – miRBase (<http://www.mirbase.org/>) – tylko dla organizmu człowieka zawiera aż 1,917 różnych cząsteczek miRNA. Cząsteczki miRNA stanowią istotny element w regulacji ekspresji niemal wszystkich genów w komórkach zarówno zwierzęcych jak i roślinnych. Są ważnym regulatorem w utrzymaniu homeostazy oraz przebiegu prawidłowego rozwoju złożonych organizmów (149). Liczne patogenne procesy, w tym nowotworzenie, związane są z zaburzonym poziomem licznych cząsteczek miRNA (108, 116, 117). W przypadku nowotworów, istotnie zmieniony poziom miRNA kontrolujących ekspresję protoonkogenów oraz onkogenów, prowadzi do niekontrolowanego wzrostu ilości komórek nowotworowych (94). Biorąc pod uwagę fizjologiczną funkcję miRNA, poznanie mechanizmów i procesów kontrolujących ich biogenezę wydaje się uzasadnione.

Proces powstawania cząsteczek miRNA jest procesem wieloetapowym, w który zaangażowana jest liczna grupa białek (50, 63, 151) (Rycina 7). U zwierząt proces ten rozpoczyna się w jądrze komórkowym, a kończy w cytoplazmie, gdzie dojrzałe miRNA pełnią swoją funkcję. Geny kodujące miRNA zlokalizowane są w różnych miejscach w genomie. Wyróżniamy miRNA zlokalizowane wewnątrz genów kodujących białka, jak również niekodujących transkryptów RNA – miRNA wewnątrz-genowe (ang. *intragenic*) (Rycina 8). Zdecydowana większość wewnątrz-genowych miRNA zlokalizowana jest w regionach intronowych genów kodujących białko oraz w lncRNA (84). Przeprowadzone badania określające miejsce inicjacji transkrypcji dla genów kodujących miRNA wykazały, że liczna grupa miRNA zlokalizowana wewnątrz genów-gospodarzy, posiada swoje własne, niezależne miejsca promotorowe (80). Jest to bardzo istotna informacja wskazująca na to, że RNA będące pierwotnymi prekursorami dla cząsteczek miRNA, mogą powstawać niezależnie i dodatkowo dojrzewać podobnie jak produkty ich genów-gospodarzy. Bardzo duża część miRNA zakodowana jest w miejscach poza-genowych (ang. *intergenic*). W tym przypadku wiedza na temat długości i budowy transkryptów prekursorowych pri-miRNA pozostaje wciąż mało poznana i wymaga dalszych

badani. Zdecydowana większość genów miRNA transkrybowana jest przez RNA polimerazę II (61). Pierwotne pri-miRNA, podobnie jak pre-mRNA podlegają potranskrypcyjnej obróbce, polegającej na dobudowaniu czapeczki na końcu 5' oraz ogona poli(A) na końcu 3'. Jako odmiana cząsteczek lncRNA pri-miRNA mogą również podlegać procesom składania (152). Innym ciekawym zjawiskiem związanym z biogenezą miRNA jeszcze w jądrze komórkowym jest efektywność działania kompleksu mikroprocesora na poszczególnych substratowych RNA. Kompleks mikroprocesora odpowiedzialny jest za rozpoznanie odpowiednich struktur prekursorów miRNA oraz wycięcie 60-90 nukleotydowych pre-miRNA ze znacznie dłuższych pri-miRNA. Białko z rodziny RNaz III - DROSHA poprzez naprowadzenie związane z działaniem białka rozpoznającego podwójną nić RNA - DGCR8, wycina strukturę RNA tworzącą układ spinki do włosów (ang. *hairpin loop*) i przyczynia się do powstania wolnych końców 5' i 3' zarówno pre-miRNA jak i pozostałych fragmentów pri-miRNA. W kilku badaniach wykazano, że aktywność mikroprocesora prowadząc do powstania niechronionego końca 5', i w konsekwencji degradacji RNA poprzez działanie egzonukleazy XRN2, co skutkuje terminacją transkrypcji pri-miRNA (62, 153). Jednakże, w dużej części przypadków pierwotne transkrypty pri-miRNA w pełni dojrzewają i na końcu 3' są poliadenylowane.

Jednym z miRNA, które ulegają wysokiemu poziomowi ekspresji w wielu tkankach jest miR-21 (Rycina 12, Rycina 13). Zaburzenia ekspresji miR-21 jest również powszechnym zjawiskiem występującym w przypadku szerokiego spektrum chorób nowotworowych (154). miR-21 zaangażowany jest w istotne procesy związane z rozwojem organizmów, w tym na etapie embriogenezy, podczas której jego poziom znacznie się zmienia (155). W ciągu dwóch dekad, badanie znaczenia cząsteczki miR-21 stało się głównym priorytetem dla wielu zespołów badawczych (prace przeglądowe dotyczące miR-21 w nowotworzeniu – (107, 122, 128, 144, 145, 156). Jednakże pomimo tak dużej wiedzy na temat zaangażowania miR-21 w wiele kluczowych procesów biologicznych, wiedza na temat dojrzewania tej cząsteczki nadal jest niekompletna i wymaga znacznego pogłębienia.

Sekwencja kodująca miR-21 zlokalizowana jest w regionie 3'UTR genu kodującego białko membranowe VMP1 (Rycina 12, Rycina 13). W 2004 roku Xuezhong Cai i wsp. (134) jako pierwsi opisali istnienie transkryptu *pri-mir-21* o długości około 3433 nukleotydów, który swój początek ma w 11. intronie genu *VMP1*, gdzie zlokalizowane jest silne miejsce promotorowe. Nie wskazano wówczas na to, aby transkrypt ten miał podlegać procesowi składania.

Dodatkowo na jego końcu 5' wykazano obecność struktury czapeczki, a na końcu 3' ogona poliadenylowego. Autorzy w swojej pracy zasugerowali, że fakt niepodlegania splicingu może mieć bezpośredni wpływ na proces cięcia i poliadenylacji (CPA) w miejscu sygnałowym poliadenylacji (PAS) znajdującym się bezpośrednio przed sekwencją kodującą miR-21.

W 2007 roku Dennis Löffler i wsp. również opisali początek transkryptu *pri-miR-21* w intronie 11 *VMP1* (119), który poprzez regulację przez czynnik transkrypcyjny STAT3 powstawałby niezależnie od swojego genu-gospodarza. Badanie przeprowadzone rok później przez Shuij Fujita i wsp. wykazało, że transkrypt *pri-miR-21* swój początek może mieć również kilkaset par zasad wcześniej, niż to było opisywane w 2007 roku (106). Autorzy opisali miejsce wiązania dla białka aktywującego 1 (AP-1), znajdujące się w 11. intronie *VMP1*, które po stymulacji związkiem PMA (ang. *Phorbol-12-myristate-13-acetate*) aktywnie promowało transkrypcję prekursora *pri-miR-21*. Ciekawą obserwacją był brak zmian w poziomie mRNA *VMP1* przy stymulacji PMA, co wiązano z brakiem zależności między biogenezą miR-21 a transkrypcją genu-gospodarza. Ponadto autorzy wykazali, że duża część transkryptów rozpoczynających się w dodatkowym miejscu promotorowym, jest poliadenylowana w proksymalnym PAS, podobnie jak mRNA *VMP1*. Jedynie długa izoforma *pri-miR-21*, z retencją intronu 11. *VMP1*, długości około 4,3 kbp sięgała za proksymalny PAS i stanowiła prekursor dla miR-21. Po raz kolejny pojawiła się zatem informacja, że niezłożone, długie izoformy *pri-miR-21* stanowią mogą matrycę dla powstania dojrzałej cząsteczki miR-21.

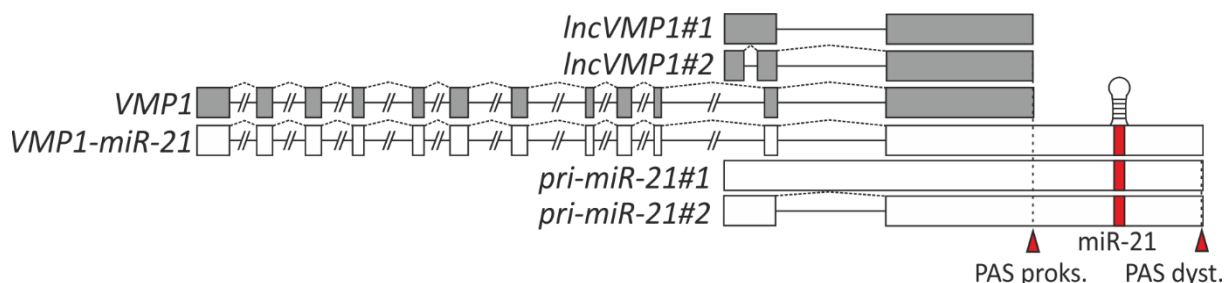
W 2012 roku Judit Ribas i wsp. powtórzyli eksperyment traktowania komórek związkiem PMA i ponownie zaobserwowali istotny wzrost ekspresji transkryptów *pri-miR-21* (82). Jednakże niezgodnie z poprzednimi badaniami (106), autorzy nowszej pracy zaobserwowali przyrost poziomu mRNA *VMP1*. Argumentowali to innym czasem ekspozycji na PMA. Na podstawie uzyskanych wyników założyli, że również transkrypt *VMP1*, może być bezpośrednio zaangażowany w biogenezę tej cząsteczki miRNA. Taką izoformę nazwali *VMP1-miR-21*. W swoim badaniu, korzystając z techniki mapowania końców transkryptów cDNA (RACE, ang. *Rapid amplification of cDNA ends*) wykazali również aktywność dwóch PAS, zlokalizowanych za sekwencją kodującą miR-21. Według autorów PAS znajdujące się bliżej sekwencji kodującej miR-21 miałyby kończyć transkrypt *VMP-miR-21*. Natomiast odsunięty o niecałe 1 kbp PAS odpowiedzialny byłby za generowanie końca transkryptów *pri-miR-21* rozpoczynających się w miejscu promotorowym wewnątrz 10. Intronu *VMP1*.

W 2018 roku przeprowadzono kolejną analizę transkryptów *pri-miR-21* (136). Opierając się na modelu komórkowym dojrzewających bydłych oocytów, autorom pracy udało się odkryć dwie izoformy *pri-miR-21*, które rozpoczynały się w tym samym miejscu promotorowym, znajdującym się w 10. Intronie *VMP1* oraz zakończone były w tym samym PAS. Jediną różnicą była obecność sekwencji intronowej pomiędzy eksonem 11. a 12. *VMP1*. Podobnie jak w przypadku pracy Dennis Löffler i wsp., transkrypcja *pri-miR-21* miałaby być regulowana poprzez czynnik transkrypcyjny STAT3. Jediną różnicą była lokalizacja miejsca promotorowego, które zmapowano przed eksonem 11. *VMP1*.

