Instytut Chemii Bioorganicznej
Polskiej Akademii Nauk

Charakterystyka i funkcja niekodujących RNA uczestniczących w rozwoju komórek nerek oraz ich karcinogenezie

Marek Kazimierczyk

Praca doktorska wykonana w **Zakładzie Biochemii RNA**
pod kierunkiem: **dr hab. Jana Wrzesińskiego, prof. ICHB PAN**

Poznań 2023

Składam najszczersze podziękowania:

Mojemu promotorowi dr hab. Janowi Wrzesińskiemu

Za opiekę merytoryczną, poświęcony czas i cierpliwość w pracy, a także za szansę by móc realizować swoją pasję jaką jest biologia

Mojemu zespołowi w osobach ***mgr inż. Martyny Kabacińskiej, dr Pauliny Machtel, dr Damiana Janeckiego, dr hab. Agaty Świątkowskiej, dr hab. Marioli Dutkiewicz, prof. dr hab. Jerzego Ciesiołki****.*

Za nieocenione wsparcie i pomoc w stawianiu pierwszych i kolejnych kroków naukowej kariery.

Kolegom i Koleżankom z ICHB PAN

Za dobre słowa, których padła niezliczona ilość.

Przyjaciołom Filipowi i Markowi

Za wsparcie w najtrudniejszych momentach.

Kochanym Rodzicom

Za stworzenie warunków do realizacji swoich pasji.

Mojej ukochanej Partnerce Kindze i córce Anastazji

Za mobilizację do pracy, wsparcie psychiczne i radość z powrotów do domu.

Swoją pracę dedykuję:

Mojej nauczycielce biologii w XII LO im. Marii Skłodowskiej-Curie ***dr Albinie Domce-Popek*** *za zaszczepienie we mnie ciekawości do świata i pasji do biologii.*

„*Nie człowiek pisze książki, lecz książki wykorzystują ludzi by zostać napisanymi*”

- Jacek Dukaj ‘Po piśmie’

Spis treści

1 Wykaz prac naukowych wchodzących w skład pracy doktorskiej

tRNA-derived fragments from the Sus scrofa tissues provide evidence of their conserved role in mammalian development

Marek Kazimierczyk, Agata Jędroszkowiak, Dorota Kowalczykiewicz, Maciej Szymański, Barbara Imiołczyk, Jerzy Ciesiołka, Jan Wrzesiński

Biochemical and Biophysical Research Communications.

2019 Dec 10;520(3):514-519. doi: 10.1016/j.bbrc.2019.10.062.

IF = 2.985, MNiSW = 100

Characteristics of Transfer RNA-Derived Fragments Expressed during Human Renal Cell Development: The Role of Dicer in tRF Biogenesis

Marek Kazimierczyk, Marta Wojnicka, Ewa Biała, Paulina Żydowicz-Machtel, Barbara Imiołczyk, Tomasz Ostrowski, Anna Kurzyńska-Kokorniak, Jan Wrzesinski

International Journal of Molecular Sciences.

2022 Mar 26;23(7):3644. doi: 10.3390/ijms23073644.

IF = 5.6, MNiSW = 140

Human Long Noncoding RNA Interactome: Detection, Characterization and Function

Marek Kazimierczyk, Marta K. Kasprowicz, Marta E. Kasprzyk, and Jan Wrzesinski

International Journal of Molecular Sciences

2020 Feb; 21(3): 1027. Published online 2020 Feb 4. doi: 10.3390/ijms21031027

IF = 5.924, MNiSW = 140

Long Non-Coding RNA Epigenetics

Marek Kazimierczyk, Jan Wrzesinski

International Journal of Molecular Sciences

2021 Jun 7;22(11):6166. doi: 10.3390/ijms22116166.

IF = 6.208, MNiSW = 140

2 Streszczenie

Badania genomowe ujawniły, że 80% genomu ulega transkrypcji, jednak tylko 2% koduje białka. Pozostała część transkryptomu do niedawna uznawana była za nieistotną, zwaną nawet śmieciowym RNA. W miarę postępów badań i odkryć kolejne frakcje nieistotnego RNA były systematyzowane i klasyfikowane pod względem funkcji i struktury. Niekodujące RNA to bardzo zróżnicowana frakcja RNA, których sekwencja nie koduje informacji o białku. Pośród niekodujących RNA wyróżniamy dwie grupy zależnie od funkcji RNA - porządkowe oraz regulatorowe. Porządkowe RNA pełnią funkcję strukturalną dla podstawowych procesów komórkowych i wyróżniamy pośród nich: tRNA, rRNA, snRNA, snoRNA. Regulatorowe RNA wpływają na procesy komórkowe i dzielą się na dwie grupy pod względem długości. Wyróżniamy małe niekodujące RNA (snRNA) oraz długie niekodujące RNA (lncRNA), a granicą między nimi jest, wyznaczona arbitralnie, długość 200 nt. Podgrupa snRNA różnicuje się również ze względu na długość na siRNA, miRNA, piRNA oraz cząsteczki tRF. Pośród lncRNA można wyszczególnić ostatnio odkrytą frakcję cyrkularnych RNA (circRNA). Zmiany w występowaniu regulatorowych RNA wpływają na procesy komórkowe, takie jak nowotworzenie czy różnicowanie. Poziom istotności regulatorowych RNA jest tak duży, że nawet zmiany w modyfikacjach pojedynczych nukleotydów są w stanie generować stany patologiczne.

W niniejszym zbiorze publikacji, wraz ze współautorami, dowodzimy o istotnym wpływie występowania małych niekodujących RNA na funkcję oraz na stan patologiczny, badając profile cząsteczek tRF w dwóch układach biologicznych. Pierwszym są różne tkanki pochodzące od zwierzęcia modelowego jakim jest *Sus scrofa (*świnia domowa). Drugim jest zaproponowany przez nas model komórkowy rozwoju ludzkiej nerki zawierający cztery linie komórkowe obrazujące różnicowanie komórek nerek oraz ich karcinogenezę. W celu poznania profilu cząsteczek tRF wykorzystujemy sekwencjonowanie nowej generacji RNAseq i metody obrazowania na membranach northern blot. W celu wykazania funkcji użyto wykorzystujemy test elektroforetycznej zróżnicowanej mobilności (EMSA) oraz analizy baz danych z wykorzystaniem narzędzi bioinformatycznych, takich jak tRFTars. W pracach przeglądowych kompletujemy najnowszą wiedzę dotyczącą interakcji oraz modyfikacji lncRNA wykazując mechanizmy wpływania na różnicowanie komórek oraz karcinogenezę.

3 Abstract

Genomic research has revealed that 80% of the genome undergoes transcription, but only 2% encodes proteins. The remaining portion of the transcriptome was long considered unimportant, and even referred to as "junk" RNA. As research has progressed and discoveries have been made, further fractions of non-coding RNA have been systematized and classified based on their function and structure. Non-coding RNA is a highly diverse fraction of RNA whose sequences do not encode protein information. Among non-coding RNAs, two groups are distinguished based on RNA function: structural and regulatory. Structural RNA including tRNA, rRNA, snRNA, and snoRNA. serves structural/play important roles in fundamental cellular processes. Regulatory RNAs influence cellular processes and are divided into two groups based on their length. These regulatory RNAs can be broadly divided into two main groups based on their length: small regulatory RNAs (snRNA) and long non-coding RNAs (lncRNAs). with the boundary between them being arbitrarily set at a length of 200 nt. The snRNA subgroup also differentiates based on length into siRNA, miRNA, piRNA, and tRF molecules, while among lncRNA, a fraction of circular RNAs (circRNA) can be distinguished. Changes in the presence of regulatory RNAs can affect cellular processes such as differentation and in consequence can lead to carcinogenesis. The significance of regulatory RNAs is so substantial that even modifications to individual nucleotides can lead to pathological conditions.

In this collection of publications, the significant impact of non-coding RNAs on function and pathological states has been demonstrated by examining the profiles of tRNA-derived fragments (tRFs) in two biological systems. The first involves various/several tissues from the model organism Sus scrofa (pig). The second is a cell model of human kidney development proposed by us, consisting of four cell lines illustrating kidney cell differentiation and carcinogenesis. To determine the tRF molecule profile, we employed next-generation RNAseq sequencing and membrane-based northern blot imaging techniques. To demonstrate function of ncRNA, we used Electrophoretic Mobility Shift Assay (EMSA) and bioinformatics tools like tRFTars by analyzing databases. Finnaly, in comprehensive review articles, we summarized the latest knowledge regarding the interactions and modifications of lncRNAs, elucidating the mechanisms influencing cell differentiation and carcinogenesis.

4 Wstęp

4.1 Od marginalizacji do cząsteczek funkcjonalnych, rys historyczny niekodujących RNA

Historia badań nad niekodującymi RNA to historia wychodzenia z cienia na pierwszy plan. Klasa niekodujących RNA przeszła drogę od nieistotnych transkryptów przez intrygujące cząsteczki po potencjalne terapeutyki. Korzenie historii ncRNA sięgają izolacji pierwszej nukleiny przez Fiedricha Mieschera [1]. Jego procedurę ulepszył Richard Altmann w 1889 roku oczyszczając nukleinę z białek i otrzymując kwas nukleinowy [2]. W miarę prowadzenia badań zaczął wyłaniać się obraz dwóch kwasów nukleinowych. Rozróżnienie DNA i RNA nastąpiło stopniowo. W 1909 roku Phoebus Levene badając kwas nukleinowy drożdży ustalił cukier wchodzący w jego skład, a mianowicie D-rybozę [3]. Levene odkrył podstawową różnicę - występowanie uracylu w roślinnym kwasie nukleinowym, zaś tyminy w zwierzęcym, przez co uważano, że istnieje specyficzny podział kwasów nukleinowych na królestwa zwierząt i roślin [4]. Ten sam badacz 20 lat później odkrył, że cukrem wchodzącym w skład zwierzęcych kwasów nukleinowych jest 2-deoksy-D-ryboza [5]. Od tamtego momentu w literaturze używane są nazwy: DNA – kwas deoksyrybonukleinowy i RNA – kwas rybonukleinowy. Pogląd przynależności RNA do królestwa roślin, a DNA do królestwa zwierząt trwał aż do roku 1939. W tym właśnie roku Jean Brachet wykazał, że zarówno DNA jak i RNA znajdują się w komórkach roślinnych i zwierzęcych [6]. Było to przełamanie jednego z wielu błędnych dogmatów. Przed 1961 rokiem, nim wykazano, że to mRNA jest pośrednim nośnikiem informacji, wielu badaczy twierdziło, że RNA ma niewielkie znaczenie [7,8]. Jako główną funkcję RNA wskazywano stabilizację białek rybosomalnych i działanie jako jednostki metaboliczne lub strukturalne. W 1969 roku Roy Britten i Eric Davidson zasugerowali istnienie RNA o znaczeniu regulacyjnym, co można uznać za poszukiwania pierwszych RNA, które nie tylko kodują białko, ale i regulują inne geny. Na tym etapie badacze spotkali się z oporem środowiska naukowego. Mimo sceptycyzmu nad regulacją RNA przez inne RNA, prace nadal trwały i w 1989 roku grupa badawcza Dalihasa opublikowała wyniki badań dowodzących, że niekodujący micF osłabia ekspresję genu ompF w komórkach *E. Coli* [10]. Jedną z kluczowych prac były badania nad ekspresją genu lin-14. Zauważono, że produkt genu lin-4 celuje w dotychczas uważany za nieistotny region między kodonem terminacyjnym a łańcuchem poli(A) mRNA lin-14, co powoduje represję docelowego mRNA [11]. Na bazie tej pracy uważa się, że lin-4 jest jednym z założycieli rodziny miRNA. W kolejnych pracach naukowcy odkrywali następne ncRNA o potencjale regulacyjnym. Kolejnym przełomem były wyniki sekwencjonowania DNA przez konsorcjum ENCODE. Naukowcy wykazali, że co najmniej 80%

genomowego DNA ssaków podlega aktywnej transkrypcji i szczegółowej regulacji, jednak mniej niż 3% genomu stanowią rejony kodujące białka [12, 13]. Wywołało to burzliwą dyskusję nad jakością prac i sekwencjonowaniem „śmieci". Dziś już wiemy, że to kolejna lekcja pokory i wiele z tych śmieci to w rzeczywistości regiony zawierające geny ncRNA.

4.2 Uniwersum niekodujących RNA

Niekodujące RNA to olbrzymia rodzina cząsteczek, które możemy podzielić ze względu na funkcję - RNA strukturalne i regulatorowe. Pośród strukturalnych wyróżniamy: transferowy (tRNA), rybosomalny (rRNA), mały jądrowy (snRNA) i mały jąderkowy (snoRNA). Regulatorowe RNA dzielimy ze względu na wielkość, przy przyjęciu arbitralnej długości cząsteczki 200 nukleotydów (nt), na małe niekodujące RNA (sncRNA) oraz na długie niekodujące RNA (lncRNA). Regulatorowe sncRNA wewnętrznie dzielą się na: siRNA, miRNA, piRNA oraz cząsteczki tRF. Długie niekodujące RNA są trudniejsze do klasyfikacji, ponieważ ich długość wynosi od kilkuset do kilku tysięcy nt. Główny podział lncRNA dotyczy miejsca ich lokalizacji genomowej [14, 15]. Nadal w środowisku naukowym podejmowana jest dyskusja, w jakim stopniu ncRNA są rzeczywiście funkcjonalne, a na ile są to jedynie produkty przypadkowej transkrypcji generujące szum informacyjny. Jeśli chodzi o funkcjonalność, zdecydowanie lepiej opisane są krótkie sncRNA, a w szczególności miRNA [16]. Może to być spowodowane dobrze opanowanymi technikami sekwencjonowania opartymi o krótkie odczyty, które swoją długością idealnie odpowiadają sncRNA, jednak ich użyteczność w przypadku analizy lncRNA znacznie spada. Nowe techniki sekwencjonowania, takie jak PackBio czy Nanopore budzą duże nadzieje, jeśli chodzi o analizę lncRNA.

Główną wątpliwości odnośnie funkcjonalności odkrywanych transkryptów budzi fakt, że jako grupa cząsteczek RNA utrzymują się w niskim stężeniu w komórce. Prawdopodobnie ich ekspresja i funkcja są wysoce specyficzne tkankowo [17]. Narastająca dyskusja wokół ncRNA skłoniła naukowców do redefinicji funkcjonalności transkryptu [18]. Jak piszą autorzy wspomnianego artykułu, funkcjonalności genu nie możemy określać poprzez mutację danego genu i zmiany w fenotypie. Dzieje się tak, ponieważ występuje zjawisko redundancji genetycznej, które powoduje przejęcie części funkcji jednego genu przez inny. Inny powód, dla którego określanie funkcjonalności genu jest trudne, to hermetyczne warunki laboratoryjne. W naturze na organizm działa wiele czynników, które mogą wpływać na ekspresję genów, a których ekspresji w warunkach „idealnych'' możemy nie odnotować. Posługując się przykładem drożdży hodowanych w warunkach standardowych, jedynie około 1/3 genów jest potrzebna do prawidłowego wzrostu [19, 20]. W przypadku hodowli ludzkich komórek w warunkach laboratoryjnych jest to jedynie około 10% [21]. Biorąc pod uwagę te wszystkie argumenty

wydaje się rozsądnym zakładać, że ncRNA są w większości funkcjonalne, nawet jeśli pojedynczy transkrypt nie wykazuje jednoznacznej funkcji.