Powyżej opisane badania dotyczące powstawania prekursorowej cząsteczki *pri-miR-21* wskazują, że ekspresja *pri-miR-21* może być procesem bardzo złożonym. Analizując wyniki uzyskane w opisanych pracach, nie można jednoznacznie stwierdzić czy i w jakim stopniu transkrypcja genu *VMP1* zaangażowana jest w proces biogenezy miR-21. Dodatkowo obecność oraz aktywność PAS znajdującego się bezpośrednio przed sekwencją kodującą miR-21 potencjalnie może odgrywać istotną rolę w biogenezie tej cząsteczki. Mechanizm alternatywnego składania transkryptów *pri-miR-21* również może mieć znaczący wpływ na proces biogenezy miR-21. Biorąc pod uwagę powyżej opisane doniesienia naukowe, w pracy doktorskiej postanowiono dokładnie przyjrzeć się potranskrypcyjnym mechanizmom biorącym udział w dojrzewaniu prekursora miR-21. W niniejszej pracy główny nacisk położono na zbadanie mechanizmów alternatywnej poliadenylacji oraz alternatywnego splicingu *pri-miR-21* oraz wzajemnej relacji między nimi

W pierwszej części pracy szczególną uwagę zwrócono na zmienny poziom ekspresji cząsteczki *pri-miR-21* w szerokim spektrum ludzkich tkanek (Rycina 12, Rycina 13, Rycina 14). Nieregularny poziom ekspresji *pri-miR-21* zarówno pod względem ilościowym jak i jakościowym może być ściśle powiązany z rolą, jaką pełni w danej tkance cząsteczka miR-21, ale może i dotąd niepoznaną rolę znacznie dłuższej niekodującej cząsteczki *IncVMP1*. Dodatkowo, opierając się na danych pochodzących z analiz RNA-seq, zaobserwowano między innymi w tkance gruczołu piersiowego obecność dużej liczby odczytów zmapowanych do obszaru intronu (Rycina 12, Rycina 15). Fakt ten, biorąc pod uwagę wcześniejsze doniesienia literaturowe, wydaje się być interesujący w odniesieniu do przebiegu biogenezy miR-21. Występowanie silnego miejsca inicjacji transkrypcji, rozpoznawanego przez liczne czynniki promotorowe (Rycina 12), wewnątrz sekwencji kodującej *VMP1*, czyni *pri-miR-21* ciekawym

obiektem badawczym. Należy zauważyć, że transkrypt *pri-miR-21* w swojej sekwencji zawiera miejsca 5' oraz 3' składania RNA, występujące również w pre-mRNA *VMP1*. Przez to *pri-miR-21* podlegać może regulacji splicingu podobnie jak produkt jego genu-gospodarza (Rycina 28).



Rycina 28 Schemat prezentujący transkrypty *IncVMP1*, *VMP1* oraz *pri-miR-21*. Kolorem szarym zaznaczone zostały izoformy kończące się w miejscu proksymalnego sygnału poliadenylacji (PAS proks.) niebędące pierwotnymi prekursorami miR-21. Kolorem białym oznaczone zostały transkrypty, z których w późniejszym etapie powstać może dojrzała cząsteczka miR-21. Transkrypty te zakończone są w dystalnym PAS (PAS dyst.).

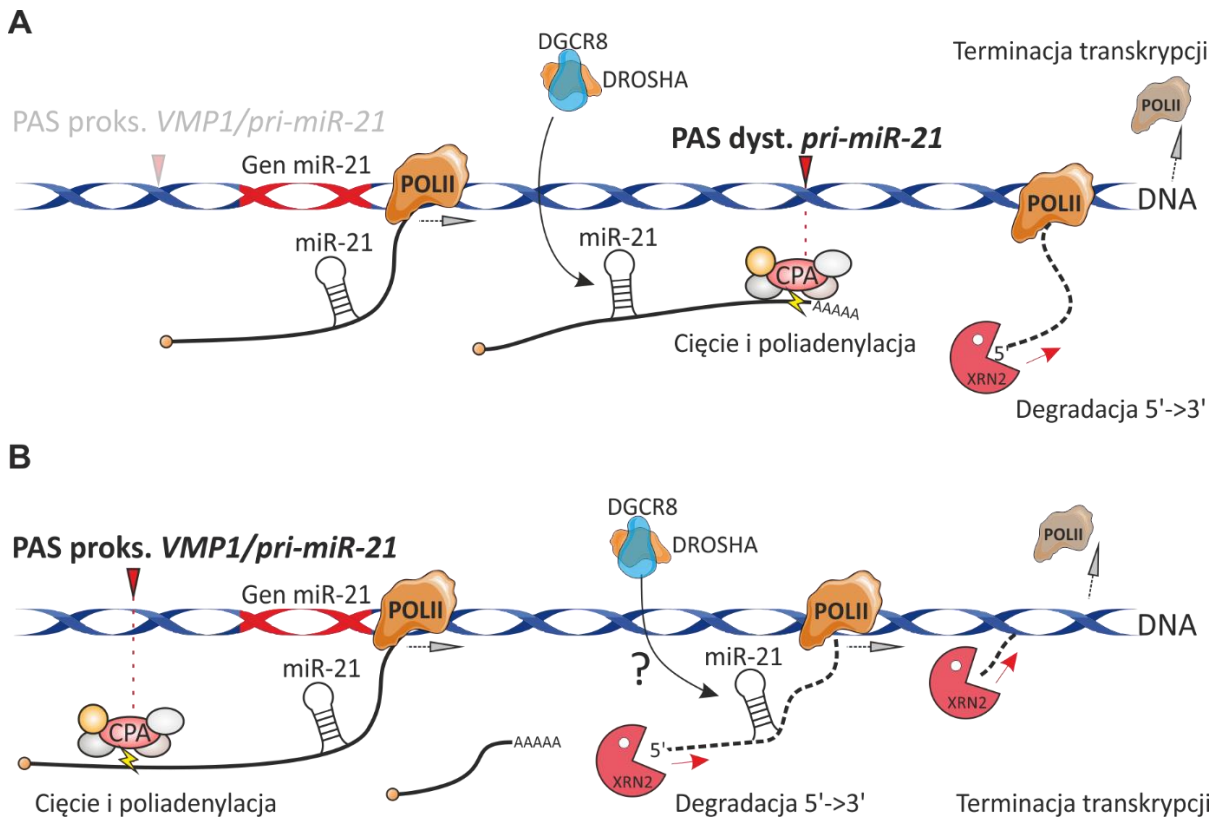
Do analizy zmian poziomu oraz możliwej detekcji transkryptów będących prekursorami miRNA, zastosowano model komórkowy z obniżonym poziomem białka DROSHA (Rycina 16), głównego enzymu kompleksu mikroprocesora (80). W przeciwieństwie do wyników uzyskanych przez Judit Ribas i wsp. (82), przy istotnym obniżeniu komórkowej aktywności DROSHA, nie zaobserwowano zmian w poziomie mRNA *VMP1* (Rycina 20). Wynik ten uzyskano dla trzech linii komórkowych (HeLa, HSkM oraz MCF-7). Brak zmian poziomu mRNA *VMP1* może wynikać z faktu, że transkrypt ten w niewielkim stopniu, o ile w ogóle, zaangażowany jest w biogenezę miR-21. Innym potencjalnym wytłumaczeniem braku zmian w poziomie ekspresji genu-gospodarza *VMP1* może być krótszy czas prowadzenia eksperymentu z użyciem dupletu siDROSHA. Przeprowadzona na potrzeby pracy analiza poziomu mRNA, wykonana na materiale wyizolowanym z frakcji jądrowej oraz cytoplazmatycznej prawidłowych ludzkich fibroblastów, wykazała bardzo efektywny transport mRNA *VMP1* do cytoplazmy. Możliwe, że jedynie niewielki procent *VMP1-miR-21* jest rozpoznany i procesowany przez kompleks mikroprocesora. Również analiza transkryptów *IncVMP1* wykazała brak wpływu obniżonego poziomu DROSHA. Wynik ten pozwala stwierdzić, że proksymalny PAS jest bardzo aktywny i wpływa na zajęcie CPA nie tylko mRNA *VMP1*, ale również powstających z silnego miejsca promotorowego w 10. intronie *VMP1* transkryptów *IncVMP1*. Jak dotąd w literaturze nie pojawiły się badania wyjaśniające mechanizm stojący za wyborem proksymalnego lub dystalnych PAS w 3'UTR *VMP1*. Z badań przeprowadzonych w tej pracy można przypuszczać,

że jednym z modulatorów tego procesu jest albo wycięcia intronu 11 albo wiązanie się białek do miejsca 3' splicingu.

W pracy z 2015 roku Peng Du i wsp. odkryli i opisali dodatkowy etap w procesie dojrzewania cząsteczek miRNA pochodzących z klastra genetycznego miR-17-92 (157). Okazało się, że białko CPSF3 oraz homolog czynnika splicingu ISY1, poprzez swoją aktywność tworzą pośredni produkt w biogenezie miRNA. Ten intermediat nazwano pro-miRNA (ang. *progenitor miRNA*). Powstanie pro-miRNA ma istotny wpływ na efektywne dojrzewanie pięciu z sześciu miRNA z tego wielocistronowego prekursora. Powstanie pro-miRNA zwiększa aktywność mikroprocesora, który wycina pojedyncze pre-miRNA. Bezpośrednie wpłynięcie na proces CPA, poprzez obniżenie poziomu ekspresji jednego z głównych genów odpowiedzialnych za cięcie RNA podczas terminacji transkrypcji - *CPSF3* - nie miało istotnego wpływu na biogenezę miR-21 (Rycina 22). W niniejszej pracy nie zaobserwowano zwiększenia ilości izoform *pri-mir-21* w badanej linii komórkowej HeLa z wyciszeniem *CPSF3*. Brak zmiany w przypadku PAS znajdującym się przed sekwencją kodującą miR-21 można tłumaczyć nieznacznym zaangażowaniem *CPSF3* w terminację transkrypcji w tym konkretnym miejscu. Prowadząc wyciszenie *CPSF3* zakładano, że wywoła to globalne zaburzenie procesu poliadenylacji i tym samym wpłynie na wzrost ilości prekursorowej i dojrzałej cząsteczki miR-21. Po przeprowadzonej analizie okazało się jednak, że nie tylko poziom miR-21, ale również wszystkich innych badanych cząsteczek miRNA uległ istotnemu obniżeniu. Sytuacja ta spowodowana mogła być globalnym zaburzeniem poliadenylacji wielu genów i tym samym wpłynąć na istotne zmiany poziom biogenezy lub stabilności wielu, jeśli nie wszystkich, cząsteczek miRNA. Okazało się zatem, że taki model komórkowy nie jest użyteczny w monitorowaniu wpływu wyboru miejsca poliadenylacji na biogenezę miR-21.