4.2.1 sncRNA - małe cząsteczki o dużym znaczeniu

Małe niekodujące RNA to najlepiej opisana grupa niekodujących regulatorowych RNA. Ich przewaga w analizie opiera się o większą łatwość w identyfikacji przy wykorzystaniu sekwencjonowania RNA opartego o krótkie odczyty. Sekwencjonowanie RNA nowej generacji umożliwia odczyty do długości nawet 300 nt [22], co pokrywa górną granicę jaką definiujemy sncRNA, a otwiera dolną granicę lncRNA. Obecność małych transkryptów możemy odnotować w komórkach eukariotycznych, archeonach i bakteriach [23]. Powszechność i względna łatwość w analizie spowodowały, że przez ostatnie dekady małe niekodujące RNA były bardzo intensywnie badane. Ogromnym sukcesem było odkrycie miRNA, a później siRNA, których właściwości i funkcję biologiczną wykorzystujemy dziś jako narzędzia biologii molekularnej. W kolejnych latach grono małych niekodujących transkryptów powiększyło się o RNA oddziałujące z białkami PIWI (piRNA) oraz RNA pochodzące z tRNA (tRFy). W kontekście leków RNA bardzo duże nadzieje wiązane są z sncRNA.

W 2018 roku FDA zatwierdziło pierwszą terapię RNAi, czyli wykorzystującą interferencyjnie działanie małych niekodujących RNA. Lek opiera się o cząsteczkę siRNA, która ma za zadanie ograniczyć ilość mRNA nieprawidłowo wytwarzanego białka transtyretyny, co łagodzi objawy choroby [24, 25]. Na początku 2023 roku dostępne były 4 leki zatwierdzone przez FDA oparte o RNAi. W III fazie badań klinicznych znajduje się 5 kolejnych terapii czekających na zatwierdzenie, a wiele kolejnych, nad którymi prowadzone są badania, jest na wcześniejszych etapach [26]. Potencjalne leki oparte o regulatorowe RNA mają szereg zalet [25]. Małe interferujące RNA otwierają możliwość docierania do podstaw choroby. Wiele chorób wynika z wadliwej informacji zakodowanej w DNA i przepisanej na RNA, i nie ma innego skutecznego sposobu na zahamowanie eskalacji błędu. Odpowiednie przygotowanie cząsteczki, poprzez szereg modyfikacji łańcucha RNA, pozwala wydłużyć czas działania leku wobec standardowych leków małocząsteczkowych. Stosowanie leków opartych o regulatorowe RNA budzi nadzieję dla wielu pacjentów cierpiących z powodu chorób rzadkich, na które brak jest skutecznych leków [27]. Zastosowania dwóch pierwszych odkrytych małych niekodujących RNA są szeroko badane, a podobny potencjał upatruje się w cząsteczkach piRNA oraz tRF. Świat małych RNA nie zostaje całkowicie poznanym i stale się rozszerza.

Ostatnie odniesienia naukowe zwracają uwagę na rosnącą grupę niekanonicznych sncRNA [28]. Te umykające klasyfikacji sncRNA pochodzą najczęściej z grupy strukturalnych sncRNA i osiągają długość od 15 do 50 nt. Znaczenie tych cząsteczek coraz częściej dostrzega

się w związku z występowaniem wielu jednostek chorobowych. Warto zauważyć, że płaszczyzny działania sncRNA często się pokrywają, dlatego ich dobra charakterystyka jest niezwykle ważna, by zauważyć i wyodrębnić różnice.

4.2.1.1 siRNA – małe interferujące RNA

Najmniejsze z regulatorowych sncRNA zawierają od 20 do 24 nt. Źródłem pierwotnym siRNA w komórce są egzogenne potencjalnie niepożądane transkrypty uruchamiające odpowiedź komórkową w postaci szlaku RNAi. System odpornościowy działa w oparciu o podwójną komplementarną nić RNA (dsRNA). W momencie, gdy dsRNA zostanie zidentyfikowane w komórce, zostaje pocięte przez rybonukleazę Dicer na mniejsze fragmenty nazywane siRNA. Charakterystycznym elementem siRNA są dwunukleotydowe niesparowane wystające końce 3'. Obecność siRNA zapoczątkowuje formowanie się kompleksu RISC (*RNA-induced silencing complex*). W skład kompleksu RISC wchodzi wiele białek, a kompletny skład kompleksu nadal nie jest rozpoznany. Kompleks RISC, zależnie od adenozyno-5′-trifosforanu (ATP), aktywnością helikazy rozplata duplex siRNA. Następnie nić wiodąca komplementarna do celowanego RNA wchodzi w skład RISC, czyniąc go aktywnym. Druga nić, zwana pasażerską, ulega degradacji. Główną rolę w kompleksie RISC odgrywa białko z rodziny Argonaute. Białko zawiera domenę PAZ wiążącą RNA oraz domenę PIWI o aktywności nukleazy. Kompleks, wykorzystując komplementarność załadowanej nici siRNA, łączy się z celem i za pośrednictwem aktywności nukleazowej tnie docelowe RNA w rejonie dwuniciowym siRNA-targetRNA. Proces biogenezy siRNA został przedstawiony na Rycinie 1.

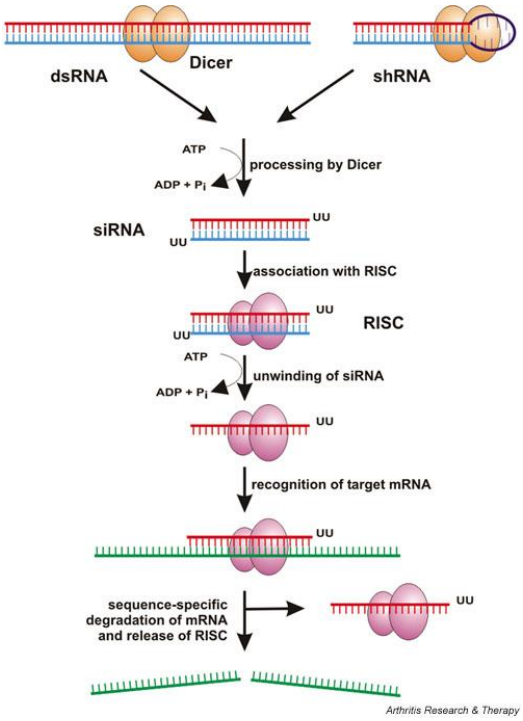


Figura 1. Mechanizmy biogenezy i działania siRNA [29].

siRNA obecnie wykorzystywane są w roli sztucznie transfekowanych cząsteczek w celu zaindukowania mechanizmu RNAi przeciw wybranym mRNA. Narzędzie RNAi jest wysoce skuteczne, o czym świadczą wcześniej wspomniane leki oparte o ten szlak. Technika RNAi nie jest pozbawiona wad. Główną wadą jest zjawisko *off-target*. Biorąc pod uwagę biologiczną obfitość transkryptów i sekwencji, istnieje prawdopodobieństwo rozpoznania przez siRNA innego transkryptu niż docelowy. Koniecznym jest wnikliwa analiza komplementarności projektowanego siRNA, zarówno do docelowego transkryptu, jak i nawet niepełna komplementarność do jakiegokolwiek transkrytpu, który może znaleźć się w badanej komórce. Nawet niepełna komplementarność jest w stanie dać efekt *off-target* [30]. Nie wszystkie efekty da się rozpoznać *in silico,* dlatego skuteczne wyeliminowanie celowanego mRNA wymaga wielu testów. Efekt wprowadzenia cząsteczki dsRNA może nie być obojętny dla organizmu. Sztucznie wprowadzone cząsteczki dsRNA jako substraty dla białka Dicer, jeśli ich długość jest większa niż 30pz, mogą indukować odpowiedź zapalną zależną od interferonu, prowadzącą do nieswoistej inhibicji translacji białek. Za taką reakcję odpowiada zdolność do rozpoznawania dsRNA, o wspomnianej długości, przez kinazę białkową zależną od RNA [31]. Rozwiązaniem kwestii *off-target* oraz nadmiernej reakcji komórkowej są najczęściej modyfikacje wprowadzane do cząsteczek dsRNA. Jednym z rodzajów modyfikacji są modyfikacje łańcucha cukrowego, gdzie wyróżniamy: 2'fluoro_RNA, 2'-O-metylo RNA, 2'-O-methoxyethyl, 4'-tio-RNA, LNA, UNA [32]. Właściwości siRNA możemy modulować również przez modyfikacje szkieletu cząsteczki, takie jak dołączanie grup fosforotionianowych, bromofosforowych czy fosforoacetonowych. Modyfikacji ulegać mogą również same zasady azotowe, ale efektywne jest wykorzystywanie modyfikacji uracylu na 4-tiouracyl, 2-tiouracyl czy pseudouracyl. Wszystkie zmiany i dodatkowe grupy mają na celu poprawić wydajność działania maszynerii RNAi wydłużając czas trwania duplexu, nie indukując nadmiernej odpowiedzi komórkowej lub ograniczając efekt *off-target*. Skuteczne zastosowanie RNAi na hodowlach komórkowych jest dużo prostsze, ponieważ jedyną barierą, jaką napotyka cząsteczka, jest bariera komórkowa. Wejście cząsteczki do komórki odbywa się głównie na drodze endocytozy, która wspomagana jest laboratoryjnie odczynnikami do transfekcji [33, 34].

W przypadku badań na organizmach pojawia się problem dostarczenia odpowiednio przygotowanej cząsteczki do docelowego miejsca działania, która musi pokonać szereg barier [35]. Wymaga się od cząsteczki działania punktowego, czasem w ściśle określonych komórkach badanego organu. Pomocne w tym celu są biokoniugaty, czyli kombinacje dwóch cząsteczek, z których przynajmniej jedna jest aktywna biologicznie [36]. Takie nośniki pozwalają uzyskać większą precyzję w stosowanych siRNA. Cząsteczki siRNA oraz ściśle związana z nimi technika

RNAi to potężne narzędzia biologii molekularnej, które zmieniają oblicze współczesnej nauki i medycyny. Za odkrycie zjawiska interferencji RNA w 2006 roku Andrew Z. Fire i Craig C. Mello dostali nagrodę Nobla w dziedzinie medycyny i fizjologii.

4.2.1.2 miRNA – pierwsze regulatorowe RNA

Cząsteczki miRNA obejmują zakres od 19 do 25 nt. Pierwszym opisanym miRNA w 1993 roku był, wcześniej wspomniany, lin-4. miRNA zostały uznane za osobną klasę cząsteczek dopiero na początku XXI wieku [37]. Potwierdza to dużą dozę sceptycyzmu wśród naukowców, towarzyszącą odkrywaniu ncRNA. MicroRNA wykazują duże podobieństwo do siRNA, ponieważ jedne i drugie uczestniczą w mechanizmie obronnym RNAi. Biologicznie jednak rola miRNA jest zupełnie inna. miRNA kodowane są w genach a ich biologiczną funkcją jest wpływanie na poziom rodzimych komórce transkryptów. Geny miRNA zlokalizować można w różnych miejscach genomów. Około 56% znajduje się w intronach genów kodujących białka, podczas gdy pozostałe możemy odnaleźć w rejonach międzygenowych posiadających własny system regulacji [38]. Polimeraza RNA II i w pewnych przypadkach, III biorą udział w transkrypcji genów miRNA. Biogenezę miRNA można podzielić na kanoniczną i niekanoniczną.

Kanoniczny szlak biogenezy jest dominującym i rozpoczyna się od transkrypcji pri-miRNA pierwszego prekursora miRNA. Kolejno pri-miRNA przetwarzane jest przez kompleks zwany mikroprocesorem, składający się z białka wiążącego DGCR8 oraz enzymu Drosha, do pre-miRNA. Białko DGCR8 wiąże pri-miRNA, a następnie Drosha przecina cząsteczkę pozostawiając charakterystyczne miejsce tworzące strukturę spinki i 2 nt niesparowane na końcu 3’. Wiele transkryptów pozornie odpowiada substratom kompleksu DGCR8/Drosha, jednak nie wszystkie wchodzą w szlak biogenezy kanonicznej miRNA. Kompleks mikroprocesora, preferuje ściśle określone struktury spinki charakteryzujące się długością 35 nt w podstawie, pętlą wierzchołkową o długości ponad 10 nt oraz regionami jednoniciowymi flankującymi trzon cząsteczki. Całość procesów zachodzi w jądrze, a eksport do cytoplazmy zachodzi z udziałem kompleksu eksportyny-5. Kolejno pre-miRNA staje się substratem dla białka Dicer, który przecina pętlę, niwelując strukturę spinki i tworząc dojrzały dupleks miRNA posiadający wystające końce 3’. Nić z końca 5’ spinki nazywamy nicią 5p, z kolei z końca 3’ nicią 3p. Obie nici mogą zostać potencjalnie połączone z białkiem AGO, które dalej wchodzi w skład kompleksu RISC, jednak z danego dupleksu tylko jedna łączy się z AGO, natomiast druga ulega degradacji. Stosunek białek AGO z poszczególnymi nićmi może być komórkowo oraz środowiskowo specyficzny, jednak dobór nici bazuje głównie o stabilność termodynamiczną [39,

40]. Następnie jedna z nici ulega odrzuceniu, a druga przechodzi do kompleksu RISC zwanego również miRISC.

Niekanoniczna ścieżka biogenezy dzieli się na dwie podścieżki - niezależną od kompleksu Drosha/DGCR8 oraz niezależną od Dicer. W przypadku tej pierwszej mówimy o mirtronach powstających podczas wycinania intronów w *splicesomie*, które strukturą przypominają substraty dla białka Dicer i nie wymagają wstępnej obróbki [41]. W przypadku drogi niezależnej od Dicer przetwarzane są shRNA, które w swojej strukturze posiadają idealnie sparowaną strukturę spinki. Obróbka przez białko Drosha pozostawia pre-miRNA o niewystarczającej długości, aby zostały substratami dla białka Dicer, a koniec dojrzewania wymaga aktywności białka AGO, które odcina pętlę kończąc dojrzewanie miRNA [42]. Zawiły proces biogenezy kanonicznej i niekanonicznej miRNA został przedstawiony na Rysunku 2.

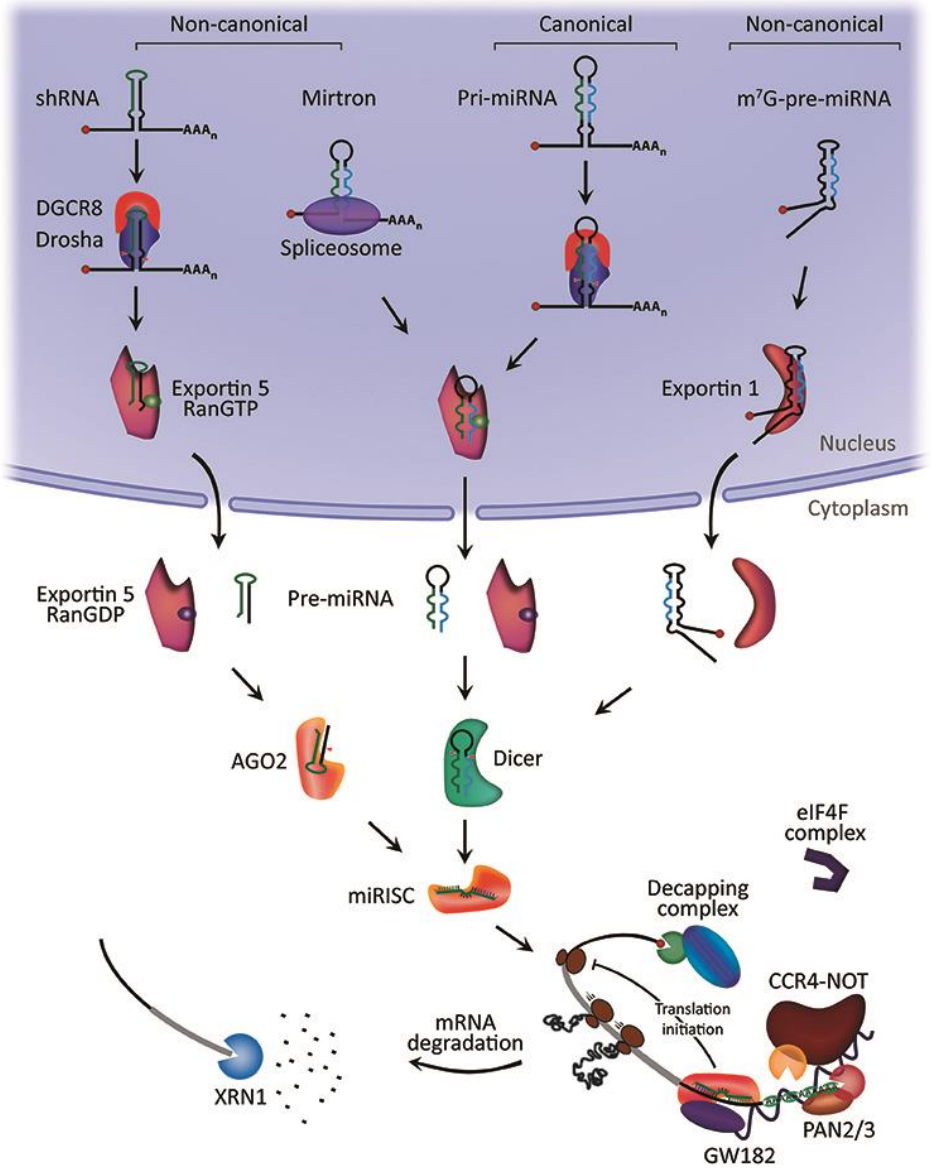


Figura 2. Kanoniczna i niekanoniczna biogeneza oraz mechanizm działania miRNA [38].

Główną funkcją miRNA jest regulacja ekspresji genów gospodarza za pomocą mechanizmów interferencji RNA. Podobnie jak siRNA, miRNA jest częścią kompleksu RISC tworząc miRISC. Kompleks miRISC oddziałuje głównie w rejonie 3'UTR docelowego mRNA, choć są doniesienia o oddziaływaniu także w rejonie 5'UTR z tym, że dużo rzadsze [43, 44]. Region komplementarny do miRNA zwany jest *seed* i stanowi od 2 do 7 nt od końca 5' miRNA. Wpływ miRNA na ekspresję genów może być dwojaki: albo przez pełną degradację celowego mRNA, albo przez inhibicję translacji. Przeciwnie do siRNA, którego komplementarność jest pełna i prowadzi do przecięcia celowanego RNA, komplementarność miRNA nie musi być całkowita. Wraz ze spadkiem komplementarności obserwujemy przesunięcie z kierunku cięcia celowanego mRNA do inhibicji translacji. Białko AGO wraz z miRNA stanowi minimalny kompleks wyciszający zdolny rozszczepić celowany mRNA.

Komórki ssaków kodują 4 białka AGO z czego AGO2 występuje najliczniej i jako jedyne posiada aktywność endonukleazy [45]. Białko AGO wraz z miRNA może tworzyć minimalny kompleks wyciszający. Białko zbudowane jest z 4 domen: N-końcowej, PAZ, MID, PIWI. Składowa N-końcowa odpowiedzialna jest za angażowanie dojrzałego miRNA i rozplatanie dupleksu. Domeny PAZ i MID powalają zakotwiczyć koniec miRNA, gdzie domena PAZ zakotwicza koniec 3' z kolei MID wiąże 5' końcową grupę monofosforową i 5' końcowy nukleotyd. Domena PIWI, kluczowa dla możliwości regulacji ekspresji genów, posiada aktywność nukleazy podobną do RNAzy H [46].

Inhibicja translacji również prowadzi do rozpadu transkryptu i jest zdecydowanie częstszym zjawiskiem. Wynika to z niewymaganego w pełni komplementarnego wiązania miRNA do celowanego mRNA. Mechanizm posiada niższy próg inicjacji, ale jego złożoność jest dużo większa, przez co zaproponowano kilka niewykluczających się scenariuszy. Scenariusze różnią się w zależności od zaangażowanych białek, jednak zasugerowano standardowy model, zakładający wiązanie się do białka AGO białka GW182 będącego pomostem dla kolejnych wiążących się enzymów. Rozpatrując od końca 3' GW182 stanowi połączenie z białkiem wiążącym poli(A), czyli PABP, równocześnie będąc kotwicą dla deadenylaz, takich jak PAN2-PAN3 i CCR4-NOT. Koniec 5' ulega dekapowaniu przez białka takie jak DCP1-DCP2, również angażowane przez GW182, a następnie całe mRNA ulega degradacji przez XRN1. Lista potencjalnych białek angażowanych nie jest zamknięta i wciąż poszerzana o kolejne białka angażowane do kompleksu RISC przez GW182 zdolne wpływać na represję translacji [44].

Obszar działania miRNA to zarówno cytoplazma jak i jądro komórkowe oraz jąderko. Ekspresja miRNA jest w istotny sposób zależna od modyfikacji epigenetycznych [47]. Bardzo często dochodzi też do zjawiska sprzężenia zwrotnego, w którym miRNA regulują poziom enzymów odpowiedzialnych za metylację DNA lub acetylację histonów odpowiedzialnych za

regulację miRNA. Jest to niezwykle ważne w kontekście stanów patologicznych, takich jak nowotwory [48]. Schemat działania miRNA nie ogranicza się tylko do wyciszania ekspresji genów za pośrednictwem mechanizmów RNAi. miRNA są w stanie angażować czynniki transkrypcyjne i być uczestnikami regulacji trójwymiarowej struktury genomowej tworząc kompleksy RNA-białko-DNA [49]. Z uwagi na niebagatelne znaczenie, miRNA są intensywnie badanymi i jednymi z najlepiej poznanych ncRNA w komórce, choć na mapie badań wciąż pozostają białe plamy czekające na odkrycia.

4.2.1.3 piRNA - strażnicy integralności genomu i nie tylko

piRNA to kolejna klasa małych niekodujących RNA. Ich długość mieści się w zakresie od 24 nt do nawet 36 nt. W literaturze odnajdujemy różne przedziały długości, co spowodowane jest brakiem jednolitości tej klasy. Pik odpowiadający piRNA występuje różnie w zależności od pracy analitycznej i badanego materiału. Poza większą długością od miRNA piRNA charakteryzują się 5’ końcową urydyną oraz adenozyną w pozycji 10 bez wykazywania charakterystycznej struktury drugorzędowej [51]. Drugą charakterystyczną cechą piRNA jest 3’ końcowa 2’-O-metlacja. Pierwsze doniesienia o frakcji ncRNA większej od miRNA pochodzą z 2001 roku, z badań nad gonadami męskimi muszki owocowej (*Drosophila melanogaster*) [52].

Początkowo łączono występowanie piRNA z komórkami germinalnymi oraz stanami patologicznymi, takimi jak nowotwory [53, 54]. Istnieją jednak doniesienia o istotnej obecności piRNA w komórkach somatycznych, co sugeruje ich rolę w szerszym zakresie [55]. Główną rolą piRNA jest reagowanie na elementy mobilne genomu, takie jak transpozony. Mechanizm reagowania jest podobny do klasycznego wyciszania ekspresji genów RNAi.

piRNA oddziałują z białkami PIWI, skąd wzięła się ich nazwa. Białka PIWI obok białek AGO należą do podrodziny białek Argonaute. U człowieka zidentyfikowano jak dotąd cztery białka Piwi, od Piwil1 do Piwil4. Schematycznie na białko Piwi składają się 3 domeny: PAZ – wiążąca specyficznie koniec 3’ piRNA, który jest 2’- O-metylowany, MID – środkowa domena wiążąca się z fosforanem końca 5’, domena C-końcowa PIWI - działająca jak RNaza H [56]. Obecność białek PIWI wydaje się być wysoce konserwatywna, gdyż ich homologi znajdujemy w komórkach organizmów modelowych, od takich jak wspomniana *D.melanogaster* (ABU, AGO3, PIWI) przez myszy (MIWI, MIWI2, MILI), aż po ludzi. Nadekspresja białek Piwi jest wiązana często ze stanami nowotworowymi oraz wpływem na parametry życiowe komórek [57]. Piwi odgrywają podobną rolę w funkcji piRNA jak białka AGO dla siRNA/miRNA. Proces wyciszania niepożądanych transkryptów bazuje na podobnej idei jak w przypadku RNAi.

Komplementarność piRNA determinuje cel białka Piwi, które z udziałem domeny PIWI degraduje docelowy transkrypt.

Zagadnienie biogenezy piRNA jest wciąż badane i uzupełniane o kolejne odkrycia. Szlaki zostały najlepiej poznane na organizmie modelowym, w którym odkryto piRNA, czyli *D. melanogaster,* jednak białka biorące udział w szlakach biosyntezy piRNA różnią się międzygatunkowo, stąd procesy oraz dodatkowe enzymy mogą być różne względem gatunku. Biosynteza piRNA odbywa się dwuetapowo: przez biosyntezę pierwotną i wtórną.

Szlak pierwotny zakłada transkrypcję piRNA z charakterystycznych rejonów w genomie, zwanych klastrami piRNA, odpowiadających sekwencyjnie tylko fragmentom transpozonów. Takie fragmenty przechowują genetyczną pamięć o transpozonie, ale same nie są mobilne w obrębie genomu [58]. Klastry piRNA to długie regiony ulegające transkrypcji pod wpływem polimerazy RNA II, o dużym zróżnicowaniu lokalizacji, nawet w obrębie gatunku [59]. Klastry mogą ulegać transkrypcji z jednej lub z obu nici DNA, co oznacza, że piRNA pochodzą zarówno z nici sensowej jak i antysensowej [60]. Istnieją również pri-piRNA pochodzące z mRNA, tRNA, lncRNA oraz snoRNA, jednak w wyniku ich transkrypcji powstają tylko nici sensowe [61]. Pierwotny transkrypt pri-piRNA przechodzi podobną do mRNA drogę, czyli podlega wycinaniu intronów oraz przyłączeniu 5'CAP i 3'Poli(A) [62]. Biosynteza pierwotna uwzględnia nici antysensowe pri-piRNA. Długie pri-piRNA transportowane są do cytoplazmy i następnie w strukturach zwanych ciałkami Yb są przycinane do odpowiedniej długości przez enzym Zicchini [63]. Do końca 5' piRNA przyłącza się białko PIWI, następnie piRNA w kompleksie jest przycinane przez egzonukleazy, a koniec 3' ulega metylacji przez HEN1 dając w pełni funkcjonalny kompleks.

W przypadku biosyntezy wtórnej, do piRNA eksportowanego do cytoplazmy przyłączają się białka Piwi, czyli Aub albo Ago3. Przyłączenie się tych białek do piRNA zapoczątkowuje cykl amplifikacyjny. Cykl amplifikacji „ping-ponga'' jest zaproponowanym drugim źródłem piRNA zwiększającym ilość regulujących cząsteczek w komórce. Łączy się on z funkcją piRNA, gdyż do swojego działania wymaga celowanego transpozonu lub innego transkryptu, który jest celem piRNA, a zarazem daje źródło nowych piRNA. Cały cykl odbywa się w cytoplazmie w ciałkach Yb, gdzie powstają dwa kompleksy piRNA/Ago3 i piRNA/Aub, których nici wiodące wzajemnie się uzupełniają w zakresie końca 5'. Antysensowy piRNA wykazujący komplementarność do nici docelowego RNA wchodzi w skład kompleksu piRNA/Aub. W wyniku przecięcia celowanego RNA przy końcu 3' powstaje transkrypt stanowiący sensowe piRNA wchodzących w kompleks z białkiem Ago3. Kompleks piRNA/Ago3 wykazuje komplementarność i prowadzi do przecięcia transkryptu antysensowego klastra piRNA. W

wyniku cięcia z końca 3’ powstaje kolejne antysensowe piRNA które wchodząc w kompleks z białkiem Aub zamyka cykl. W toku działania cyklu powstają nadmiarowe piRNA które dalej pełnią funkcję w komórce. piRNA powstające w cyklu „ping-ponga’’ posiadają charakterystyczne modyfikacje. Kompleks piRNA/Ago3 produkuje piRNA cechujące się obecnością adeniny w pozycji 10, z kolei piRNA/Aub obecnością urydyny na końcu 5’ [64].