W literaturze możemy znaleźć informację, że miR-21 może być jednym z tzw. miRNA „przeciekających” (ang. *readthrough*) (158). Grupa ta stanowi miRNA, których sekwencje są zlokalizowane tuż za PAS i mogą powstawać mimo efektywnego cięcia transkryptu w tym PAS. Sytuacja taka jest potencjalnie możliwa, ponieważ pomimo zajścia procesu CPA, transkrypcja nadal jest prowadzona przez RNA polimerazę II (159). Pamiętać należy bowiem, że RNA powstający po zajściu CPA nie jest już zabezpieczony po stronie 5' i bardzo szybko ulega degradacji głównie przez egzonukleazę XRN2 (160). Jest to zjawisko bardzo powszechne i znane jest jako model „torpedo” terminacji transkrypcji. Dzięki temu zjawisku,

RNA polimeraza II może zakończyć syntezę RNA i oddysocjować od matrycy DNA (**Rycina 29**). Można założyć, że niewielka odległość sekwencji kodującej miR-21 od proksymalnego PAS (około 200 pz) może być kluczowym elementem dla efektywnego przecieku transkryptu przez PAS i tym samym stworzenie możliwości utworzenia struktury RNA będącej substratem dla mikroprocesora.



Rycina 29 Terminacja transkrypcji oraz model „torpedo” w odniesieniu do biogenezy miR-21. (A) Schematyczne przedstawienie zachodzącego procesu cięcia i poliadenylacji (CPA) w dystalnym miejscu sygnału poliadenylacji (PAS dyst.). Ominięciem przez polimerazę II proksymalnego miejsca sygnału poliadenylacji (Pas proks.) i zajęcie CPA w PAS dyst. stwarza matrycę w postaci dojrzałego prekursora *pri-miR-21* dla kompleksu mikroprocesora (DROSHA/DGCR8) co prowadzi w dalszej kolejności do biogenezy miR-21. Uwolniony, niezmodyfikowany koniec 5' RNA zostaje rozpoznany przez egzonukleazę XRN2, która degraduje RNA z aktywnością 5'→3' i uwalnia polimerazę II (POLII). **(B)** Schematyczne przedstawienie zachodzącego procesu (CPA) w PAS proks. genu *VMP1* oraz *pri-miR-21*, który prowadzi do uwolnienia niemodyfikowanego końca 5' powstającego RNA. Za PAS proks. znajduje się sekwencja kodująca pre-miR-21 rozpoznawana przez kompleks mikroprocesora. Po zajęciu CPA dochodzi do powstania niezabezpieczonego końca 5' rozpoznawanego przez egzonukleazę XRN2, która degraduje taki RNA (model *torpedo*). XRN2 katalizująca trawienie RNA prowadzi do kolizji z polimerazą RNA II (POLII i oddysocjowanie tego holoenzymu) od matrycowego DNA, co skutkuje terminacją transkrypcji. Nie wiadomo w jaki sposób aktywność XRN2 i degradacja RNA wpłynąć może na działanie kompleksu mikroprocesora. Wydajniej zachodząca degradacja RNA z udziałem XRN2 może negatywnie wpływać na biogenezę miR-21. Biogeneza miR-21 jest zatem zależna od kinetyki powstania kompleksu terminacyjnego i cięcia w PAS proks.

Jednak proces degradacji RNA przez XRN2 jest bardzo wydajny i dynamiczny. Co ciekawe, opisywane wcześniej białko CPSF3 biorące udział w CPA, również pełni istotną rolę w mechanizmie terminacji transkrypcji zgodnie z modelem „torpedo”. Brak tego białka prowadzi

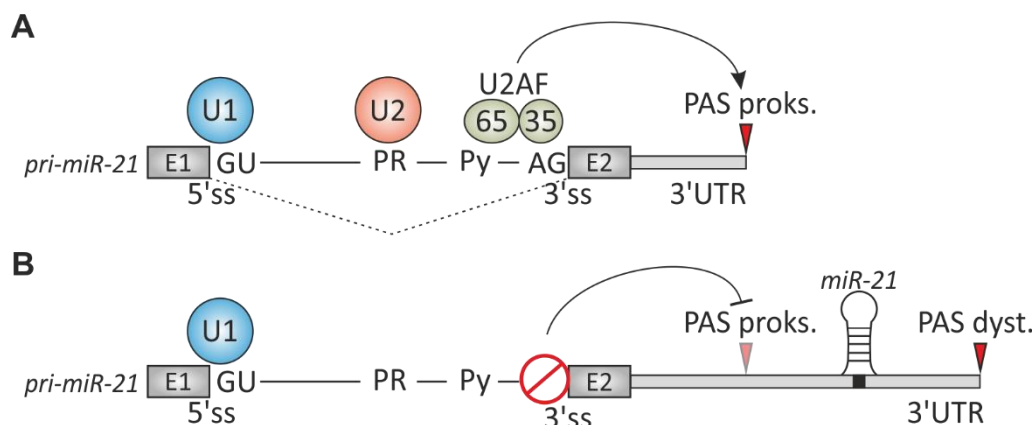
do globalnego zaburzenia terminacji transkrypcji i wyraźnego wydłużenia pierwotnych transkryptów większości genów. W pracy (161) badacze użyli modelu, w którym bardzo wydajnie i szybko byli w stanie usunąć białko CPSF3. Okazało się, że rolą CPSF3 jest spowolnienie działania polimerazy RNA II i tym samym efektywne zajście degradacji RNA przez działania egzonukleazy XRN2.

Przeprowadzone w pracy analizy multipleks PCR (Rycina 17), z wykorzystaniem komórek z obniżonym poziomem białka DROSHA, wykazały że *pri-miR-21*, będący przeciekiem za proksymalny PAS jest poliadenylowany (użycie cDNA otrzymanego przez zastosowanie starterów oligo-dT) i ulega akumulacji przy niedoborze DROSHA. Wynik ten pokazuje, że prekursor *pri-miR-21* stanowią pulę transkryptów, które w dużej mierze są poliadenylowane i w takiej postaci przetwarzane są przez kompleks mikroprocesora.

Uwzględniając wcześniejsze doniesienia o aktywności PAS znajdującym się przed sekwencją kodującą miR-21 (82, 106) i aby sprawdzić czy proces alternatywnej poliadenylacji może być bezpośrednio zaangażowany w biogenezę miR-21, w pracy doktorskiej postanowiono wykorzystać blokery, którymi były oligonukleotydy 2'OMe-RNA-PS (AON, ang. *antisense oligonucleotide*). Możliwość specyficznego wiązania się blokerów do cząsteczek RNA umożliwia zamaskowanie wybranego, funkcjonalnego regionu dla oddziaływania z białkami (162). Na potrzeby pracy zaprojektowano bloker, który specyficznie wiązał się do proksymalnego PAS *VMP1* (PAS proks. AON) i w założeniu miał znacząco ograniczyć działanie kompleksów białek odpowiedzialnych za CPA w celowanym miejscu RNA (Rycina 24). Okazało się, że zastosowanie blokera PAS proks. istotnie wpływa na obniżenie wydajności terminacji syntezy RNA w proksymalnym PAS i zwiększa ilość wrażliwych na poziom DROSHA transkryptów *pri-miR-21*. To przekłada się na zwiększoną liczbę cząsteczek miR-21 w komórkach traktowanych PAS proks. AON. Analiza puli transkryptów powstających z konstruktu genetycznego ze zmutowanym PAS proksymalnym potwierdziła ten wynik. Obserwacja ta okazała się zgodna ze wcześniejszymi doniesieniami wskazującymi na bezpośrednie zaangażowanie proksymalnego PAS w biogenezę miR-21. Jednakże eksperymenty opisane w niniejszej pracy pokazują tę zależność w sposób bezpośredni.

Na chwilę obecną istnieje wiele dowodów potwierdzających bezpośrednią zależność pomiędzy procesem składania pre-mRNA, a efektywnością procesu CPA (prace przeglądowe (150, 163)). Oba te procesy, zachodzące podczas transkrypcji prowadzonej przez polimerazę

RNA II, wzajemnie się regulują poprzez skomplikowane sieci oddziaływań między specyficznymi białkami. Jedną z takich relacji jest bezpośredni związek czynnika pomocniczego U2 (U2AF) z czynnikiem hydrolizującym RNA podczas procesu CPA (CF Im59) (31). Badanie przeprowadzone przez Stefanie Millevoi i wsp. w 2006 roku wykazało, że koniec aminowy czynnika U2AF65 wchodzi w interakcje z CF Im59 i tym samym stymuluje proces CPA. Autorzy w swojej pracy wykazali, że brak U2AF65 istotnie zmienił wydajność CPA. Co ciekawe, inne badanie ujawniło, że U2AF65 na swoim końcu N zawiera motyw blokujący działanie polimerazy poli(A) (164). Z drugiej strony wydajność dojrzewania końca 3' RNA ma duży wpływ na składanie nowopowstającego transkryptu. Już w 1991 roku Maho Niwa i wsp. opisali wpływ mutacji sekwencji sygnału poliadenylacji AAUAAA, znajdującej się w regionie 3'UTR na zaburzenie wycinania ostatniego intronu (33). Przeprowadzona w rozprawie doktorskiej analiza poziomu RNA zawierającego ostatni intron *VMP1* zarówno w komórkach traktowanych blokerem PAS proks. jak i komórkach ekspresujących konstrukt genetyczny z mutacją PAS proks. wykazały zwiększoną retencję tego intronu (Rycina 26). Wynik ten potwierdza istnienie silnego związku między procesem splicingu ostatniego intronu i poliadenylacją w PASproks w dojrzewaniu *VMP1/pri-miR-21* (Rycina 30).