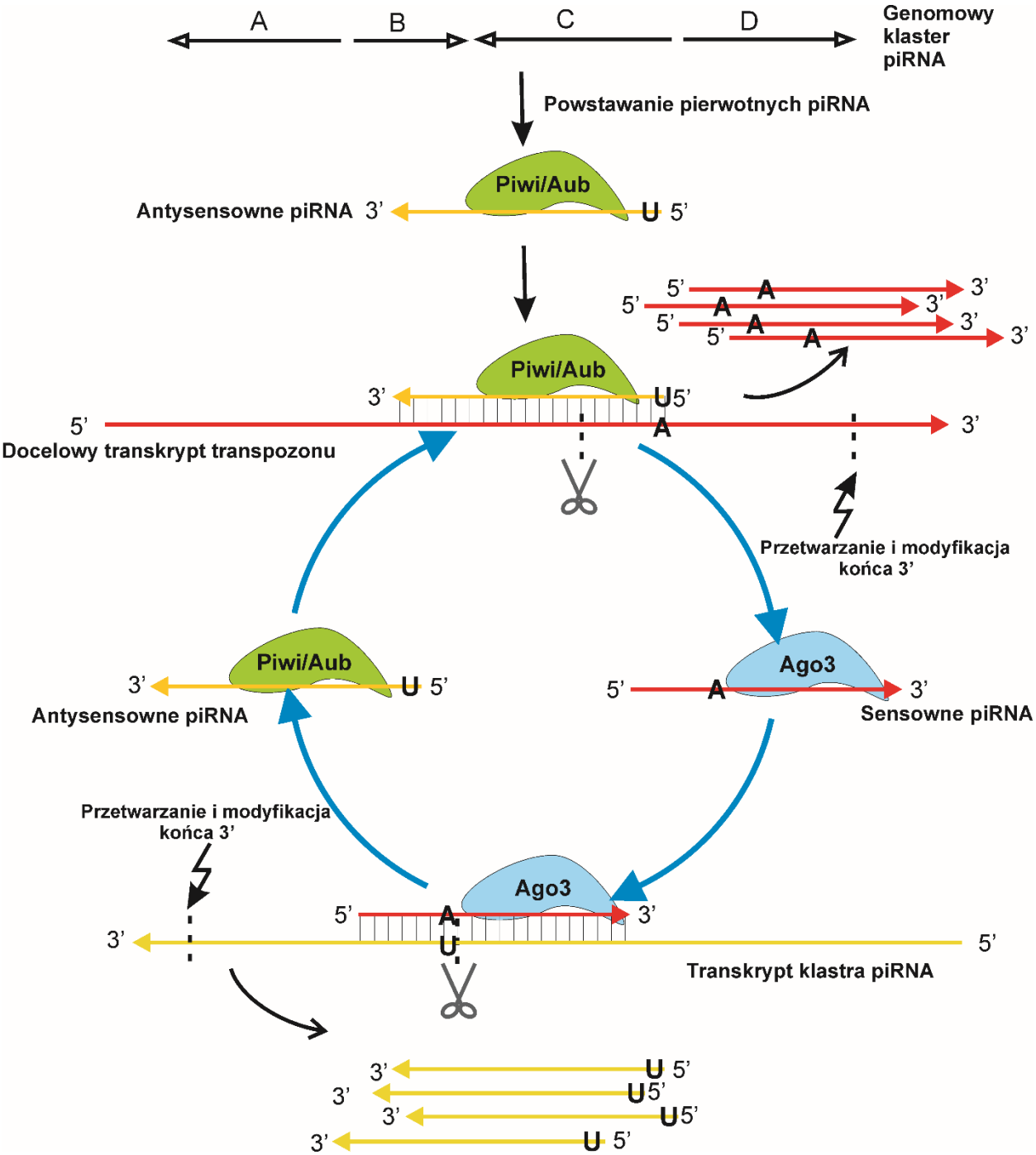


Figura 3. Biosynteza wtórna piRNA (Kowalczykiewicz D., Wrzesinski J. Post Bioch,2011, 57:3)

Regulacja za pośrednictwem piRNA występuje na kilku poziomach. Wyciszanie na poziomie transkrypcyjnym odbywa się za pośrednictwem enzymów modyfikujących histony lub metylotransferaz DNA. Kompleksy piRNA/Piwi i wcześniej wspomniane enzymy, wprowadzają modyfikacje epigenetyczne wpływające na procesy transkrypcji poszczególnych genów [65]. Na poziomie regulacji potranskrypcyjnej, piRNA funkcjonują podobnie jak miRNA prowadząc do degradacji celowanego RNA, którym może być mRNA, lncRNA lub transkrybowane pseudogeny [66, 67, 68]. Największym zainteresowaniem cieszą się piRNA będące kandydatami na biomarkery nowotworowe. Ekspresja piRNA i białek Piwi jest mocno związana z postępującymi zmianami parametrów komórkowych, a obecne techniki pozwalają szybko wykrywać pojawiające się w nadekspresji cząsteczki [61]. Wciąż wiele zagadnień dotyczących funkcji i biogenezy piRNA pozostaje niejasnych i poszlakowych, co otwiera pola do dalszych badań i odkryć.

4.2.1.4 tRF - najbardziej zróżnicowana frakcja sncRNA

Istnienie frakcji tRF jest ściśle związane z metabolizmem cząsteczek tRNA. Cząsteczki tRF podzielone są na kilka podklas, w zależności od procesu biogenezy, obejmując i-tRF, tRF-1, tRF-2, tRF-3 i tRF-5. Cząsteczki tRF otrzymujące swoją nazwę od końca 5' macierzystego tRNA są określane jako tRF-5 i występują w trzech podtypach: tRF-5a (o długości 14 -16 nt), tRF-5b (o długości 22 - 24 nt) i tRF-5c (o długości 28 - 30 nt). Cząsteczki tRF-5 powstają w wyniku cięcia pętli D macierzystego tRNA, a za ten proces sugeruje się udział białka Dicer [69]. Analogicznie, tRF zawierające koniec 3' macierzystego tRNA nazywane są tRF-3 i występują w dwóch podtypach: tRF-3a (o długości 18 nt) i tRF-3b (o długości 22 nt). Ich powstanie jest rezultatem cięcia w pętli TψC. Potencjalnymi enzymami uczestniczącymi w procesie formowania tRF-3 mogą być Angiogenina, Dicer lub inne egzonukleazy [70]. Fragmenty tRF-2 pochodzą z pętli antykodonowej macierzystego tRNA, ale nie zawierają charakterystycznych grup kończących ani na 5', ani na 3'. Proces tworzenia tRF-2 najczęściej ma miejsce w warunkach niskiego poziomu tlenu w środowisku komórki [71]. tRF-1 powstają z pre-tRNA za pośrednictwem RNAzy Z. W trakcie dojrzewania pre-tRNA, obszar 3' UTR jest cięty przez RNAzę Z w charakterystycznym miejscu zawierającym poli(U). i-tRF to cząsteczki pochodzące z różnych wewnętrznych regionów macierzystego tRNA, ale nie zawierają one charakterystycznych grup zakończeń 3' i 5'. Ich formowanie zależy od różnych enzymów, w

zależności od miejsca ich pochodzenia [72]. Oprócz tRF o określonej strukturze, istnieją także połówki tRNA, znane jako tiRNA. Połówki tRNA powstają w reakcji na warunki stresowe komórki, a za cięcia w pętli antykodonowej odpowiedzialna jest głównie Angiogenina, chociaż mogą w tym procesie uczestniczyć także inne RNAzy. Analogicznie do tRF-3 i tRF-5, rozróżniamy 3' i 5' tiRNA [73].

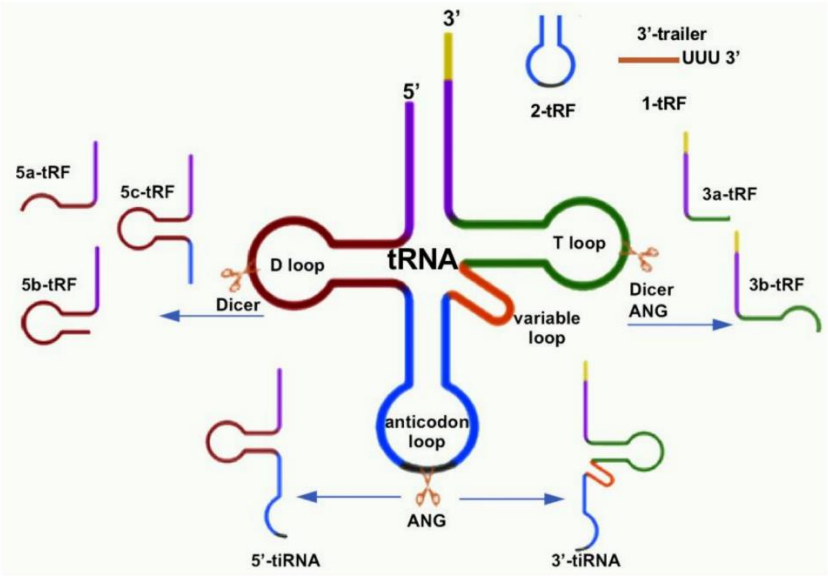


Figura 4. Klasyfikacja cząsteczek tRF występujących w komórce [73].

Mnogość cząsteczek sugeruje różnorodne funkcje. tRFy mogą wpływać na wiele procesów na różnych etapach regulacji ekspresji genów. Regulacja na poziomie translacji może zachodzić poprzez wiele mechanizmów. Na przykład, wiązanie tRF do podjednostek rybosomów może obniżać aktywność translacji [74]. Inny mechanizm to cząsteczki tRF zawierające motyw oligoguaniny na końcu 5' (TOG), które hamują translację poprzez tworzenie struktury kwadrupleksu RG4 i wypierają kompleks inicjacji translacji eIF4G/eIF4E [75]. Przykładem innego mechanizmu jest tRF-3, który może oddziaływać z białkami Ago3 i Ago4, tworząc kompleks, w którym tRF działa jako nić wiodąca dla nukleaz, co prowadzi do degradacji docelowego transkryptu [76]. Ponadto, fragmenty tRNA mogą uczestniczyć w tworzeniu granulek stresowych, które wpływają na wyciszanie ekspresji mRNA [77]. Istnieją także doniesienia o roli tRF w supresji wirusa HIV. Odpowiedź tRF na obecność wirusa może być zróżnicowana. Na przykład, tRF-3006 może działać jako reporter zakażenia wirusem, a także łączyć się i blokować region 56-67 nt tRNALys, który jest rozpoznawany jako primer odwrotnej transkrypcji przez wirusa HIV [78]. Odnajdujemy także cząsteczki tRF w regulacjach epigenetycznych. Przykładem regulacji poza sekwencyjnej jest tRF Glu(TTC), który wywiera wpływ na metabolizm tłuszczu [79]. tRFy mogą również wpływać na cykl komórkowy, gdzie niektóre z nich zdolne są do wiązania się z cytochromem C, co prowadzi do zablokowania

kaspazy-9 i zahamowania procesu apoptozy [80]. Poziomy cząsteczek tRF wzrasta w stanach patologicznych, dlatego też są silnie związane z tematem biomarkerów. Bioróżnorodność cząsteczek tRF sugeruje wiele funkcji, jednak uniemożliwia to stworzenie jednego kanonicznego szlaku funkcji cząsteczek tRF. To jednocześnie stanowi wyzwanie w jednoznacznym określeniu istotności fragmentów pochodzących z tRNA i mechanizmów w danym procesie, ale również stanowi zaletę, sugerując, że zawsze warto rozważyć potencjalny wpływ cząsteczek tRF na dany proces.

4.2.2 lncRNA jedna grupa o wielu obliczach

Brak ściśle określonego zakresu długości czyni lncRNA niezwykle zróżnicowaną grupą. Ze względu na tę różnorodność, podział lncRNA jest trudny, a żadna cecha nie wydaje się być na tyle charakterystyczna, aby mogła służyć jako wyraźna linia podziału. Pod wieloma względami, lncRNA są podobne do mRNA. Ulegają one transkrypcji głównie przez Polimerazę RNA II, mogą być poliadenylowane, ulegają *splicingu* oraz posiadają CAP na końcu 5' [81]. Jednak, w przeciwieństwie do mRNA, gdzie sekwencja i układ ramek odczytu mają pierwszorzędne znaczenie, w przypadku lncRNA istota ich roli skupia się na strukturze drugo- i trzeciorzędowej.

Geny lncRNA można odnaleźć w wielu miejscach w całym genomie. Odkryto, że długie niekodujące RNA są wielodomenowymi złożonymi strukturami, a ich domeny mogą różnić się stopniem konserwacji w obrębie tej samej cząsteczki [82]. Wybiórcza konserwacja wynika często z udziału elementów mobilnych genomu, takich jak transpozony. Wiele eksonów lncRNA ma swoje pochodzenie w transpozonach [83, 84]. Cząsteczki lncRNA mogą pochodzić z różnych obszarów genomu, takich jak sekwencje między genowe, nici antysensowe, introny [81]. Odkryto nawet lncRNA, których geny zawierają sekwencje kodujące białka [81]. Istnieją także doniesienia o lncRNA pochodzących z pseudogenów [85]. Ekspresja lncRNA jest komórkowo bardzo specyficzna, do tego stopnia, że komórki tego samego rodzaju, ale w różnych stadiach rozwoju, różnią się profilem ekspresji lncRNA [86]. Krajobraz długich niekodujących RNA jest kształtowany również przez czynniki zewnętrzne, takie jak środowisko, stres komórkowy oraz dostępność składników odżywczych [87].

Główną rolą lncRNA jest regulacja procesów komórkowych. Jednak określenie konkretnej roli lub funkcji danego transkryptu jest niezwykle trudne. Często zmiany fenotypowe są subtelne i pojawiają się jedynie na poziomie molekularnym. lncRNA regulują procesy na różnych poziomach życia komórkowego, oddziałując zarówno z innymi RNA, DNA, białkami, jak również peptydami i związkami niskocząsteczkowymi [88, 89]. Wiele funkcji lncRNA jest wynikiem różnorodnych możliwości tych cząsteczek, ale trzy główne role można uznać za

wspólne mianowniki dla lncRNA: udział w modyfikacjach chromatyny, tworzenie koacerwatów oraz działanie jako wzmacniacze ekspresji genów.

Jednym z kluczowych obszarów, w których lncRNA odgrywają istotną rolę, jest epigenetyczna modyfikacja chromatyny. Przykładem jest lncRNA o nazwie DACOR1, która wchodzi w interakcję z metylotransferazą DNA DNMT1 i jest wysoce specyficzna dla raka jelita grubego [90]. Inny mechanizm epigenetyczny, w którym uczestniczą lncRNA, to modyfikacje histonów. Zidentyfikowano wiele lncRNA, które oddziałują z kompleksami modyfikującymi białka histonowe. Jednym z dobrze zrozumianych mechanizmów jest oddziaływanie z kompleksem PRC2, dla którego przewiduje się wiele potencjalnych interakcji z lncRNA [91].

Długie niekodujące transkrypty często pełnią rolę transkryptów regulujących ekspresję genów, zwanych wzmacniaczami. Wzmacniacze działają jako lokalne centra transkrypcji i *splicingu*, angażując czynniki transkrypcyjne i wpływając na strukturę chromatyny w określonych miejscach [92]. System aktywnych wzmacniaczy tworzy unikalne wzory ekspresji genów w różnych komórkach, co wpływa na charakterystyczne cechy każdej z tych komórek [93]. Przykłady lncRNA, takie jak Firre, Peril czy Maenli, wykazujące tę rolę, zostały już opisane, choć wiele pozostaje do zbadania [94, 95, 96].

Kolejną ważną rolą lncRNA jest tworzenie koacerwatów bimolekularnych. Kondensaty te to zgrupowania makrocząsteczek oddzielonych fazowo, charakteryzujące się dużą dynamiką i potencjalnym funkcjonalnym [97]. W ich skład wchodzą zarówno RNA, jak i białka. Białka składające się na koacerwaty są często bogate w regiony wewnętrznie nieuporządkowane [98]. Te niestabilne i elastyczne struktury trzeciorzędowe pozwalają na interakcje z różnymi cząstkami, co sprawia, że kompleksy te są potencjalnie wielofunkcyjne. lncRNA odgrywają rolę architektonicznego szkieletu w tych kondensatach, nadając im odpowiednią strukturę i determinując ich funkcje [99]. Przykłady takiej działalności można znaleźć w przypadku lncRNA takich jak SLERT i LENT w jąderkach czy MALAT1 w plamkach jądrowych, a także NORAD, który zapobiega wyłapywaniu białek Pumilio [81].

Jedną z wyróżniających cech lncRNA jest ich długość oraz zdolność do przyjmowania różnorodnych struktur. Wielu z tych długich niekodujących RNA podlega również alternatywnemu *splicingowi*, podobnie jak mRNA. To otwiera szeroki zakres możliwości i zmienności na poziomie regulacyjnym. Warto zauważyć pewien paradoks, który dotyczy niektórych reprezentantów tej klasy cząsteczek. Podczas analizy sekwencji lncRNA odkryto obecność sekwencji ORF (open reading frame) [100]. To sugeruje, że niektóre lncRNA mogą jednak zawierać kodujące fragmenty. W trakcie badań udowodniono istnienie krótkich peptydów pochodzących z lncRNA, a co więcej, wykazano, że te peptydy mogą mieć istotny wpływ na funkcjonowanie komórki [101]. Przykładem takiego peptydu jest SPAR kodowany przez

LINC00961. SPAR działa jako bloker v-ATPazy, co uniemożliwia aktywację mTORC1, regulatora uwalniania aminokwasów z lizosomów [102]. Ciekawe jest to, że dalsze badania nad LINC00961 i SPAR wykazały, że zarówno lncRNA, jak i jego peptyd, mogą pełnić funkcje niezależnie od siebie. Co więcej, ich funkcje są antagonistyczne, ponieważ LINC00961 wiąże się z białkiem Tβ4, które wiąże aktynę i hamuje angiogenezę, podczas gdy SPAR wiąże się z innym białkiem wiążącym aktynę, SYNE1, i promuje angiogenezę [103]. Peptydy pochodzące z lncRNA uczestniczą również w mechanizmach kontroli jakości mRNA. Takim polipeptydem jest NoBody kodowany przez LINC01420, który bierze udział w degradacji mRNA z przedwczesnymi kodonami terminacji translacji poprzez mechanizm usuwania 5' Cap [104]. Warto rozważyć możliwość wyodrębnienia frakcji lncRNA kodujących peptydy od frakcji lncRNA, ponieważ paradoks ten, który się pojawia w miarę odkrywania nowych funkcji peptydów pochodzących z lncRNA, może prowadzić do wielu nieporozumień.

4.2.2.1 circRNA – nietypowe RNA

Jednym z wyróżniających się podtypów lncRNA są cyrkularne niekodujące RNA, znane jako circRNA. CircRNA powstają głównie z loci kodujących białka za pośrednictwem trzech mechanizmów: bezpośredniego *splicingu* wstecznego, cyrkulacji sterowanej parowaniem intronów oraz cyrkulacji sterowanej lariatem [105]. W przypadku circRNA alternatywne składanie transkryptów może prowadzić do powstania kolistych RNA, które mogą być złożone zarówno z kompilacji intronów i egzonów, jak i zawierać jedynie jednego typu sekwencje [106]. CircRNA wykazują zdolność do interakcji z białkami oraz kwasami nukleinowymi. Analizy bioinformatyczne sugerują, że circRNA mogą potencjalnie pełnić rolę molekularnych "gąbek" dla miRNA, choć ta teza nadal oczekuje na potwierdzenie eksperymentalne [106].

W interakcjach z białkami circRNA spełniają dwie główne funkcje. Mogą wiązać się specyficznie z białkami, zmieniając ich aktywność fizjologiczną lub służyć jako rusztowanie dla kompleksów białkowych [107]. Na przykład circNDUFB2 uczestniczy w degradacji białek wiążących mRNA insulinopodobnego czynnika wzrostu 2 (IGF2BP), tworząc kompleks między wspomnianym białkiem a białkiem zawierającym motyw trójdzielny 25 (TRIM25). Skutkiem tego jest degradacja IGF2BP i pozytywny efekt hamowania postępów nowotworzenia [108]. Innym przykładem korzystnego działania circRNA jest hamowanie przerzutów nowotworu żołądka przez circ-TNPO3, który łączy się z czynnikiem IGF2BP3, który promuje proliferację i przerzuty nowotworowe [109].

Podobnie jak liniowe lncRNA, cyrkularne RNA mogą kodować krótkie peptydy. Charakterystyczne motywy wewnętrznego miejsca wejścia rybosomu (IRES) umożliwiają

translację, jednak brak 5'Cap powoduje, że jest to translacja oporna na białka regulujące ten proces, takie jak białko 4E-BP [110]. CircRNA są intensywnie badane w kontekście nowotworów, gdzie mogą służyć jako specyficzne biomarkery stanów patologicznych [111]. Przejście od struktury liniowej do struktury cyrkularnej otwiera szeroki zakres potencjalnych zastosowań i stawia circRNA jako cząsteczki z dużym potencjałem.

5 Cel naukowy przedstawionego cyklu prac

Nadrzędnym celem niniejszej rozprawy doktorskiej jest poznanie wpływu niekodujących RNA na rozwój nerek.

Po pierwsze zamierzano poznać profil występowania cząsteczek tRF w różnych organach zwierzęcia modelowego jakim jest S. scrofa oraz wykazać możliwość interakcji cząsteczki tRF z białkami Piwi (Kazimierczyk, M., et al., BBRC, 2019).

Drugim celem była charakterystyka występowania cząsteczek tRF w zaproponowanym modelu komórkowym rozwoju nerki oraz poznanie ich potencjalnej funkcji. Zamierzałem sprawdzić wcześniejsze doniesienia o potencjalnym udziale nukleazy Dicer w biogenezie cząsteczek tRF (Kazimierczyk, M., et al., Int. J. Mol. Sci., 2022).

Kolejną częścią badań było zgromadzenie aktualnych informacji dotyczących interaktomu lncRNA oraz charakterystyka modyfikacji epigenetycznych lncRNA i analiza zebranych infromacji w kontekście różnicowania i rozwoju komórki (Kazimierczyk, M., et al., Int. J. Mol. Sci., 2020, Kazimierczyk, M Wrzesinski, J., Int. J. Mol. Sci., 2021).

6 Skrótowy opis przedstawionych prac

6.1 tRNA-derived fragments from the Sus scrofa tissues provide evidence of their conserved role in mammalian development.

Kazimierczyk, M., Jędroszkowiak, A., Kowalczykiewicz, D., Szymanski, M., Imiołczyk, B., Ciesiołka, J., Wrzesinski J. (2019) Biochem. Biophys. Res. Commun. 520, 514-519.

W niniejszej publikacji przedstawiono badania dotyczące występowania cząsteczek tRF, które są reprezentantami grupy małych niekodujących RNA, w różnych tkankach zwierzęcia modelowego, jakim jest S. scrofa. Dodatkowo, w pracy dowiedziono, że cząsteczki tRF mają zdolność do wiązania się z białkami rodziny Argonaute. Badania opierały się na wcześniejszych wynikach analizy bioinformatycznej, która opisywała występowanie tRFów w gonadach. Następnie te wyniki zostały uzupełnione danymi eksperymentalnymi.

W wyniku wcześniejszych badań ustalono, że połówki tRNA, takie jak 5'tRFGly (GCC), 5'tRFVal (CAC), 5'tRFHis (GUG), 5'tRFLys (CUU) i 5'tRFGlu (UUC), występują w największych ilościach. Te wyniki zostały potwierdzone ilościowo poprzez analizę eksperymentalną przy użyciu techniki northern blot. Następnie, przystąpiono do zaprojektowania par sond znakowanych izotopem fosforu ^{32}P, które były komplementarne do końców 3' i 5' wybranego tRNAValCAC. To pozwoliło na wykrycie obu połówek tRNA w próbkach pochodzących z różnych tkanek Sus Scrofa, takich jak żołądek, wątroba, serce, płuca, nerki, mózg i mięśnie. Wyniki tych analiz wykazały, że zarówno dla sondy 5', jak i 3', wykryto podobny wzorzec występowania tRFów pochodzących z tRNAVal(CAC) w gonadach i nerkach. Podobne badania przeprowadzono także dla cząsteczek 5'tRF, takich jak Gly(GCC), Lys(CUU) i Glu(UUC). Wybór tych cząsteczek opierał się na wcześniejszej analizie bioinformatycznej, która wykazała podobny stosunek występowania tRNA/tRF na poziomie około 6% w próbkach pochodzących z gonad. Te wyniki ponownie potwierdziły podobny profil ekspresji tRF w gonadach i nerkach, co sugeruje potencjalne istnienie istotnych funkcji tych cząsteczek.

W pracy zbadano także interakcje małych niekodujących RNA z białkami rodziny Argonaute, która dzieli się na białka Ago i Piwi. Białka Piwi są biologicznie związane z funkcjami w gonadach. Wykorzystano technikę EMSA do sprawdzenia potencjalnego wiązania tRF z białkami Piwi. W celu uzyskania białek Piwi, przeprowadzono nadekspresję tych białek w komórkach E. coli, wykorzystując wektor pET151/D-TOPO zawierający otwartą ramkę odczytu dla białek Piwil1, Piwil2 i Piwil4. Ostatecznie, trudności w nadekspresji pełnych białek skłoniły do ograniczenia się do domeny PAZ wybranego białka. Domena PAZ jest znana z wiązania

kwasów nukleinowych i stanowi istotny element białka Piwi. Doniesienia literaturowe sugerowały, że powiązanie białka Piwil 4 z tRF w przypadku raka piersi MDAMB231 oraz w jądrach myszy typują domenę PAZ białka Piwil4 jako najlepszą do procesu oczyszczania i dalszych badań. W eksperymentach przeprowadzonych w tej pracy oczyszczoną domenę PAZ Piwil4 wykorzystano jako kontrolę pozytywną z piRNA, jako kontrolę negatywną z miRNA, oraz eksperymentalnie z tRFVal(CAC). Wyniki tych badań zostały zaprezentowane na Figurze 1. Stała dysocjacji dla każdego wariantu była znacznie większa dla cząsteczki tRF (6,69 ± 1,36 μM) niż dla kontrolnego piRNA (1,70 ± 0,46 μM), co sugeruje silne powinowactwo między badanymi cząsteczkami tRF a białkami Piwi.

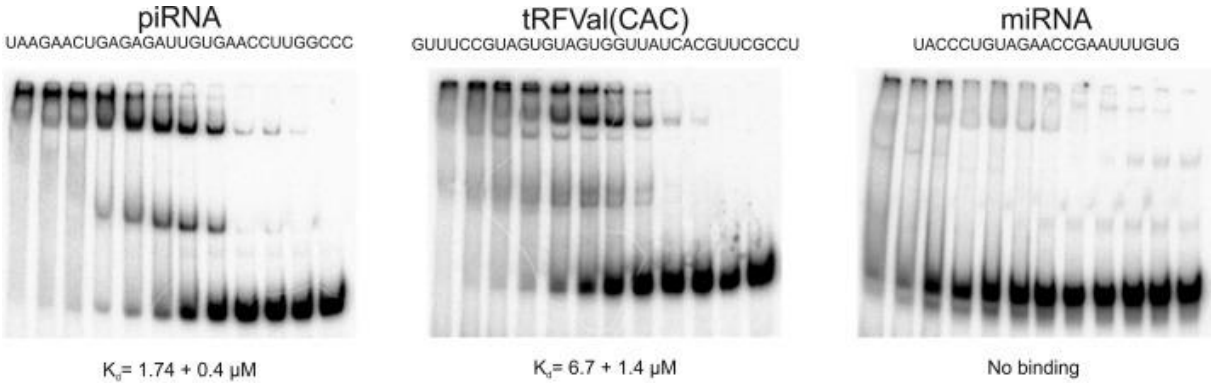


Figura 5. Zestawienie wyników eksperymentu EMSA przeprowadzonego dla domeny PAZ białka Piwil4 oraz poszczególnych niekodujących RNA.

Dotychczas opisane interakcje białek Piwi-RNA skupiały się głównie na gonadach, jednakże funkcja tych interakcji w komórkach somatycznych była nieznana. Podobne wyrażanie cząsteczek tRF w komórkach gonad i nerkach może być związane z bliskością obu układów na etapie rozwoju embrionalnego, kiedy to narządy te są położone obok siebie i korzystają z tego samego przewodu prowadzącego. Warto zaznaczyć, że w przypadku nowotworu piersi MDAMB231 kompleksy Piwil 4 i tRFGly(GCC) oraz tRFVal(CAC) uczestniczą w regulacji translacji. Wnioski z przeprowadzonych badań stanowią fundamenty dla dalszych eksploracji roli małych niekodujących RNA w komórkach nerek.

6.2 Differential expression of transfer RNA-derived fragments (tRF) during renal cell development. The role of Dicer in tRF biogenesis.

Kazimierczyk, M., Wojnicka, M., Biała, E., Żydowicz-Machtel, P., Ostrowski, T., Kurzyńska-Kokorniak, A., Wrzesiński, J. (2021) Int. J. Mol. Sci. 23(7), 3644.

Publikacja ta wynika z wcześniejszych badań przeprowadzonych na modelu zwierzęcym, *S.Scrofa*, gdzie założenie o wspólnym kanale rozwoju dla układu wydalniczego i płciowego zostało przeniesione na poziom modelu komórkowego człowieka. W pracy zaakceptowano proponowany model rozwoju nerki, który skomponowany jest z czterech różnych linii komórkowych reprezentujących poszczególne etapy rozwoju nerki. W skład tego modelu wchodzą komórki: hESC (komórki macierzyste), HEK293T (komórki embrionalne nerkowe), HK-2 (komórki kompletnie wykształconej nerki) oraz A-498 (komórki nowotworowe nerki, jako forma odstępstwa od normy).

Wspólnie z firmą Macrogen (Korea Południowa), przeprowadzono sekwencjonowanie specjalnie przygotowanych bibliotek małych RNA pochodzących z każdej linii komórkowej. Analizę bioinformatyczną prowadziłem wspólnie z dr Michałem Szcześniakiem (Zakład Bioinformatyki, Uniwersytetu im. Adama Mickiewicza w Poznaniu) oraz dr Natalią Szostak (Pracownia Bioinformatyki, Instytutu Chemii Bioorganicznej w Poznaniu). W zaprezentowanej publikacji skupiłem się na cząsteczkach tRF i ich wyrażaniu na poszczególnych etapach wspomnianego modelu. Analiza bioinfromatyczna wykazała najwięcej zidentyfikowanych cząsteczek tRF specyficznie dla komórek hESC i A-498. Największą wspólną grupą cząsteczek tRF była grupa pomiędzy komórkami hESC, HEK293T i A498. Z analiz bioinformatycznych wynikało, że najwyżej wyrażanymi w całym modelu cząsteczkami tRF pochodzą z tRNA: Gly(CCC), Val(AAC), Arg(CCU). W kolejnym kroku potwierdziłem dane z analizy bioinformatycznej za pomocą eksperymentu northern blot. Zaprojektowałem sondy DNA znakowane izotopem fosforu ^{32}P tak aby wykrywały wszystkie fragmenty pochodzące z końca 5' i 3' badanych tRNA. Wykrywane tą strategią cząsteczki tRF można podzielić na trzy grupy pochodzenia z cięcia zachodzącego w rejonie pętli D, antykodonowej i TΨC. W zależności od końca hybrydyzującego sondę cięcia w tych samych rejonach będą odwrotnie widoczne, jeśli chodzi o pętlę D i TΨC i tak samo widoczne dla pętli antykodonowej. Uzyskane wyniki prążki pogrupowałem zależnie od rejonu pochodzenia i odpowiednio zliczyłem w stosunku do macierzystego tRNA za pomocą programu Multi Gauge V3.0 tak aby znormalizować wyniki i umożliwić ich porównanie pomiędzy liniami komórkowymi. Wzory hybrydyzacji zanotowanych na membranach przeniosłem na zaimportowane z GtRNAdb modele 2D badanych cząsteczek

tRNA co umożliwiło identyfikację ich sekwencji i ID w bazie tRFdb. Dalsze analizy w bazach danych prowadziłem tylko na cząsteczkach tRF które udało zidentyfikować na obu membranach i których ID udało się ustalić we wspomnianej bazie danych traktując wyniki jako wiarygodne.