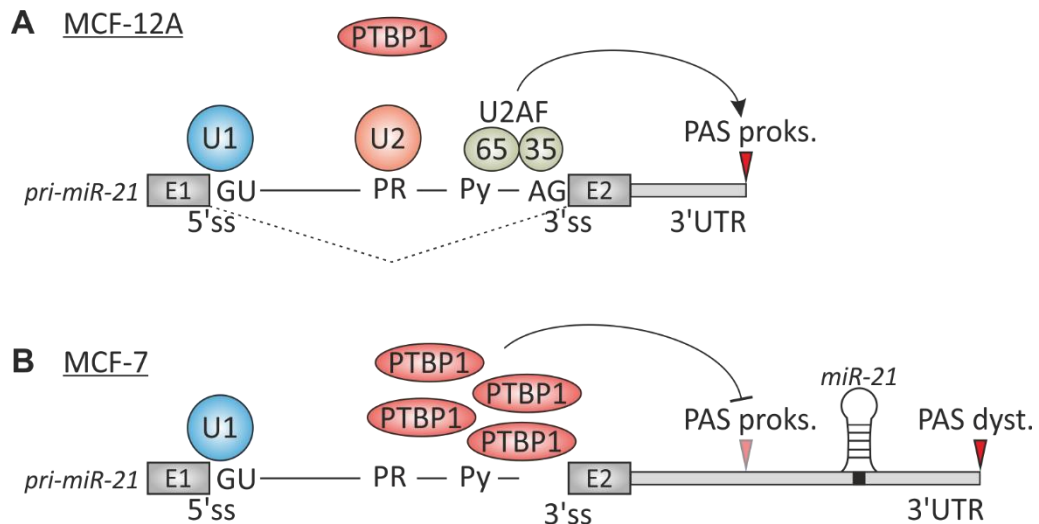
Rycina 30 Schematyczna prezentacja potencjalnego działa czynników U2AF na biogenezę miR-21. (A) wiązanie się czynników splicingowych do miejsc 5' (5'ss) oraz 3' (3'ss) powoduje efektywne wycinanie intronu 11 i zajęcie procesu składania *pri-miR-21*. Dodatkowo, poprzez bezpośrednią aktywność U2AF65 z CF Im59 biorącym udział w procesie cięcia oraz poliadenylacji (CPA), dochodzić może do aktywnego CPA w proksymalnym miejscu sygnałowym (PAS proks.). W takiej sytuacji przeważają izoformy nie zawierają sekwencji kodującej miR-21. (B) W sytuacji, gdy wiązanie U2AF do miejsca 3'ss jest niemożliwe (mutacja tego miejsca lub zastosowanie blokera) dochodzi do zaburzenia w procesie składania i retencji intronu. Ponadto, brak wiązania się U2AF w 3'ss powodować może zaburzenie CPA w PAS proks. i transkrypcja *pri-miR-21* terminowana jest w okolicach dystalnego miejsca sygnału poliadenylacji (PAS dyst.). W takiej sytuacji wzrasta poziom *pri-miR-21* i miR-21. E1 – ekson 1; E2 – ekson 2; PR – punkt rozgałęzień; Py – trakt polipirymidynowy

Eksperyment z wykorzystaniem blokera maskującego wiązanie się białek U2 oraz U2AF35/65 w ostatnim intronie *VMP1* wykazał istotny wpływ na zaburzenie splicingu

pri-miR-21 (Rycina 27). Równocześnie ograniczenie wiązania U2 i U2AF35/65 do miejsca splicingu po stronie 3' zwiększyło poziom transkryptów *pri-miR-21* oraz dojrzałych cząsteczek miR-21. Efekt ten jest zgodny z założeniem, że wiązanie się U2AF65 wzmacnia aktywność CPA w niedaleko położonym PAS (31). Również analiza efektu wywołanego przez usunięcie miejsca składania po stronie 3' w konstrukcie genetycznym pokazała całkowite zahamowanie splicingu intronu 11 *VMP1* jak również pozytywny wpływ na biogenezę miR-21. Okazało się, że po zmutowaniu 3'ss ilość transkryptów kończących się na dystalnych PAS była istotnie podwyższona. Występowanie długiej izoformy *pri-miR-21* z retencją intronu, może wynikać z braku wiązania się i aktywności U2AF65. Powstawanie izoformy RNA z retencją intronu może powodować aktywację CPA w dystalnym miejscu PAS i tym samym regulować efektywność biogenezy miR-21 (Ryc. 1A-B).

Białko wiążące się do traktów polipirymidynowych 1 (PTBP1, ang. *polypyrimidine tract-binding protein 1*) przyłącza się do sekwencji bogatych w polipirymidyny znajdującej się powszechnie na końcach 3' intronów i tym samym blokuje możliwość wiązania się czynników U2AF. Tym samym, wiązanie się PTBP1 może modyfikować proces wycinania intronów (165). PTBP1 poprzez swoje działanie może nie tylko zmieniać efektywność procesu składania, ale również ma wpływ na wybór alternatywnych PAS (APAS) (166). Badanie przeprowadzone na materiale pobranym z glejaka wykazały, że PTBP1 po związaniu z RNA może zamaskować znajdujące się w odległości około 100 nt PAS i tym samym istotnie wpłynąć na aktywność CPA (166). Poziom białka PTPB1 jest istotnie podwyższony w komórkach raka piersi (MCF-7), co według autorów pracy ma wpływ na zaburzoną proliferację komórek (167). Wyciszenie ekspresji *PTBP1* spowodowało znaczące zahamowanie proliferacji oraz podziałów komórek nowotworowych w badaniach *in vitro*. Jedną z obserwowanych zmian w wynikach uzyskanych w pracy doktorskiej był istotnie podwyższony poziom pierwotnych prekursorów *pri-miR-21* oraz poziom dojrzałych cząsteczek miR-21 w komórkach MCF-7 (Rycina 21). Ponadto w tych komórkach nowotworowych w porównaniu do komórek prawidłowych MCF-12A, stwierdzono obecność izoform *pri-miR-21* z retencją intronu. Uwzględniając informację o istotnie podwyższonym poziomie białka PTBP1 można przyjąć hipotezę, że PTBP1 który poprzez swoją aktywność wpływa zarówno na zaburzenie splicingu jak i zmiany w aktywności APAS, może regulować poziom miR-21, poprzez zaburzenie dojrzewania prekursora *pri-miR-21* (Rycina 31). Jego podwyższony poziom w komórkach MCF-7 asocjuje z podwyższonym poziomem miR-21.

Uwzględniając obserwacje niniejszej pracy doktorskiej oraz doniesienia literaturowe, przyszłe badania powinny skupiać się na zbadaniu bezpośredniej interakcji PTBP1 z transkryptami *pri-miR-21* oraz efektem tej interakcji na dojrzewanie tego RNA. Nie można wykluczyć, że czynniki regulujące potranskrypcyjne dojrzewanie *pri-miR-21*, będą w istotny sposób wpływać na efektywność biogenezy miR-21.



Rycina 31 Model potencjalnego wpływu białka PTBP1 na biogenezę miR-21 w prawidłowych (MCF-12A) oraz nowotworowych zmienionych komórkach raka piersi (MCF-7). (A) W komórkach MCF-12A ilość białka PTBP1, w porównaniu do komórek MCF-7 nie jest podwyższona i nie konkuruje ono z czynnikami U2AF w wiązaniu się do traktów polipirymidynowych. Tym samym nie prowadzi do zaburzeń w splicingu, co może powodować zwiększenie efektywności cięcia oraz poliadenylacji (CPA) w proksymalnym miejscu sygnału poliadenylacji (PAS proks.) (B) W przypadku komórek raka piersi (MCF-7) ilość PTBP1 jest istotnie podwyższona i poprzez wzmożone wiązanie się do traktów polipirymidynowych uniemożliwia działanie czynników U2AF. Prowadzić to może do zaburzeń w splicingu oraz aktywności poliadenylacji w PAS proks. Tym samym CPA zachodzić wówczas będzie w PAS dyst. i doprowadzi tym samym do zwiększenia poziomu miR-21. E1 – ekson 1; E2 – ekson 2; PR – punkt rozgałęzień; Py – trakt polipirymidynowy

Podsumowując można stwierdzić, że biogeneza miR-21 jest niezwykle złożonym procesem. Znajdujący się około 200 pz przed sekwencją kodującą miR-21 sygnał poliadenylacji jest bezpośrednio zaangażowany w efektywność wycinania pre-miR-21. Po zastosowaniu blokera RNA wykazano, że poprzez maskowanie proksymalnego PAS *VMP1* można istotnie wpłynąć na poziom dojrzałej cząsteczki miR-21. Tym samym udowodniono, że zachodzący proces cięcia i poliadenylacji w 3'UTR *VMP1* w bezpośredni sposób wpływa na powstawanie cząsteczki miR-21. Analiza powstających prekursorów miR-21 wykazała obecność dwóch izoform – z retencją intronu 11 oraz bez. Dodatkowo z przeprowadzonych badań wynika, że bardzo duża część powstających ze znajdującego się wewnątrz 10 intronu *VMP1* miejsca inicjacji transkrypcji RNA ulega poliadenylacji w proksymalnym PAS i stanowi pulę lncVMP1. Poziom akumulacji obydwu izoform *pri-miR-21*, w sytuacji obniżonego poziomu DROSHA,

okazał się być do siebie zbliżony. Można założyć, że w proces biogenezy miR-21 zaangażowane są obydwie izoformy. Obecność alternatywnej izoformy, z retencją intronu 11, może stanowić dowód na to, że podczas biogenezy miR-21 zajście alternatywnego splicingu odgrywa rolę. Silny związek pomiędzy procesem składania RNA oraz jego poliadenylacji, w przypadku dojrzewania prekursorów *pri-miR-21* może stanowić ciekawe zjawisko, które w przyszłości warto dogłębnie przebadать.

6. Wnioski

Na podstawie wyników badań uzyskanych w trakcie realizacji projektu doktorskiego sformułowano następujące wnioski:

1. Transkrypty *lncVMP1* oraz dłuższe *pri-miR-21* powstają niezależnie od transkrypcji swojego genu-gospodarza – *VMP1*.
2. Znacząca część transkryptów rozpoczynających się w miejscu inicjacji transkrypcji charakterystycznym dla *pri-miR-21* ulega terminacji przed sekwencją kodującą miR-21 i tworzy pulę długich, niekodujących RNA – *lncVMP1*.
3. Transkrypty *VMP1* nie są najprawdopodobniej cząsteczkami źródłowymi dla biogenezy miR-21.
4. Analizując dane RNA-seq, pochodzące z wielu ludzkich tkanek można stwierdzić, że poziom ekspresji *pri-miR-21* charakteryzuje się dużą zmiennością zarówno pod względem ilościowym jak i jakościowym.
5. Zidentyfikowano dwie izoforny transkryptów *lncVMP1* oraz *pri-miR-21*, różniące się obecnością sekwencji intronowej, znajdującej się pomiędzy eksonem 11 i 12 *VMP1*.
6. Terminacja transkrypcji w proksymalnym miejscu poliadenylacji (PAS proks.), zlokalizowanym około 200 nukleotydów od sekwencji kodującej miR-21, ma negatywny wpływ na efektywność biogenezy miR-21.
7. Brak miejsca splicingowego po stronie 3' (3'ss) intronu 11 *VMP1* prowadzi do retencji tego intronu i pozytywnie wpływa na ilość *pri-miR-21*, co prowadzi do wzrostu ilości dojrzałych cząsteczek miR-21.
8. Procesy splicingu i poliadenylacji *pri-miR-21* wzajemnie na siebie wpływają i stanowią potencjalne cele dla strategii terapeutycznych skierowanych na regulację poziomu miR-21.