Korzystając z ID cząsteczek tRF wykorzystałem bazę danych tRFTars wykrywającą potencjalne mRNA mogące oddziaływać z potwierdzonymi cząsteczkami tRF. Z pomocą narzędzia jakim jest tRFTars oraz przeglądu literaturowego opracowałem tabelę potencjalnych możliwych oddziaływań zidentyfikowanych cząsteczek tRF na transkrypty białek zaangażowanych w rozwój i różnicowanie komórek nerek u człowieka. Przykładową tabelę dla 5' i 3'tRFGly(CCC) zawartą w publikacji zamieszczono poniżej.

5'tRF	tRFdb ID	Length	Positive probability	Gene	Protein	Function in the kidneys
GlyCCC	5004b	22 nt	0,8696	MEIS2	Homeobox protein Meis2	• Overexpression during aging and kidney damage • Participates in nephrogenesis • Meis1 is required for correct vascular network formation in the embryo
			0,8685	SH3TC	SH3 domain and tetratricopeptide repeat-containing protein 2	• miR-584, which is located within the SH3TC intron, is a tumor suppressor, directly targeting the Rock-1 oncogene and reducing the invasion capacity in human clear cell renal cell carcinoma
			0,7550	FMN1	Formin-1	• Genomic rearrangements of the FMN1 locus cause kidney defects • Associated with acro-renal disorders
			0,7099	CTDSPL 2	CTD small phosphatase-like protein 2	• Dephosphorylation of the FoxOs protein by stimulating proteolysis during nephritis
	5004c	31 nt	0,8077	ZBTB20	Zinc finger and BTB domain-containing protein 20	• Glucose homeostasis
			0,7730	KSR2	Kinase suppressor of Ras 2	• Cell energy homeostasis

3'tRF	tRFdb ID	Length	Probability	Gene	Protein	Function in the kidneys
GlyCCC	3027a	tRF3b 17nt	1,0000	ORAI2	Protein orai-2	• Transporting the Ca^{2+} ion
			0,9868	CCDC9B0	Coiled-coil domain-containing protein 9B	• Cell morphology, cellular assembly and organization, cellular function and maintenance
			0,9854	CYP20A1	Cytochrome P450 20A1	• Involved in diabetic nephropathy
			0,9841	CACNG8	Voltage-dependent calcium channel gamma-8 subunit	• Involved in transporting calcium ions
			0,9806	GAS7	Growth arrest-specific protein 7	• Involved in proteinuria and chronic kidney disease
	3027b	22	0,9404	PLD6	Phospholipase D Family Member 6	• Member of a family of proteins correlated with Kidney Fibrosis
			0,8715	LY6G6C	Lymphocyte antigen 6 complex locus protein G6c	• LY6G6C belongs to a cluster of leukocyte antigen-6 (LY6) genes located in the major histocompatibility complex (MHC) class III region

Przy okazji realizacji głównego nurtu badań przeprowadziłem hybrydyzację northern blot wykorzystując materiał pochodzących z komórek pozbawionych ekspresji białka Dicer. Białko Dicer dotychczas uważane jako enzym tworzący cząsteczki tRF z pętli D. Układ doświadczalny zakładał wykorzystanie komórek HEK293T typu dzikiego, HEK293T pozbawionych genu białka Dicer, HEK293T nadprodukujące białko Dicer z wykorzystaniem plazmidu oraz wariant wyrażający białko Dicer tylko z plazmidu. Wyniki zaprezentowano w Figurze 2. Brak sygnału hybrydyzacji na wysokości około 60 nt potwierdziły niezbędną rolę białka Dicer w powstawaniu cząsteczek tRF z pętli D. Nadekspresja białka Dicer nie odznaczała się przyrostem cząsteczek tRF z pętli D. Należy zauważyć, że proces nadekspresji nie był znacząco wydajny co może świadczyć o mechanizmach regulujących poziom tego białka odgórnie.

6.3 **Long Noncoding RNA Interactome: Detection, Characterization and Function.**

Kazimierczyk, M., Kasprowicz, M.K., Kasprzyk, K.E., Wrzesinski, J. (2020) Human Int. J. Mol. Sci. 21, 1027;

Wzrost ilości publikacji naukowych dotyczących długich niekodujących RNA (lncRNA) w biologii był zauważalny w ciągu ostatnich kilku lat. Dane pochodzące z PubMed wskazują na wyraźny wzrost liczby publikacji, które zawierają termin "lncRNA" w tytule. W roku 2013 liczba takich publikacji wynosiła 642, natomiast w roku 2020 osiągnęła imponujący wynik 5559. Ten rosnący trend stał się motywacją do przeprowadzenia przeglądowej publikacji, w której zebrano dostępną wiedzę na temat lncRNA. W niniejszej publikacji, przedstawiono pojęcie "interaktomu lncRNA" oraz zaproponowano techniki badawcze do badania tych interakcji. Rozwijając koncepcję interaktomu lncRNA, uwzględniono różne klasy RNA, zarówno kodujące, jak i niekodujące, DNA, białka, związki niskocząsteczkowe oraz krótkie peptydy. W kontekście oddziaływań lncRNA z RNA, szczególną uwagę poświęcono interakcjom z miRNA i mRNA.

Omawiając oddziaływanie miRNA-lncRNA należy wspomnieć teorię hipotezy ceRNA (konkurencyjne endogenne RNA). Zakłada ona, że lncRNA może działać jako "gąbka" dla miRNA, absorbując je z otoczenia i wpływając pośrednio na procesy regulowane przez konkretne miRNA. Te interakcje są kluczowe w kontekście wielu stanów patologicznych. Analizy potencjalnych oddziaływań miRNA-lncRNA opierają się często na analizach *in silico*, wykorzystujących miejsca wiązania miRNA do lncRNA. Narzędzia takie jak MARIO czy PARIS okazały się niezwykle przydatne. Jednakże, każde takie oddziaływanie wymaga potwierdzenia za pomocą eksperymentów in vivo. Do badania interakcji między RNA-RNA, przydatne są techniki takie jak wysokoprzepustowe sekwencjonowanie RNA izolowanego metodą immunoprecypitacji (HITS-CLIP) oraz metoda sieciowania z wykorzystaniem fotoaktywowanych rybonukleozydów połączona z immunoprecypitacją (PAR-CLIP). W kontekście oddziaływań z mRNA, lncRNA wpływa na proces składania mRNA oraz reguluje stabilność matrycowego RNA. Analizy bioinformatyczne wykorzystują narzędzia takie jak LncTar i RNAplex.

LncRNA mają zdolność do oddziaływania z DNA, co prowadzi do bezpośredniej regulacji ekspresji genów poprzez tworzenie potrójnych helis RNA-DNA-DNA. Analizy bioinformatyczne, przewidujące tworzenie się takich potrójnych helis, wykorzystują narzędzia jak GRIDseq, Triplexator czy LongTarget. Eksperymentalne metody wykrywania miejsc wiązania lncRNA do DNA obejmują izolację chromatyny przez oczyszczanie RNA (ChIRP), wytrącanie przez powinowactwo chromatyny oligo (ChOP), analizę hybrydyzacji wychwytu docelowego RNA (CHART) i oczyszczanie antysensownego RNA (RAP). Ze względu na

różnorodność długości lncRNA, mogą one przybierać różne struktury przestrzenne, co umożliwia szeroki zakres oddziaływań z białkami. W zależności od początkowego punktu wyjścia, przeprowadza się różne techniki badawcze, które obejmują utworzenie kompleksu białko-RNA, sieciowanie cząsteczek, a następnie ich immunoprecypitację i sekwencjonowanie specyficznych kwasów nukleinowych. Istnieją różne warianty tych technik, w zależności od potrzeb i kontekstu biologicznego. Do analiz bioinformatycznych sekwencjonowania zalecamy wykorzystanie wiarygodnych baz danych takich jak LNCipedia, NONCODE 2018 czy lncRNAdb.

LncRNA teoretycznie są niekodującymi RNA jednak niektóre z nich znajdujące się w cytozolu posiadają krótkie ramki odczytów (ORF). Udowodniono, że możliwość taka nie tylko istnieje, ale i zachodzi. Produkowane cząsteczki są to zazwyczaj krótkie peptydy. Analiza aktywnych ORF jest niezwykle trudna i wymaga połączenia zarówno narzędzi bioinformatycznych jak i eksperymentalnych. Długość lncRNA powyżej 200 nt powoduje, że bardzo często zawierają one w sobie kodony start AUG. Dobrym rozwiązaniem jest prowadzenie translacji in vitro a następnie detekcja np. metodą western blot czy o wiele czulszej metody spektrometrii mas. Innym podejściem jest sekwencjonowanie fragmentów związanych z polisomami i identyfikacja sekwencji ORF. Istotnym problemem jest stwierdzenie czy powstały peptyd pełni funkcję w kompleksie z macierzystym lncRNA czy osobno lub czy funkcję są w ogóle zbieżne. Literatura przytaczana przez nas dostarcza informacji o istnieniu każdego wariantu zdarzeń.

Możliwość przybierania różnych struktur przestrzennych otwiera wiele możliwości oddziaływania z grupą związków małocząsteczkowych. Niezwykle różnorodna grupa związków małocząsteczkowych zrzesza cząsteczki o masie mniejszej niż 1000Da i rozmiarze poniżej 1nm. Oddziaływania związków o bardzo małej masie na długie niekodujące RNA udowodniono dla najlepiej zbadanych lncRNA XIST oraz MALAT1. Badania nad wpływem małych cząsteczek prowadzone były z wykorzystaniem wysokowydajnej techniki przesiewowej związków zwanej Amplified Luminescent Proximity Homogeneous Assay. Zakres interakcji lncRNA z różnymi grupami cząstek jest ogromny i będzie on rósł wraz ze wzrostem wiedzy na temat struktur i właściwości lncRNA gdzie wciąż pozostaje wiele do odkrycia.

6.4 Long Non-Coding RNA Epigenetics.

Kazimierczyk, M., Wrzesinski, J. Int. J. Mol. Sci. 22(1, 6166;) (2021)

Epigenetyka staje się terminem, który podbił nie tylko umysły naukowców, ale coraz częściej przedostaje się do literatury popularno-naukowej. Wpływ epigenetyczny na stany patologiczne takie jak nowotwory czy choroby przewlekłe coraz częściej zaczyna się stawiać na genetycznymi uwarunkowaniami. Z racji na bogaty garnitur modyfikacji chemicznych RNA których odkryto dotychczas ponad 170 postanowiliśmy zbadać jak rosnący trend zainteresowania epigenetyką zmienił stan wiedzy o modyfikacjach epigenetycznych lncRNA. Modyfikacje nukleotydów angażują 3 typy białek. Wprowadzanie zmian epigenetycznych następuje za pośrednictwem enzymów zwanych writers. Zmiany wprowadzone do sekwencji odczytywane są przez białka zwane readers. Zmiany epigenetyczne mogą być usuwane za pośrednictwem enzymów zwanych erasers. Wśród modyfikacji najczęściej odnajdujemy cztery: 6-metyloadenozyna (m^6A), 5- metylocytydana (m^5C), pseudourydyna (Ψ), inozyna (I).

6-metyloadenozyna nanoszona jest przez: kompleks METTL3/METTL14, METTL16, METTL1. Należy zauważyć, że kompleks METTL3/METTL14 obiera za cel kodujące mRNA I wymaga szeregu białek regulujących wydajność i specyficzność. METTL16 obiera na cel zarówno kodujące jak I niekodujące RNA z kolei METTL1 pełni funkcję zarówno writer jak i reader. Białkami odczytującymi modyfikację m^6A są białka YTHDF1–3 i YTHDC1–2 dzięki domenie YT521-B. YTHDF1–3 są białkami o właściwościach helikazy. Białko YTHDC1 bierze udział w składaniu genów YTHDC2 stanowi podstawę strukturalną kompleksu wyciszającego chromosom X przez XIST. Białka znoszące modyfikację m^6A to głównie białko FTO i ALKBH5. Pierwsze usuwa ślad metylacji za pośrednictwem utleniania m^6A do N6-hydroksymetylodozyny lub N6-formyladenozyny które z uwagi na niską stabilność hydrolizują do adeniny. ALKBH5 katalizuje bezpośrednie przeniesienie grupy metylowej z adeniny.

Modyfikacja m^5C wystepuje tak samo w DNA jak i w RNA. Nanoszenie grupy metylowej na cytozynę odbywa się z udziałem metylotransferaz dwóch grup. Przedstawiciel rodziny metylotransferaz DNA a mianowicie DNMT2 może nanosić grupę metylową na RNA. Druga rodzina NSUN zawiera siedmioro członków z czego NSUN2 oddziałuje z mRNA i lncRNA. Białkiem pełniącym funkcję reader dla m^5C jest ALYREF które jest niezbędne do eksportu jądrowego do cytoplazmy. Brak białka znoszącego modyfikację m^5C jednak białka TET katalizują utlenianie m^5C do 5-hydroksymetylocytozyny (hm^5C).

Pseudourydyna pojawia się głownie w rRNA oraz w tRNA nadając charakter jednej z pętli. W obrębie komórki szacuje się, że pseudourydyna występuje w 5%. Tworzenie pseudourydyny następuję za pośrednictwem enzymów zwanych syntazami pseudourydyny (PUS) w reakcji izomeryzacji urydyny. Rodzina białek PUS uwzględnia 10 białek podzielonych na 5 rodzin. Pseudourydyna tworzy unikalne wiązanie glikozydowe C-C. Stabilność tego wiązania jest bardzo duża i nie odkryto białka znoszącego które można zakwalifikować jako eraser dla pseudourydyny. Podobny problem dotyczy białek typu reader gdyż podobieństwo pomiędzy U a Ψ jest wysokie.