7. Bibliografia

1. M. Brugiolo, L. Herzelt, K. M. Neugebauer, Counting on co-transcriptional splicing. *F1000Prime Rep.* **5**, 1–6 (2013).
2. Q. Pan, O. Shai, L. J. Lee, B. J. Frey, B. J. Blencowe, Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
3. M. C. Wahl, C. L. Will, R. Lührmann, The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell.* **136**, 701–718 (2009).
4. G. Dujardin, C. Lafaille, M. de la Mata, L. E. Marasco, M. J. Muñoz, C. Le Jossic-Corcós, L. Corcos, A. R. Kornblihtt, How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping. *Mol. Cell.* **54**, 683–690 (2014).
5. E. T. Wang, R. Sandberg, S. Luo, I. Khrebtkova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth, C. B. Burge, Alternative Isoform Regulation in Human Tissue Transcriptomes. *Science* **306**, 470–476 (2009).
6. M. Akerman, Y. Mandel-Gutfreund, Alternative splicing regulation at tandem 3' splice sites. *Nucleic Acids Res.* **34**, 23–31 (2006).
7. Z. Wang, C. B. Burge, Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA.* **14**, 802–813 (2008).
8. J. W. Park, K. Parisky, A. M. Celotto, R. A. Reenan, B. R. Graveley, Identification of alternative splicing regulators by RNA interference in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15974–15979 (2004).
9. O. V. Makarova, The 65 and 110 kDa SR-related proteins of the U4/U6U5 tri-snRNP are essential for the assembly of mature spliceosomes. *EMBO J.* **20**, 2553–2563 (2001).
10. N. Rooke, V. Markovtsov, E. Cagavi, D. L. Black, Roles for SR Proteins and hnRNP A1 in the Regulation of c-src Exon N1. *Mol. Cell. Biol.* **23**, 1874–1884 (2003).
11. B. Amir-Ahmady, P. L. Boutz, V. Markovtsov, M. L. Phillips, D. L. Black, Exon repression by polypyrimidine tract binding protein. *Rna.* **11**, 699–716 (2005).
12. C. R. Mandel, Y. Bai, L. Tong, Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci.* **65**, 1099–1122 (2008).
13. C. R. Eckmann, C. Rammelt, E. Wahle, Control of poly(A) tail length. *Wiley Interdiscip. Rev. RNA.* **2**, 348–361 (2011).

14. J. Zhao, L. Hyman, C. Moore, Formation of mRNA 3' Ends in Eukaryotes: Mechanism, Regulation, and Interrelationships with Other Steps in mRNA Synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405–445 (1999).
15. J. Rogers, P. Early, C. Carter, K. Calame, M. Bond, L. Hood, R. Wall, Two mRNAs with different 3' ends encode membrane-bound and secreted forms of immunoglobulin μ chain. *Cell.* **20**, 303–312 (1980).
16. D. R. Setzer, M. McGrogan, J. H. Nunberg, R. T. Schimke, Size heterogeneity in the 3' end of dihydrofolate reductase messenger RNAs in mouse cells. *Cell.* **22**, 361–370 (1980).
17. G. Edwalds-Gilbert, K. L. Veraldi, C. Milcarek, Alternative poly(A) site selection in complex transcription units: Means to an end? *Nucleic Acids Res.* **25**, 2547–2561 (1997).
18. W. Chen, Q. Jia, Y. Song, H. Fu, G. Wei, T. Ni, Alternative Polyadenylation: Methods, Findings, and Impacts Methods and Findings of Alternative Polyadenylation. *Genomics, Proteomics Bioinforma.* **15**, 287–300 (2017).
19. B. Tian, J. Hu, H. Zhang, C. S. Lutz, A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **33**, 201–212 (2005).
20. D. C. Di Giammartino, K. Nishida, J. L. Manley, Mechanisms and Consequences of Alternative Polyadenylation. *Mol. Cell.* **43**, 853–866 (2011).
21. H. Zhang, J. Y. Lee, B. Tian, Biased alternative polyadenylation in human tissues. *Genome Biol.* **6**, 1–13 (2005).
22. C. Mayr, D. P. Bartel, Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell.* **138**, 673–684 (2009).
23. P. Singh, T. L. Alley, S. M. Wright, S. Kamdar, W. Schott, R. Y. Wilpan, K. D. Mills, J. H. Graber, Global Changes in Processing of mRNA 3' Untranslated Regions Characterize Clinically Distinct Cancer Subtypes. *Cancer Res.* **69**, 9422–9430 (2009).
24. Y. Fu, Y. Sun, Y. Li, J. Li, X. Rao, C. Chen, A. Xu, Differential genome-wide profiling of tandem 3' UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res.* **21**, 741–747 (2011).
25. D. M. Jensen, Diagnosis and treatment of definitive diverticular hemorrhage (DDH). *Am. J. Gastroenterol.* **113**, 1570–1573 (2018).
26. T. Naganuma, S. Nakagawa, A. Tanigawa, Y. F. Sasaki, N. Goshima, T. Hirose, Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles.

- EMBO J.* **31**, 4020–4034 (2012).
27. S. Millevoi, S. Vagner, Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res.* **38**, 2757–2774 (2009).
 28. D. D. Licatalosi, A. Mele, J. J. Fak, J. Ule, M. Kayikci, S. Wook, T. A. Clark, A. C. Schweitzer, J. E. Blume, X. Wang, C. Jennifer, R. B. Darnell, HITS-CLIP insights into processing. *Nature.* **456**, 464–469 (2009).
 29. C. S. Lutz, K. G. K. Murthy, N. Schek, J. P. O'Connor, J. L. Manley, J. C. Alwine, Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev.* **10**, 325–337 (1996).
 30. A. Kyburz, A. Friedlein, H. Langen, W. Keller, Direct Interactions between Subunits of CPSF and the U2 snRNP Contribute to the Coupling of Pre-mRNA 3' End Processing and Splicing. *Mol. Cell.* **23**, 195–205 (2006).
 31. S. Millevoi, C. Loulergue, S. Dettwiler, S. Z. Karaa, W. Keller, M. Antoniou, S. Vagner, An interaction between U2AF 65 and CF Im links the splicing and 3' end processing machineries. *EMBO J.* **25**, 4854–4864 (2006).
 32. M. J. Dye, N. J. Proudfoot, Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Mol. Cell.* **3**, 371–378 (1999).
 33. M. Niwa, S. M. Berget, Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev.* **5**, 2086–2095 (1991).
 34. P. Kapranov, J. Cheng, S. Dike, D. A. Nix, R. Dutttagupta, A. T. Willingham, P. F. Stadler, J. Hertel, J. Hackermüller, I. L. Hofacker, I. Bell, E. Cheung, J. Drenkow, E. Dumais, S. Patel, G. Helt, M. Ganesh, S. Ghosh, A. Piccolboni, V. Sementchenko, H. Tammana, T. R. Gingeras, RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science (80-.).* **316**, 1484–1488 (2007).
 35. K. Struhl, Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* **14**, 103–105 (2007).
 36. K. C. Pang, M. C. Frith, J. S. Mattick, Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends Genet.* **22**, 1–5 (2006).
 37. T. R. Mercer, M. E. Dinger, J. S. Mattick, Long non-coding RNAs: Insights into functions. *Nat. Rev. Genet.* **10**, 155–159 (2009).
 38. J. Feng, C. Bi, B. S. Clark, R. Mady, P. Shah, J. D. Kohtz, The evf-2 noncoding RNA is

- transcribed from the Dlx-5/6 ultraconserved region and functions as a Idx2 coactivator. **1**, 1470–1484 (2006).
39. R. Lyle, D. Watanabe, D. Te Vrugte, W. Lerchner, O. W. Smrzka, A. Wutz, J. Schageman, L. Hahner, C. Davies, D. P. Barlow, The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat. Genet.* **25**, 19–21 (2000).
 40. E. Hacisuleyman, L. A. Goff, C. Trapnell, A. Williams, J. Henao-Mejia, L. Sun, P. Mcclanahan, D. G. Hendrickson, M. Sauvageau, D. R. Kelley, M. Morse, J. Engreitz, E. S. Lander, M. Guttman, H. F. Lodish, R. Flavell, A. Raj, J. L. Rinn, Topological Organization of Multi-chromosomal Regions by Firre HHS Public Access. *Nat Struct Mol Biol.* **21**, 198–206 (2014).
 41. M. C. Siomi, K. Sato, D. Pezic, A. A. Aravin, PIWI-interacting small RNAs: The vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* **12**, 246–258 (2011).
 42. S. Kuramochi-Miyagawa, T. Watanabe, K. Gotoh, Y. Totoki, A. Toyoda, M. Ikawa, N. Asada, K. Kojima, Y. Yamaguchi, T. W. Ijiri, K. Hata, E. Li, Y. Matsuda, T. Kimura, M. Okabe, Y. Sakaki, H. Sasaki, T. Nakano, DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* **22**, 908–917 (2008).
 43. R. Wilson, J. A. Doudna, Molecular mechanisms of RNA interference A BIOLOGICAL VIEW OF RNA INTERFERENCE • Small regulatory RNAs in cellular function and dysfunction HHS Public Access. *Annu. Rev. Biophys.* **42**, 217–239 (2013).
 44. A. J. Hamilton, D. C. Baulcombe, A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science (80-.)*. **286**, 950–952 (1999).
 45. S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, T. Tuschl, Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature.* **411**, 494–498 (2001).
 46. R. Feinbaum, V. Ambros, R. Lee, The C. elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to lin-14. *Cell.* **116**, 843–854 (2004).
 47. B. Wightman, T. R. Bürglin, J. Gatto, P. Arasu, G. Ruvkun, Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. *Genes Dev.* **5**, 1813–1824 (1991).
 48. B. Wightman, I. Ha, G. Ruvkun, Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell.* **75**, 855–862

- (1993).
49. R. C. Lee, V. Ambros, An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* (80-.). **294**, 862–864 (2001).
 50. D. P. Bartel, MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*. **116**, 281–297 (2004).
 51. T. Watanabe, T. Sato, T. Amano, Y. Kawamura, N. Kawamura, H. Kawaguchi, N. Yamashita, H. Kurihara, T. Nakaoka, Dnm3os, a non-coding RNA, is required for normal growth and skeletal development in mice. *Dev. Dyn.* **237**, 3738–3748 (2008).
 52. R. Sanuki, A. Onishi, C. Koike, R. Muramatsu, S. Watanabe, Y. Muranishi, S. Irie, S. Uneo, T. Koyasu, R. Matsui, Y. Chérasse, Y. Urade, D. Watanabe, M. Kondo, T. Yamashita, T. Furukawa, MiR-124a is required for hippocampal axogenesis and retinal cone survival through Lhx2 suppression. *Nat. Neurosci.* **14**, 1125–1136 (2011).
 53. S. Barbato, E. Marrocco, D. Intartaglia, M. Pizzo, S. Asteriti, F. Naso, D. Falanga, R. S. Bhat, N. Meola, A. Carissimo, M. Karali, H. M. Prosser, L. Cangiano, E. M. Surace, S. Banfi, I. Conte, MiR-211 is essential for adult cone photoreceptor maintenance and visual function. *Sci. Rep.* **7**, 1–13 (2017).
 54. K. Tsujimura, K. Irie, H. Nakashima, Y. Egashira, Y. Fukao, M. Fujiwara, M. Itoh, M. Uesaka, T. Imamura, Y. Nakahata, Y. Yamashita, T. Abe, S. Takamori, K. Nakashima, MiR-199a Links MeCP2 with mTOR Signaling and Its Dysregulation Leads to Rett Syndrome Phenotypes. *Cell Rep.* **12**, 1887–1901 (2015).
 55. N. Liu, S. Bezprozvannaya, A. H. Williams, X. Qi, J. A. Richardson, R. Bassel-Duby, E. N. Olson, microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev.* **22**, 3242–3254 (2008).
 56. N. Liu, A. H. Williams, J. M. Maxeiner, S. Bezprozvannaya, J. M. Shelton, J. A. Richardson, R. Bassel-Duby, E. N. Olson, MicroRNA-206 promotes skeletal muscle regeneration and delays progression of Duchenne muscular dystrophy in mice. *J. Clin. Invest.* **122**, 2054–2065 (2012).
 57. R. Song, P. Walentek, N. Sponer, A. Klimke, J. S. Lee, G. Dixon, R. Harland, Y. Wan, P. Lishko, M. Lize, M. Kessel, L. He, MiR-34/449 miRNAs are required for motile ciliogenesis by repressing cp110. *Nature*. **510**, 115–120 (2014).
 58. C. Henique, G. Bollée, X. Loyer, F. Grahammer, N. Dhaun, M. Camus, J. Vernerey, L. Guyonnet, F. Gaillard, H. Lazareth, C. Meyer, I. Bensaada, L. Legrès, T. Satoh, S. Akira, P.

- Bruneval, S. Dimmeler, A. Tedgui, A. Karras, E. Thervet, D. Nochy, T. B. Huber, L. Mesnard, O. Lenoir, P. L. Tharaux, Genetic and pharmacological inhibition of microRNA-92a maintains podocyte cell cycle quiescence and limits crescentic glomerulonephritis. *Nat. Commun.* **8** (2017), doi:10.1038/s41467-017-01885-7.
59. X. Chen, Y. Ling, Y. Wei, J. Tang, Y. Ren, B. Zhang, F. Jiang, H. Li, R. Wang, W. Wen, G. Lv, M. Wu, L. Chen, L. Li, H. Wang, Dual regulation of HMGB1 by combined JNK1/2-ATF2 axis with miR-200 family in nonalcoholic steatohepatitis in mice. *FASEB J.* **32**, 2722–2734 (2018).
 60. M. Latreille, P. Rorsman, M. Stoffel, M. Latreille, J. Hausser, I. Stützer, Q. Zhang, B. Hastoy, S. Gargani, J. Kerr-conte, F. Pattou, M. Zavolan, J. L. S. Esguerra, L. Eliasson, T. Rüllicke, P. Rorsman, M. Stoffel, MicroRNA-7a regulates pancreatic β cell function. *J. Clin. Invest.* **124**, 2722–2735 (2014).
 61. Y. Lee, M. Kim, J. Han, K. H. Yeom, S. Lee, S. H. Baek, V. N. Kim, MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051–4060 (2004).
 62. M. Ballarino, F. Pagano, E. Girardi, M. Morlando, D. Cacchiarelli, M. Marchioni, N. J. Proudfoot, I. Bozzoni, Coupled RNA Processing and Transcription of Intergenic Primary MicroRNAs. *Mol. Cell. Biol.* **29**, 5632–5638 (2009).
 63. W. Fang, D. P. Bartel, The Menu of Features that Define Primary MicroRNAs and Enable De Novo Design of MicroRNA Genes. *Mol. Cell.* **60**, 131–145 (2015).
 64. E. Lund, S. Güttinger, A. Calado, J. E. Dahlberg, U. Kutay, Nuclear Export of MicroRNA Precursors. *Science (80-.).* **303**, 95–98 (2004).
 65. R. Yi, Y. Qin, I. G. Macara, B. R. Cullen, Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011–3016 (2003).
 66. R. F. Ketting, S. E. J. Fischer, E. Bernstein, T. Sijen, G. J. Hannon, R. H. A. Plasterk, Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659 (2001).
 67. H. Kobayashi, Y. Tomari, RISC assembly: Coordination between small RNAs and Argonaute proteins. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859**, 71–81 (2016).
 68. D. P. Bartel, MicroRNAs: Target Recognition and Regulatory Functions. *Cell.* **136**, 215–233 (2009).
 69. Y. K. Kim, B. Kim, V. N. Kim, Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1881–E1889 (2016).

70. J. G. Ruby, C. H. Jan, D. P. Bartel, Intronic microRNA precursors that bypass Drosha processing. *Nature*. **448**, 83–86 (2007).
71. C. Lu, S. S. Tej, S. Luo, C. D. Haudenschild, B. C. Meyers, P. J. Green, Genetics: Elucidation of the small RNA component of the transcriptome. *Science (80-.)*. **309**, 1567–1569 (2005).
72. C. D. Malone, G. J. Hannon, Small RNAs as Guardians of the Genome. *Cell*. **136**, 656–668 (2009).
73. J. G. Ruby, C. Jan, C. Player, M. J. Axtell, W. Lee, C. Nusbaum, H. Ge, D. P. Bartel, Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in *C. elegans*. *Cell*. **127**, 1193–1207 (2006).
74. S. Griffiths-Jones, the microRNA sequence database. *Methods Mol Biol*. **342**, 129–138 (2006).
75. A. Kozomara, M. Birgaoanu, S. Griffiths-Jones, MiRBase: From microRNA sequences to function. *Nucleic Acids Res*. **47**, D155–D162 (2019).
76. J. Alles, T. Fehlmann, U. Fischer, C. Backes, V. Galata, M. Minet, M. Hart, M. Abu-Halima, F. A. Grässer, H. P. Lenhof, A. Keller, E. Meese, An estimate of the total number of true human miRNAs. *Nucleic Acids Res*. **47**, 3353–3364 (2019).
77. B. Fromm, T. Billipp, L. E. Peck, M. Johansen, J. E. Tarver, B. L. King, J. M. Newcomb, L. F. Sempere, K. Flatmark, E. Hovig, K. J. Peterson, A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annu. Rev. Genet*. **49**, 213–242 (2015).
78. C. Backes, T. Fehlmann, F. Kern, T. Kehl, H. P. Lenhof, E. Meese, A. Keller, MiRCarta: A central repository for collecting miRNA candidates. *Nucleic Acids Res*. **46**, D160–D167 (2018).
79. C. He, Z. Li, P. Chen, H. Huang, L. D. Hurst, J. Chen, Young intragenic miRNAs are less coexpressed with host genes than old ones: Implications of miRNA-host gene coevolution. *Nucleic Acids Res*. **40**, 4002–4012 (2012).
80. G. Georgakilas, I. S. Vlachos, M. D. Paraskevopoulou, P. Yang, Y. Zhang, A. N. Economides, A. G. Hatzigeorgiou, microTSS: accurate microRNA transcription start site identification reveals a significant number of divergent pri-miRNAs. *Nat. Commun*. **5**, 5700 (2014).
81. L. C. G. Hinske, P. A. F. Galante, W. P. Kuo, L. Ohno-Machado, A potential role for

- intragenic miRNAs on their hosts' interactome. *BMC Genomics*. **11** (2010), doi:10.1186/1471-2164-11-533.
82. J. Ribas, X. Ni, M. Castanares, M. M. Liu, D. Esopi, S. Yegnasubramanian, R. Rodriguez, J. T. Mendell, S. E. Lupold, A novel source for miR-21 expression through the alternative polyadenylation of VMP1 gene transcripts. *Nucleic Acids Res.* **40**, 6821–6833 (2012).
 83. A. C. Seila, J. M. Calabrese, S. S. Levine, G. W. Yeo, P. B. Rahl, R. A. Flynn, R. A. Young, P. A. Sharp, Divergent transcription from active promoters. *Science (80-.)*. **322**, 1849–1851 (2008).
 84. G. S. França, M. D. Vibranovski, P. A. F. Galante, Host gene constraints and genomic context impact the expression and evolution of human microRNAs. *Nat. Commun.* **7**, 11438 (2016).
 85. J. Béthune, C. G. Artus-Revel, W. Filipowicz, Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. *EMBO Rep.* **13**, 716–723 (2012).
 86. S. Djuranovic, A. Nahvi, R. Green, miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science (80-.)*. **336**, 237–240 (2012).
 87. R. C. Friedman, K. K. H. Farh, C. B. Burge, D. P. Bartel, Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105 (2009).
 88. E. Bernstein, S. Y. Kim, M. A. Carmell, E. P. Murchison, H. Alcorn, M. Z. Li, A. A. Mills, S. J. Elledge, K. V. Anderson, G. J. Hannon, Dicer is essential for mouse development. *Nat. Genet.* **35**, 215–217 (2003).
 89. A. M. Ardekani, M. M. Naeini, The role of microRNAs in human diseases. *Avicenna J. Med. Biotechnol.* **2**, 161–179 (2010).
 90. A. J. Giraldez, R. M. Cinalli, M. E. Glasner, A. J. Enright, J. M. Thomson, S. Baskerville, S. M. Hammond, D. P. Bartel, A. F. Schier, MicroRNAs regulate brain morphogenesis in zebrafish. *Science (80-.)*. **308**, 833–838 (2005).
 91. K. Goljanek-Whysall, D. Sweetman, A. M. Münsterberg, microRNAs in skeletal muscle differentiation and disease. *Clin. Sci.* **123**, 611–625 (2012).
 92. S. D. Hatfield, H. R. Shcherbata, K. A. Fischer, K. Nakahara, R. W. Carthew, H. Ruohola-Baker, Stem cell division is regulated by the microRNA pathway. *Nature*. **435**, 974–978 (2005).

93. G. A. Calin, C. Sevignani, C. D. Dumitru, T. Hyslop, E. Noch, S. Yendamuri, M. Shimizu, S. Rattan, F. Bullrich, M. Negrini, C. M. Croce, Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 2999–3004 (2004).
94. B. Zhang, X. Pan, G. P. Cobb, T. A. Anderson, microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* **302**, 1–12 (2007).
95. M. Garofalo, C. Quintavalle, G. Romano, C. M. Croce, G. Condorelli, miR221/222 in Cancer: Their Role in Tumor Progression and Response to Therapy. *Curr. Mol. Med.* **12**, 27–33 (2011).
96. I. K. Guttilla, B. A. White, Coordinate regulation of FOXO1 by miR-27a, miR-96, and miR-182 in breast cancer cells. *J. Biol. Chem.* **284**, 23204–23216 (2009).
97. I. A. Asangani, S. A. K. Rasheed, D. A. Nikolova, J. H. Leupold, N. H. Colburn, S. Post, H. Allgayer, MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene.* **27**, 2128–2136 (2008).
98. Z. Lu, M. Liu, V. Stribinskis, C. M. Klinge, K. S. Ramos, N. H. Colburn, Y. Li, MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene. *Oncogene.* **27**, 4373–4379 (2008).
99. S. Zhu, H. Wu, F. Wu, D. Nie, S. Sheng, Y. Y. Mo, MicroRNA-21 targets tumor suppressor genes in invasion and metastasis. *Cell Res.* **18**, 350–359 (2008).
100. M. Fuse, N. Nohata, S. Kojima, S. Sakamoto, T. Chiyomaru, K. Kawakami, H. Enokida, M. Nakagawa, Y. Naya, T. Ichikawa, N. Seki, Restoration of miR-145 expression suppresses cell proliferation, migration and invasion in prostate cancer by targeting FSCN1. *Int. J. Oncol.* **38**, 1093–1101 (2011).
101. R. Mizuno, K. Kawada, Y. Sakai, The Molecular Basis and Therapeutic Potential of Let-7 MicroRNAs against Colorectal Cancer. *Can. J. Gastroenterol. Hepatol.* **2018** (2018), doi:10.1155/2018/5769591.
102. A. De Cola, S. Volpe, M. C. Budani, M. Ferracin, R. Lattanzio, A. Turdo, D. D'Agostino, E. Capone, G. Stassi, M. Todaro, C. Di Ilio, G. Sala, M. Piantelli, M. Negrini, A. Veronese, V. De Laurenzi, MIR-205-5p-mediated downregulation of ERBB/HER receptors in breast cancer stem cells results in targeted therapy resistance. *Cell Death Dis.* **6** (2015), doi:10.1038/cddis.2015.192.

103. M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, Identification of novel genes coding for small expressed RNAs. *Science (80-.)*. **294**, 853–858 (2001).
104. P. P. Medina, M. Nolde, F. J. Slack, OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. *Nature*. **467**, 86–90 (2010).
105. F. Ozsolak, L. L. Poling, Z. Wang, H. Liu, X. S. Liu, R. G. Roeder, X. Zhang, J. S. Song, D. E. Fisher, Chromatin structure analyses identify miRNA promoters, 3172–3183 (2008).
106. S. Fujita, T. Ito, T. Mizutani, S. Minoguchi, N. Yamamichi, K. Sakurai, H. Iba, miR-21 Gene Expression Triggered by AP-1 Is Sustained through a Double-Negative Feedback Mechanism. *J. Mol. Biol.* **378**, 492–504 (2008).
107. G. Mudduluru, J. N. George-William, S. Muppala, I. A. Asangani, R. Kumarswamy, L. D. Nelson, H. Allgayer, Curcumin regulates miR-21 expression and inhibits invasion and metastasis in colorectal cancer. *Biosci. Rep.* **31**, 185–197 (2011).
108. J. M. Thomson, M. Newman, J. S. Parker, E. M. Morin-Kensicki, T. Wright, S. M. Hammond, Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev.* **20**, 2202–2207 (2006).
109. M. V. Iorio, R. Visone, G. Di Leva, V. Donati, F. Petrocca, P. Casalini, C. Taccioli, S. Volinia, C. G. Liu, H. Alder, G. A. Calin, S. Ménard, C. M. Croce, MicroRNA signatures in human ovarian cancer. *Cancer Res.* **67**, 8699–8707 (2007).
110. B. N. Davis, A. C. Hilyard, G. Lagna, A. Hata, SMAD proteins control DROSHA-mediated microRNA maturation. *Nature*. **454**, 56–61 (2008).
111. A. F. Schier, The maternal-zygotic transition: Death and birth of RNAs. *Science (80-.)*. **316**, 406–407 (2007).
112. P. Y. Chen, H. Manninga, K. Slanchev, M. Chien, J. J. Russo, J. Ju, R. Sheridan, B. John, D. S. Marks, D. Gaidatzis, C. Sander, M. Zavolan, T. Tuschl, The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev.* **19**, 1288–1293 (2005).
113. R. K. Ramachandra, M. Salem, S. Gahr, C. E. Rexroad, J. Yao, Cloning and characterization of microRNAs from rainbow trout (*Oncorhynchus mykiss*): Their expression during early embryonic development. *BMC Dev. Biol.* **8**, 1–11 (2008).
114. S. K. Singh, M. N. Kagalwala, J. Parker-Thornburg, H. Adams, S. Majumder, REST maintains self-renewal and pluripotency of embryonic stem cells. *Nature*. **453**, 223–227 (2008).

115. T. Hayashi, N. Koyama, Y. Azuma, M. Kashimata, Mesenchymal miR-21 regulates branching morphogenesis in murine submandibular gland in vitro. *Dev. Biol.* **352**, 299–307 (2011).
116. J. Lu, G. Getz, E. A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B. L. Ebert, R. H. Mak, A. A. Ferrando, J. R. Downing, T. Jacks, H. R. Horvitz, T. R. Golub, MicroRNA expression profiles classify human cancers. *Nature.* **435**, 834–838 (2005).
117. S. Volinia, G. A. Calin, C.-G. Liu, S. Ambs, A. Cimmino, F. Petrocca, R. Visone, M. Iorio, C. Roldo, M. Ferracin, R. L. Prueitt, N. Yanaihara, G. Lanza, A. Scarpa, A. Vecchione, M. Negrini, C. C. Harris, C. M. Croce, A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci.* **103**, 2257–2261 (2006).
118. J. A. Chan, A. M. Krichevsky, K. S. Kosik, MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res.* **65**, 6029–6033 (2005).
119. D. Löffler, K. Brocke-Heidrich, G. Pfeifer, C. Stocsits, J. Hackermüller, A. K. Kretzschmar, R. Burger, M. Gramatzki, C. Blumert, K. Bauer, H. Cvijic, A. K. Ullmann, P. F. Stadler, F. Horn, Interleukin-6-dependent survival of multiple myeloma cells involves the Stat3-mediated induction of microRNA-21 through a highly conserved enhancer. *Blood.* **110**, 1330–1333 (2007).
120. P. Wan, F. Zou, X. Zhang, H. Li, A. Dulak, J. Robert, T. Jr, J. S. Lazo, Z. Wang, L. Zhang, J. Yu, miR21 negatively regulates CDC25A and cell cycle progression in colon cancer cells. *Cancer Res.* **69**, 8157–8165 (2009).
121. J. Ren, T. H. Kuang, J. Chen, J. W. Yang, Y. X. Liu, The diagnostic and prognostic values of microRNA-21 in patients with gastric cancer: a meta-analysis. *Eur. Rev. Med. Pharmacol. Sci.* **21**, 120–130 (2017).
122. D. Sekar, R. Krishnan, K. Thirugnanasambantham, B. Rajasekaran, V. I. H. Islam, P. Sekar, Significance of microRNA 21 in gastric cancer. *Clin. Res. Hepatol. Gastroenterol.* **40**, 538–545 (2016).
123. P. Xu, Y. Zhu, B. Sun, Z. Xiao, Colorectal cancer characterization and therapeutic target prediction based on microRNA expression profile. *Sci. Rep.* **6**, 1–11 (2016).
124. K. Qu, T. Lin, Q. Pang, T. Liu, Z. Wang, M. Tai, F. Meng, J. Zhang, Y. Wan, P. Mao, X. Dong, C. Liu, W. Niu, S. Dong, Extracellular miRNA-21 as a novel biomarker in glioma: Evidence from meta-analysis, clinical validation and experimental investigations. *Oncotarget.* **7**, 33994–34010 (2016).

125. H. Wang, Z. Tan, H. Hu, H. Liu, T. Wu, C. Zheng, X. Wang, Z. Luo, J. Wang, S. Liu, Z. Lu, J. Tu, MicroRNA-21 promotes breast cancer proliferation and metastasis by targeting LZTFL1. *BMC Cancer*. **19**, 1–13 (2019).
126. I. Medimegh, I. Omrane, M. Privat, N. Uhrhammer, H. Ayari, F. Belaiba, F. Benayed, K. Benromdhan, S. Mader, I. J. Bignon, A. B. Elgaaied, MicroRNAs expression in triple negative vs non triple negative breast cancer in Tunisia: Interaction with clinical outcome. *PLoS One*. **9** (2014), doi:10.1371/journal.pone.0111877.
127. W. Zhao, J. J. Zhao, L. Zhang, Q. F. Xu, Y. M. Zhao, X. Y. Shi, A. G. Xu, Serum miR-21 level: A potential diagnostic and prognostic biomarker for non-small cell lung cancer. *Int. J. Clin. Exp. Med*. **8**, 14759–14763 (2015).
128. C. Bica-Pop, R. Cojocneanu-Petric, L. Magdo, L. Raduly, D. Gulei, I. Berindan-Neagoe, Overview upon miR-21 in lung cancer: focus on NSCLC. *Cell. Mol. Life Sci*. **75**, 3539–3551 (2018).
129. B. L. Jackson, A. Grabowska, H. L. Ratan, MicroRNA in prostate cancer: Functional importance and potential as circulating biomarkers. *BMC Cancer*. **14**, 1–10 (2014).
130. K. T. Gagnon, L. Li, B. A. Janowski, D. R. Corey, Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. *Nat. Protoc*. **9**, 2045–2060 (2014).
131. N. J. Dusetti, Y. Jiang, R. Tomasini, A. Azizi Samir, E. L. Calvo, G. V. Mallo, J. C. Dagorn, J. L. Iovanna, M. I. Vaccaro, A. Ropolo, F. Fiedler, Cloning and expression of the rat vacuole membrane protein 1 (VMP1), a new gene activated in pancreas with acute pancreatitis, which promotes vacuole formation. *Biochem. Biophys. Res. Commun*. **290**, 641–649 (2002).
132. A. Ropolo, D. Grasso, R. Pardo, M. L. Sacchetti, C. Archange, A. Lo Re, M. Seux, J. Nowak, C. D. Gonzalez, J. L. Iovanna, M. I. Vaccaro, The pancreatitis-induced vacuole membrane protein 1 triggers autophagy in mammalian cells. *J. Biol. Chem*. **282**, 37124–37133 (2007).
133. A. Amirfallah, A. Arason, H. Einarsson, E. T. Gudmundsdottir, E. S. Freysteinsdottir, K. A. Olafsdottir, O. T. Johannsson, B. A. Agnarsson, R. B. Barkardottir, I. Reynisdottir, High expression of the vacuole membrane protein 1 (VMP1) is a potential marker of poor prognosis in HER2 positive breast cancer. *PLoS One*. **14**, 1–19 (2019).
134. X. Cai, C. H. Hagedorn, B. R. Cullen, Human microRNAs are processed from capped,

- polyadenylated transcripts that can also function as mRNAs. *RNA*. **10**, 1957–1966 (2004).
135. R. Wang, R. Nambiar, D. Zheng, B. Tian, PolyA-DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res.* **46**, D315–D319 (2018).
 136. A. Tscherner, A. C. Brown, L. Stalker, J. Kao, I. Dufort, M. A. Sirard, J. LaMarre, STAT3 signaling stimulates miR-21 expression in bovine cumulus cells during in vitro oocyte maturation. *Sci. Rep.* **8**, 1–15 (2018).
 137. R. I. Gregory, K. P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch, R. Shiekhattar, The Microprocessor complex mediates the genesis of microRNAs. *Nature*. **432**, 235–240 (2004).
 138. M. Morlando, M. Ballarino, N. Gromak, F. Pagano, I. Bozzoni, N. J. Proudfoot, Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. & Mol. Biol.* **15**, 902–909 (2008).
 139. A. Louloui, E. Ntini, J. Liz, U. A. Ørom, Microprocessor dynamics shows co- and post-transcriptional processing of pri-miRNAs. *Rna*. **23**, 892–898 (2017).
 140. J. Rohwedel, V. Maltsev, E. Bober, H.-H. Arnold, J. Hescheler, A. M. Wobus, Muscle Cell Differentiation of Embryonic Stem Cells Reflects Myogenesis in Vivo: Developmentally Regulated Expression of Myogenic Determination Genes and Functional Expression of Ionic Currents. *Dev. Biol.* **164**, 87–101 (1994).
 141. P. K. Rao, R. M. Kumar, M. Farkhondeh, S. Baskerville, H. F. Lodish, Myogenic factors that regulate expression of muscle-specific microRNAs. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 8721–6 (2006).
 142. M. Horak, J. Novak, J. Bienertova-Vasku, Muscle-specific microRNAs in skeletal muscle development. *Dev. Biol.*, 1–13 (2015).
 143. G. F. Mok, D. Sweetman, Many routes to the same destination: Lessons from skeletal muscle development. *Reproduction*. **141**, 301–312 (2011).
 144. Y. H. Feng, C. J. Tsao, Emerging role of microRNA-21 in cancer (Review). *Biomed. Reports*. **5**, 395–402 (2016).
 145. C. Zhang, K. Liu, T. Li, J. Fang, Y. Ding, L. Sun, T. Tu, X. Jiang, S. Du, J. Hu, W. Zhu, H. Chen, X. Sun, MiR-21: A gene of dual regulation in breast cancer. *Int. J. Oncol.* **48**, 161–172 (2016).

146. A. V. Lee, S. Oesterreich, N. E. Davidson, MCF-7 Cells - Changing the Course of Breast Cancer Research and Care for 45 Years. *J. Natl. Cancer Inst.* **107**, 1–4 (2015).
147. A. J. Gruber, M. Zavolan, Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* (2019), doi:10.1038/s41576-019-0145-z.
148. C. R. Mandel, S. Kaneko, H. Zhang, D. Gebauer, V. Vethantham, J. L. Manley, L. Tong, Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature.* **444**, 953–956 (2006).
149. C. Alberti, L. Cochella, A framework for understanding the roles of miRNAs in animal development. *Dev.* (2017), , doi:10.1242/dev.146613.
150. D. Kaida, The reciprocal regulation between splicing and 3'-end processing. *Wiley Interdiscip. Rev. RNA.* **7**, 499–511 (2016).
151. M. Ha, V. N. Kim, Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 509–524 (2014).
152. Z. Krchňáková, P. K. Thakur, M. Krausová, N. Bieberstein, N. Haberman, M. Müller-McNicoll, D. Staněk, Splicing of long non-coding RNAs primarily depends on polypyrimidine tract and 5' splice-site sequences due to weak interactions with SR proteins. *Nucleic Acids Res.* **47**, 911–928 (2019).
153. A. Dhir, S. Dhir, N. J. Proudfoot, C. L. Jopling, Microprocessor mediates transcriptional termination in genes encoding long noncoding microRNAs. *Nat. Struct. Mol. Biol.* **22**, 319–27 (2015).
154. A. M. Krichevsky, G. Gabriely, miR-21: A small multi-faceted RNA. *J. Cell. Mol. Med.* **13**, 39–53 (2009).
155. C. Lv, W. X. Yu, Y. Wang, D. J. Yi, M. H. Zeng, H. M. Xiao, MiR-21 in extracellular vesicles contributes to the growth of fertilized eggs and embryo development in mice. *Biosci. Rep.* **38**, 1–11 (2018).
156. S. R. Pfeffer, C. H. Yang, L. M. Pfeffer, The Role of MIR-21 in Cancer. *Drug Dev. Res.* **76**, 270–277 (2015).
157. P. Du, L. Wang, P. Sliz, R. I. Gregory, A Biogenesis Step Upstream of Microprocessor Controls miR-17~92 Expression. *Cell.* **162**, 885–899 (2015).
158. B. Liu, Y. Shyr, J. Cai, Q. Liu, Interplay between miRNAs and host genes and their role in cancer. *Brief. Funct. Genomics.* **18**, 255–266 (2018).
159. P. NJ, Transcriptional termination in mammals: Stopping the RNA polymerase II

- juggernaut. *Science (80-.)*. **324**, 89–91 (2016).
160. J. D. Eaton, S. West, An end in sight? Xrn2 and transcriptional termination by RNA polymerase II. *Transcription*. **9**, 321–326 (2018).
 161. J. D. Eaton, L. Francis, L. Davidson, S. West, A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes Dev*. **34**, 132–145 (2020).
 162. R. Kole, A. R. Krainer, S. Altman, RNA therapeutics: Beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov*. **11**, 125–140 (2012).
 163. N. J. Proudfoot, A. Furger, M. J. Dye, Integrating mRNA processing with transcription. *Cell*. **108**, 501–512 (2002).
 164. B. Ko, S. I. Gunderson, Identification of new poly(A) polymerase-inhibitory proteins capable of regulating pre-mRNA polyadenylation. *J. Mol. Biol*. **318**, 1189–1206 (2002).
 165. R. Singh, J. Valcárcel, M. R. Green, Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science (80-.)*. **268**, 1173–1176 (1995).
 166. A. J. Gruber, R. Schmidt, S. Ghosh, G. Martin, A. R. Gruber, E. van Nimwegen, M. Zavolan, Discovery of physiological and cancer-related regulators of 3' UTR processing with KAPAC. *Genome Biol*. **19**, 44 (2018).
 167. X. He, A. D. Arslan, T. T. Ho, C. Yuan, M. R. Stampfer, W. T. Beck, Involvement of polypyrimidine tract-binding protein (PTBP1) in maintaining breast cancer cell growth and malignant properties. *Oncogenesis*. **3**, 1–8 (2014).

8. Spis rycin i tabel

Rycina 1 Podział RNA występujący w komórkach.....	15
Rycina 2 Schemat składania RNA z zaznaczeniem zaangażowanych snRNP przeprowadzających cały proces.....	19
Rycina 3 Schematyczne przedstawienie możliwych procesów alternatywnego składania RNA.....	21
Rycina 4 Białka regulujące proces składania RNA, w tym alternatywny splicing.	22
Rycina 5 Kompleksy białkowe zaangażowane w proces poliadenylacji.	23
Rycina 6 Schematyczne przedstawienie możliwych procesów alternatywnej poliadenylacji.....	24
Rycina 7 Schematyczne przedstawienie procesu biogenezy miRNA	30
Rycina 8 Genomowy podział cząsteczek miRNA.	35
Rycina 9 Schematyczne przedstawienie konstruktów genetycznych GFP_ <i>pri-miR-21</i> wykorzystywanych w rozprawie.....	44
Rycina 10 Schemat procedury odwrotnej transkrypcji otrzymania cDNA z krótkich RNA jako matrycy	49
Rycina 11 Schemat prezentujący lokalizacje starterów do badania poziomu zmian VMP1/ <i>pri-miR-21</i>	51
Rycina 12 Poziom ekspresji VMP1 i <i>pri-miR-21</i> w ludzkich tkankach.....	56
Rycina 13 Poziom ekspresji różnej długości izoform genu VMP1 w panelu ludzkich tkanek.	58
Rycina 14 Schemat izoform RNA pochodzących z VMP1 włączając prekursor <i>pri-miR-21</i>	59
Rycina 15 Poziom ekspresji VMP1, IncVMP1 oraz <i>pri-miR-21</i> w wybranych tkankach ludzkich.....	62
Rycina 16 Komórki z obniżonym poziomem białka DROSHA jako model badań zmian poziomu prekursorów miRNA.....	64

Rycina 17 Zaangażowanie proksymalnego miejsca poliadenylacji VMP1/ <i>pri-miR-21</i> w biogenezę miR-21.....	66
Rycina 18 Względny poziom transkryptów VMP1 oraz <i>pri-miR-21</i> we frakcji jądrowej.....	68
Rycina 19 Zmiany ekspresji VMP1, <i>pri-miR-21</i> oraz zmiany poziomu miR-21 podczas różnicowania ludzkich mioblastów.....	69
Rycina 20 Bezpośredni wpływ obniżenia poziomu białka DROSHA na ekspresję VMP1 i <i>pri-miR-21</i> ..	71
Rycina 21 Ekspresja VMP1 na poziomie jego mRNA i <i>pri-miR-21</i> oraz dojrzałego miR-21 w komórkach raka piersi.	73
Rycina 22 Wpływ wyciszenia CPSF3 na poziom ekspresji VMP1, <i>pri-miR-21</i> oraz cząsteczek miRNA.	77
Rycina 23 Badanie konstruktów genetycznych GFP_ <i>pri-miR-21</i>	78
Rycina 24 Wpływ blokowania i mutacji regionu PAS proks na biogenezę miR-21.	80
Rycina 25 Retencja ostatniego intronu pre-mRNA VMP1 oraz <i>pri-miR-21</i> po zablokowaniu i mutacji PAS proks.	83
Rycina 26 Obecność izoforny <i>pri-miR-21/lncVMP1</i> z retencją intronu w różnych modelach komórkowych (A)	84
Rycina 27 Wpływ zastosowania blokowania miejsca 3'ss w ostatnim intronie VMP1/ <i>pri-miR-21</i> oraz mutacji 3'ss w konstrukcie genetycznym GFP_ <i>pri21_Δ3'ss</i> na biogenezę miR-21.	86
Rycina 28 Schemat prezentujący transkrypty lncVMP1, VMP1 oraz <i>pri-miR-21</i>	92
Rycina 29 Terminacja transkrypcji oraz model „torpedo” w odniesieniu do biogenezy miR-21.	94
Rycina 30 Schematyczna prezentacja potencjalnego działania czynników U2AF na biogenezę miR-21.	96
Rycina 31 Model potencjalnego wpływu białka PTBP1 na biogenezę miR-21 w prawidłowych (MCF-12A) oraz nowotworowych zmienionych komórkach raka piersi (MCF-7).	98

Tabela 1 Znaczenie miR-21 w różnych typach nowotworów	41
Tabela 2 Spis konstruktów używanych w pracy doktorskiej.	43
Tabela 3 Linie komórkowe używane w pracy doktorskiej	46
Tabela 4 Antysensowne oligonukleotydy (AON) stosowane w pracy doktorskiej	48
Tabela 5 Dupleksy siRNA stosowane w pracy doktorskiej.....	48
Tabela 6 Wykorzystywane w reakcjach RT-PCR oraz RT-qPCR startery	53