Inozyna odróżnia się od adeniny grupą karbonylową zamiast grupy aminowej. Proces modyfikacji katalizowany jest przez deaminazę adenozyny działającą na RNA (ADAR). Rodzina ADAR u kręgowców składa się z 3 enzymówi ADAR1-3. Białka wspomnianej rodziny zawierają domenę aktywności deaminazy na C-końcu i domenę wiążącą dsRNA ze strony N-końcowej. Z uwagi na podobieństwo guanozyny i inozyny we właściwościach białka ADAR mogą katalizować reakcję substytucji A do G.

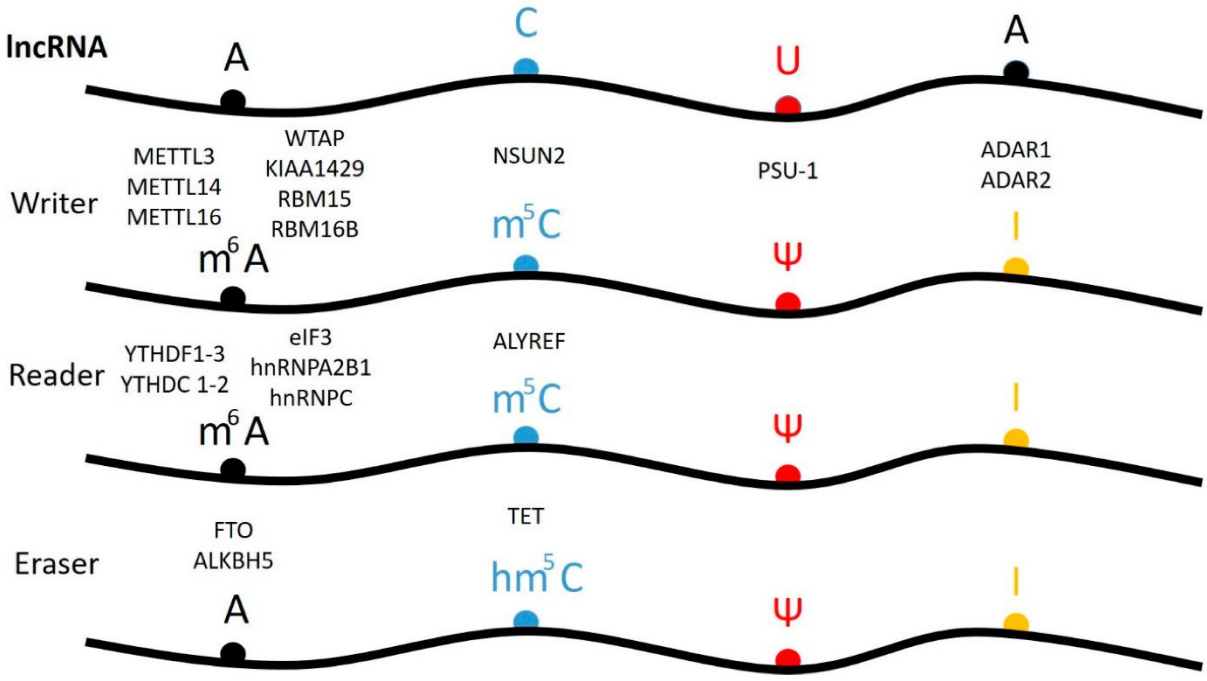


Figura 6. Schemat epigenetycznej modyfikacji lncRNA

Badanie modyfikacji mają swoje korzenie w 1957 roku, kiedy wykorzystywano chromatografię bibułową (TLC) do odkrycia pseudourydyny w całkowitym RNA u drożdży. Metoda TLC została ulepszona a jej wariant nazwany SCARLET. Nowy wariant TLC umożliwia lokalizację reszt m6A, m5C oraz pseudourydyny. Z racji na niską przepustowość aktualnie stosuje się połączenie ultrawydajnej chromatografi cieczowej z spektrometrią mas (UPLC–MS). Niektóre metody opierały się o wykorzystanie sekwencjonowania NGS gdzie wykorzystuje się

przeciwciała specyficzne dla modyfikacji, aby utworzyć bibliotekę a nastąpienie prowadzi się sekwencjonowanie. Istnieje szereg metod pośrednich w wykrywaniu modyfikacji bazujący o specyficzne działanie globalnego czynnika na daną modyfikację które następnie można wychwycić analizą wysoko przepustową. Przykładami takich metod są NOseq gdzie m^6A jest odporna na działanie kwasu azotowego a adenozyna bez modyfikacji ulega reakcji deaminacji.

Jedną z metod wykrywających modyfikację m^5C jest metoda AZA-se w której wykorzystuje się analog nukleotydowy 5-azacytydyny, aby wpłynąć na proces nanoszenia grupy metylowej na cytozyny przez białka typu writer. Lokalizację pseudourydyny identyfikuje się stosując metodę pośrednią wykorzystującą związki chemii karbodiimidów które w specyficznych warunkach selektywnie reagują (obrazują pseudourydynę.

Inozynę można identyfikować wykorzystując analizę porównawczą sekwencjonowania, gdzie w cDNA widoczna będzie zamiana A na G. Wykorzystując techniki bioinformatyczne możliwym jest odróżnienie takiej zamiany od wyników fałszywie pozytywnych lub mutacji typu SNP. Pewną rewolucję obiecuje technika sekwencjonowania NanoPore gdzie informacje o sekwencji odczytywane są jako zmiany napięcia w błonie. Modyfikowane nukleotydy zmieniają napięcie w innym stopniu niż niemodyfikowane. Jest to tylko niewielki wycinek zaprezentowanych metod analizy modyfikacji zawartych w publikacji.

Modyfikacje lncRNA mogą przyczynić się do zmiany funkcji pomiędzy modyfikowanym a niemodyfikowanym lncRNA. Najlepiej poznane lncRNA ulegają silnym modyfikacjom. XIST zaangażowany w inaktywację chromosome X posiada 78 miejsc m6A, 5 miejsc m5C i jedno miejsce Ψ. MALAT1 zaangażowany w powstawanie i progresję nowotworu płuc zawiera bardzo dużo modyfikacji m6A istotnych dla wiązania się z białkami. MALAT1 posiada również pięć miejsc m5C regulujących wiązanie z chromatyną. W HOTAIR (antysensowny RNA transkryptu HOX) identyfikowano 14 pojedynczych miejsc m6A oraz wysoce specyficzną metylację cytozyny w pozycji 1683. Inny między genowy lincRNA1281 zawiera ciekawą modyfikacją m6A na 3'-końcu mogącą działać jako przełącznik dla białek wiążących RNA. lincRNA1281 jest niezbędny do różnicowania embrionalnych komórek macierzystych myszy. Zmniejszony poziom m6A w lncRNA H19 spowodował zwiększenie żywotności komórek i zdolności przeciwapoptotycznej a został wywołany przez knockdown METTL3 I METTL14. Przytoczone przykłady są jedynie wybranymi spośród wielu znanych i ważnych przykładów, gdzie obecność modyfikacji odgrywa istotną role w kierowaniu funkcją danego lncRNA.

7 Bibliografia

1. Miescher F (1871) Ueber die chemische Zusmmensetzung der Eiterzellen. Medicinisch-chemische Untersuchungen 4: 441-460
2. Altmann R (1889) Ueber Nukleinsäuren. Arch f Anatomie u Physiol, str. 524–536
3. Levene PA, Jacobs WA (1909) Über die pentose in den nucleinsäure. Ber Deut Chem Ges 42: 3247–3251
4. Levene PA (1910) On the biochemistry of nucleic acids. J Am Chem Soc 32: 231-240
5. Levene PA, Mikeska LA, Mori T (1930) On the carbohydrate of thymonucleic acid. J Biol Chem 85: 785–787
6. Brachet J (1942) La localisation des acides pentosennucliiques dans les tissus animaux et les oeufs d'Amphibiens en voie de dcveloppement. Arch Biol 53: 207-257
7. Brenner, S., Jacob, F. & Meselson, M. Nature 190, 576–581 (1961).
8. Marx, V. How noncoding RNAs began to leave the junkyard. Nat Methods 19, 1167–1170 (2022). https://doi.org/10.1038/s41592-022-01627-8
9. Britten, R. J. & Davidson, E. H. Science 165, 349–357 (1969).
10. Andersen, J., Forst, SA, Zhao, K., Inouye, M. & Delihas, NJ Biol. chemia 264 , 17961-17970 (1989).
11. Wickens, M. & Takayama, K. Nature 367, 17–18 (1994).
12. ENCODE Project Consortium. Nature 489, 47–74 (2012).
13. Li J., Liu C. Coding or noncoding, the converging concepts of RNAs. Front. Genet. 2019;10:496. doi: 10.3389/fgene.2019.00496.
14. Fernandes JCR, Acuña SM, Aoki JI, Floeter-Winter LM, Muxel SM. Long Non-Coding RNAs in the Regulation of Gene Expression: Physiology and Disease. Noncoding RNA. 2019;5(1):17. Published 2019 Feb 17. doi:10.3390/ncrna5010017
15. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. RNA Biol. 2013;10(6):925-933. doi:10.4161/rna.24604
16. Bartel DP. Metazoan MicroRNAs. Cell. 2018;173(1):20-51. doi:10.1016/j.cell.2018.03.006
17. Palazzo AF, Lee ES. Non-coding RNA: what is functional and what is junk?. Front Genet. 2015;6:2. Published 2015 Jan 26. doi:10.3389/fgene.2015.00002
18. Lee H, Zhang Z, Krause HM. Long Noncoding RNAs and Repetitive Elements: Junk or Intimate Evolutionary Partners?. Trends Genet. 2019;35(12):892-902. doi:10.1016/j.tig.2019.09.006
19. Kellis, Manolis, Barbara Wold, Michael P. Snyder, Bradley E. Bernstein, Anshul Kundaje, Georgi K. Marinov, Lucas D. Ward, i in. „Defining functional DNA elements in the human genome". Proceedings of the National Academy of Sciences 111, nr 17 (29 January 2014): 6131–38. https://doi.org/10.1073/pnas.1318948111.
20. Hillenmeyer ME, Fung E, Wildenhain J, et al. The chemical genomic portrait of yeast: uncovering a phenotype for all genes. Science. 2008;320(5874):362-365. doi:10.1126/science.1150021
21. Wang T, Birsoy K, Hughes NW, et al. Identification and characterization of essential genes in the human genome. Science. 2015;350(6264):1096-1101. doi:10.1126/science.aac7041
22. Żmieńko, Agnieszka, i Anastasiia Satyr. „Sekwencjonowanie nanoporowe i jego zastosowanie w biologii". Postępy Biochemii 66, nr 3 (9 August 2020): 193-204. https://doi.org/10.18388/pb.2020_328.

23. Grosshans H, Filipowicz W. Molecular biology: the expanding world of small RNAs. Nature. 2008;451(7177):414-416. doi:10.1038/451414a
24. Commissioner, Office of the. „FDA Approves First-of-Its Kind Targeted RNA-Based Therapy to Treat a Rare Disease". FDA. FDA, 24 march 2020.
25. Kristen AV, Ajroud-Driss S, Conceição I, Gorevic P, Kyriakides T, Obici L. Patisiran, an RNAi therapeutic for the treatment of hereditary transthyretin-mediated amyloidosis. Neurodegener Dis Manag. 2019;9(1):5-23. doi:10.2217/nmt-2018-0033
26. Traber GM, Yu AM. RNAi-Based Therapeutics and Novel RNA Bioengineering Technologies. J Pharmacol Exp Ther. 2023;384(1):133-154. doi:10.1124/jpet.122.001234
27. Kim, YK. RNA therapy: rich history, various applications and unlimited future prospects. Exp Mol Med 54, 455–465 (2022). https://doi.org/10.1038/s12276-022-00757-5
28. Shi J, Zhou T, Chen Q. Exploring the expanding universe of small RNAs. Nat Cell Biol. 2022;24(4):415-423. doi:10.1038/s41556-022-00880-5
29. Rutz S, Scheffold A. Towards in vivo application of RNA interference - new toys, old problems. Arthritis Res Ther. 2004;6(2):78-85. doi:10.1186/ar1168
30. Chen J, Peng Y, Zhang H, et al. Off-target effects of RNAi correlate with the mismatch rate between dsRNA and non-target mRNA. RNA Biol. 2021;18(11):1747-1759. doi:10.1080/15476286.2020.1868680
31. Gantier MP, Williams BR. The response of mammalian cells to double-stranded RNA. Cytokine Growth Factor Rev. 2007;18(5-6):363-371. doi:10.1016/j.cytogfr.2007.06.016
32. Lam JK, Chow MY, Zhang Y, Leung SW. siRNA Versus miRNA as Therapeutics for Gene Silencing. Mol Ther Nucleic Acids. 2015;4(9):e252. Published 2015 Sep 15. doi:10.1038/mtna.2015.23
33. Juliano RL, Ming X, Nakagawa O. Cellular uptake and intracellular trafficking of antisense and siRNA oligonucleotides. Bioconjug Chem. 2012;23(2):147-157. doi:10.1021/bc200377d
34. Cardarelli, F., Digiacomo, L., Marchini, C. et al. The intracellular trafficking mechanism of Lipofectamine-based transfection reagents and its implication for gene delivery. Sci Rep 6, 25879 (2016). https://doi.org/10.1038/srep25879
35. Sajid MI, Moazzam M, Kato S, Yeseom Cho K, Tiwari RK. Overcoming Barriers for siRNA Therapeutics: From Bench to Bedside. Pharmaceuticals (Basel). 2020;13(10):294. Published 2020 Oct 7. doi:10.3390/ph13100294
36. Singh, Yashveer, Pierre Murat, i Eric Defrancq. „Recent Developments in Oligonucleotide Conjugation". Chemical Society Reviews 39, nr 6 (25 maj 2010): 2054–70. https://doi.org/10.1039/B911431A.
37. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science. 2001;294(5543):853-858. doi:10.1126/science.1064921
38. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. Front Endocrinol (Lausanne). 2018;9:402. Published 2018 Aug 3. doi:10.3389/fendo.2018.00402
39. Meijer HA, Smith EM, Bushell M. Regulation of miRNA strand selection: follow the leader?. Biochem Soc Trans. 2014;42(4):1135-1140. doi:10.1042/BST20140142
40. Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias [published correction appears in Cell. 2003 Nov 14;115(4):505]. Cell. 2003;115(2):209-216. doi:10.1016/s0092-8674(03)00801-8

41. Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature. 2007;448(7149):83-86. doi:10.1038/nature05983
42. Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature. 2010;465(7298):584-589. doi:10.1038/nature09092
43. Baumann V, Winkler J. miRNA-based therapies: strategies and delivery platforms for oligonucleotide and non-oligonucleotide agents. Future Med Chem. 2014;6(17):1967-1984. doi:10.4155/fmc.14.116
44. Pajares MJ, Alemany-Cosme E, Goñi S, Bandres E, Palanca-Ballester C, Sandoval J. Epigenetic Regulation of microRNAs in Cancer: Shortening the Distance from Bench to Bedside. Int J Mol Sci. 2021;22(14):7350. Published 2021 Jul 8. doi:10.3390/ijms22147350
45. Jungers CF, Djuranovic S. Modulation of miRISC-Mediated Gene Silencing in Eukaryotes. Front Mol Biosci. 2022;9:832916. Published 2022 Feb 14. doi:10.3389/fmolb.2022.832916
46. Nakanishi K. Anatomy of four human Argonaute proteins. Nucleic Acids Res. 2022;50(12):6618-6638. doi:10.1093/nar/gkac519
47. Ozsolak F, Poling LL, Wang Z, et al. Chromatin structure analyses identify miRNA promoters. Genes Dev. 2008;22(22):3172-3183. doi:10.1101/gad.1706508
48. Liu T, Cai J, Cai J, Wang Z, Cai L. EZH2-miRNA Positive Feedback Promotes Tumor Growth in Ovarian Cancer. Front Oncol. 2021;10:608393. Published 2021 Feb 25. doi:10.3389/fonc.2020.608393
49. Ramassone A, Pagotto S, Veronese A, Visone R. Epigenetics and MicroRNAs in Cancer. Int J Mol Sci. 2018;19(2):459. Published 2018 Feb 3. doi:10.3390/ijms19020459
50. Liu, Y., Dou, M., Song, X. et al. The emerging role of the piRNA/piwi complex in cancer. Mol Cancer 18, 123 (2019). https://doi.org/10.1186/s12943-019-1052-9
51. Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr Biol. 2001;11(13):1017-1027. doi:10.1016/s0960-9822(01)00299-8
52. Houwing S, Kamminga LM, Berezikov E, et al. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell. 2007;129(1):69-82. doi:10.1016/j.cell.2007.03.026
53. Wu X, Pan Y, Fang Y, et al. The Biogenesis and Functions of piRNAs in Human Diseases. Mol Ther Nucleic Acids. 2020;21:108-120. doi:10.1016/j.omtn.2020.05.023
54. Perera BPU, Tsai ZT, Colwell ML, et al. Somatic expression of piRNA and associated machinery in the mouse identifies short, tissue-specific piRNA. Epigenetics. 2019;14(5):504-521. doi:10.1080/15592294.2019.1600389
55. Tian Y, Simanshu DK, Ma JB, Patel DJ. Structural basis for piRNA 2'-O-methylated 3'-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains. Proc Natl Acad Sci U S A. 2011;108(3):903-910. doi:10.1073/pnas.1017762108
56. Thompson HJ, Jiang W, Zhu Z. Candidate mechanisms accounting for effects of physical activity on breast carcinogenesis. IUBMB Life. 2009;61(9):895-901. doi:10.1002/iub.233
57. Jia DD, Jiang H, Zhang YF, Zhang Y, Qian LL, Zhang YF. The regulatory function of piRNA/PIWI complex in cancer and other human diseases: The role of DNA methylation. Int J Biol Sci. 2022;18(8):3358-3373. Published 2022 May 9. doi:10.7150/ijbs.68221
58. Özata DM, Yu T, Mou H, et al. Evolutionarily conserved pachytene piRNA loci are highly divergent among modern humans. Nat Ecol Evol. 2020;4(1):156-168. doi:10.1038/s41559-019-1065-1
59. Rosenkranz D, Zischler H, Gebert D. piRNAclusterDB 2.0: update and expansion of the piRNA cluster database. Nucleic

Acids Res. 2022;50(D1):D259-D264. doi:10.1093/nar/gkab622
60. Chen S, Ben S, Xin J, et al. The biogenesis and biological function of PIWI-interacting RNA in cancer. J Hematol Oncol. 2021;14(1):93. Published 2021 Jun 12. doi:10.1186/s13045-021-01104-3
61. Czech B, Hannon GJ. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. Trends Biochem Sci. 2016;41(4):324-337. doi:10.1016/j.tibs.2015.12.008
62. Han BW, Wang W, Li C, Weng Z, Zamore PD. Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. Science. 2015;348(6236):817-821. doi:10.1126/science.aaa1264
63. Stein CB, Genzor P, Mitra S, et al. Decoding the 5' nucleotide bias of PIWI-interacting RNAs. Nat Commun. 2019;10(1):828. Published 2019 Feb 19. doi:10.1038/s41467-019-08803-z
64. Ninova M, Chen YA, Godneeva B, et al. Su(var)2-10 and the SUMO Pathway Link piRNA-Guided Target Recognition to Chromatin Silencing. Mol Cell. 2020;77(3):556-570.e6. doi:10.1016/j.molcel.2019.11.012
65. Peng L, Song L, Liu C, et al. piR-55490 inhibits the growth of lung carcinoma by suppressing mTOR signaling. Tumour Biol. 2016;37(2):2749-2756. doi:10.1007/s13277-015-4056-0.
66. Tan L, Mai D, Zhang B, et al. PIWI-interacting RNA-36712 restrains breast cancer progression and chemoresistance by interaction with SEPW1 pseudogene SEPW1P RNA. Mol Cancer. 2019;18(1):9. Published 2019 Jan 12. doi:10.1186/s12943-019-0940-3
67. Liu X, Zheng J, Xue Y, et al. PIWIL3/OIP5-AS1/miR-367-3p/CEBPA feedback loop regulates the biological behavior of glioma cells [published correction appears in Theranostics. 2022 Jan 21;12(4):1732-1733]. Theranostics. 2018;8(4):1084-1105. Published 2018 Jan 1. doi:10.7150/thno.21740
68. Cole C, Sobala A, Lu C, et al. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. RNA. 2009;15(12):2147-2160. doi:10.1261/rna.1738409
69. Kumar P, Anaya J, Mudunuri SB, Dutta A. Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. BMC Biol. 2014;12:78. Published 2014 Oct 1. doi:10.1186/s12915-014-0078-0
70. Kumar P, Kuscu C, Dutta A. Biogenesis and Function of Transfer RNA-Related Fragments (tRFs). Trends Biochem Sci. 2016;41(8):679-689. doi:10.1016/j.tibs.2016.05.004
71. Park EJ, Kim T-H. Fine-Tuning of Gene Expression by tRNA-Derived Fragments during Abiotic Stress Signal Transduction. International Journal of Molecular Sciences. 2018; 19(2):518. https://doi.org/10.3390/ijms19020518
72. Li S, Xu Z, Sheng J. tRNA-Derived Small RNA: A Novel Regulatory Small Non-Coding RNA. Genes (Basel). 2018;9(5):246. Published 2018 May 10. doi:10.3390/genes9050246
73. Luo S, He F, Luo J, et al. Drosophila tsRNAs preferentially suppress general translation machinery via antisense pairing and participate in cellular starvation response. Nucleic Acids Res. 2018;46(10):5250-5268. doi:10.1093/nar/gky189
74. Lyons SM, Gudanis D, Coyne SM, Gdaniec Z, Ivanov P. Identification of functional tetramolecular RNA G-quadruplexes derived from transfer RNAs [published correction appears in Nat Commun. 2017 Dec 5;8(1):2020]. Nat Commun. 2017;8(1):1127. Published 2017 Oct 24. doi:10.1038/s41467-017-01278-w
75. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived

RNA fragments (tRFs). Genes Dev. 2009;23(22):2639-2649. doi:10.1101/gad.1837609
76. Emara MM, Ivanov P, Hickman T, et al. Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. J Biol Chem. 2010;285(14):10959-10968. doi:10.1074/jbc.M109.077560
77. Yeung ML, Bennasser Y, Watashi K, Le SY, Houzet L, Jeang KT. Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid. Nucleic Acids Res. 2009;37(19):6575-6586. doi:10.1093/nar/gkp707
78. La Ferlita A, Alaimo S, Veneziano D, et al. Identification of tRNA-derived ncRNAs in TCGA and NCI-60 panel cell lines and development of the public database tRFexplorer. Database (Oxford). 2019;2019:baz115. doi:10.1093/database/baz115
79. Saikia M, Jobava R, Parisien M, et al. Angiogenin-cleaved tRNA halves interact with cytochrome c, protecting cells from apoptosis during osmotic stress. Mol Cell Biol. 2014;34(13):2450-2463. doi:10.1128/MCB.00136-14
80. Mattick, J.S., Amaral, P.P., Carninci, P. et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat Rev Mol Cell Biol 24, 430–447 (2023). https://doi.org/10.1038/s41580-022-00566-8
81. Lu Z, Zhang QC, Lee B, et al. RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. Cell. 2016;165(5):1267-1279. doi:10.1016/j.cell.2016.04.028
82. Hacisuleyman E, Shukla CJ, Weiner CL, Rinn JL. Function and evolution of local repeats in the Firre locus. Nat Commun. 2016;7:11021. Published 2016 Mar 24. doi:10.1038/ncomms11021
83. Johnson R, Guigó R. The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. RNA. 2014;20(7):959-976. doi:10.1261/rna.044560.114
84. Cheetham SW, Faulkner GJ, Dinger ME. Overcoming challenges and dogmas to understand the functions of pseudogenes. Nat Rev Genet. 2020;21(3):191-201. doi:10.1038/s41576-019-0196-1
85. Rinn JL, Wang JK, Allen N, et al. A dermal HOX transcriptional program regulates site-specific epidermal fate. Genes Dev. 2008;22(3):303-307. doi:10.1101/gad.1610508
86. Connerty P, Lock RB, de Bock CE. Long Non-coding RNAs: Major Regulators of Cell Stress in Cancer. Front Oncol. 2020;10:285. Published 2020 Mar 20. doi:10.3389/fonc.2020.00285
87. Hacisuleyman E, Shukla CJ, Weiner CL, Rinn JL. Function and evolution of local repeats in the Firre locus. Nat Commun. 2016;7:11021. Published 2016 Mar 24. doi:10.1038/ncomms11021
88. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science. 2008;322(5902):750-756. doi:10.1126/science.1163045
89. Merry CR, Forrest ME, Sabers JN, et al. DNMT1-associated long non-coding RNAs regulate global gene expression and DNA methylation in colon cancer. Hum Mol Genet. 2015;24(21):6240-6253. doi:10.1093/hmg/ddv343
90. Khalil AM, Guttman M, Huarte M, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009;106(28):11667-11672. doi:10.1073/pnas.0904715106
91. Souaid, C., Bloyer, S. & Noordermeer, D. in Nuclear Architecture and Dynamics Vol. 2 (eds Christophe, L. & Jean-Marc, V.) Ch. 19, 435–456 (Academic Press, 2018).
92. Ducoli L, Agrawal S, Sibler E, et al. LETR1 is a lymphatic endothelial-specific lncRNA governing cell proliferation and migration through KLF4 and SEMA3C. Nat

Commun. 2021;12(1):925. Published 2021 Feb 10. doi:10.1038/s41467-021-21217-0
93. Lewandowski JP, Lee JC, Hwang T, et al. The Firre locus produces a trans-acting RNA molecule that functions in hematopoiesis. Nat Commun. 2019;10(1):5137. Published 2019 Nov 13. doi:10.1038/s41467-019-12970-4
94. Groff AF, Barutcu AR, Lewandowski JP, Rinn JL. Enhancers in the Peril lincRNA locus regulate distant but not local genes. Genome Biol. 2018;19(1):219. Published 2018 Dec 11. doi:10.1186/s13059-018-1589-8
95. Allou L, Balzano S, Magg A, et al. Non-coding deletions identify Maenli lncRNA as a limb-specific En1 regulator. Nature. 2021;592(7852):93-98. doi:10.1038/s41586-021-03208-9
96. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol. 2017;18(5):285-298. doi:10.1038/nrm.2017.7
97. Niklas KJ, Dunker AK, Yruela I. The evolutionary origins of cell type diversification and the role of intrinsically disordered proteins. J Exp Bot. 2018;69(7):1437-1446. doi:10.1093/jxb/erx493
98. Roden C, Gladfelter AS. RNA contributions to the form and function of biomolecular condensates. Nat Rev Mol Cell Biol. 2021;22(3):183-195. doi:10.1038/s41580-020-0264-6
99. Chen Y, Long W, Yang L, et al. Functional Peptides Encoded by Long Non-Coding RNAs in Gastrointestinal Cancer. Front Oncol. 2021;11:777374. Published 2021 Nov 23. doi:10.3389/fonc.2021.777374
100. Hartford CCR, Lal A. When Long Noncoding Becomes Protein Coding. Mol Cell Biol. 2020;40(6):e00528-19. Published 2020 Feb 27. doi:10.1128/MCB.00528-19
101. Matsumoto A, Pasut A, Matsumoto M, et al. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature. 2017;541(7636):228-232. doi:10.1038/nature21034
102. Xing, Jiani, Haizhou Liu, Wei Jiang, i Lihong Wang. „LncRNA-Encoded Peptide: Functions and Predicting Methods". Frontiers in Oncology 10 (2021).
103. D'Lima NG, Ma J, Winkler L, et al. A human microprotein that interacts with the mRNA decapping complex. Nat Chem Biol. 2017;13(2):174-180. doi:10.1038/nchembio.2249
104. Dawoud A, Ihab Zakaria Z, Hisham Rashwan H, Braoudaki M, Youness RA. Circular RNAs: New layer of complexity evading breast cancer heterogeneity. Noncoding RNA Res. 2022;8(1):60-74. Published 2022 Oct 12. doi:10.1016/j.ncrna.2022.09.011
105. Ragan C, Goodall GJ, Shirokikh NE, Preiss T. Insights into the biogenesis and potential functions of exonic circular RNA. Sci Rep. 2019;9(1):2048. Published 2019 Feb 14. doi:10.1038/s41598-018-37037-0
106. Huang A, Zheng H, Wu Z, Chen M, Huang Y. Circular RNA-protein interactions: functions, mechanisms, and identification. Theranostics. 2020;10(8):3503-3517. Published 2020 Feb 10. doi:10.7150/thno.42174
107. Li B, Zhu L, Lu C, et al. circNDUFB2 inhibits non-small cell lung cancer progression via destabilizing IGF2BPs and activating anti-tumor immunity. Nat Commun. 2021;12(1):295. Published 2021 Jan 12. doi:10.1038/s41467-020-20527-z
108. Yu T, Ran L, Zhao H, et al. Circular RNA circ-TNPO3 suppresses metastasis of GC by acting as a protein decoy for IGF2BP3 to regulate the expression of MYC and SNAIL. Mol Ther Nucleic Acids. 2021;26:649-664. Published 2021 Aug 28. doi:10.1016/j.omtn.2021.08.029
109. Pamudurti NR, Bartok O, Jens M, et al. Translation of CircRNAs. Mol Cell. 2017;66(1):9-21.e7. doi:10.1016/j.molcel.2017.02.021
110. Luo N, Liu S, Li X, Hu Y, Zhang K. Circular RNA circHIPK3 promotes breast cancer progression via sponging MiR-326.

Cell Cycle. 2021;20(13):1320-1333. doi:10.1080/15384101.2021.1939476

Prace wchodzące w skład rozprawy doktorskiej

Biochemical and Biophysical Research Communications 520 (2019) 514–519

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

tRNA-derived fragments from the *Sus scrofa* tissues provide evidence of their conserved role in mammalian development

Marek Kazimierczyk [a], Agata Jędroszkowiak [a, b], Dorota Kowalczykiewicz [a], Maciej Szymański [c], Barbara Imiołczyk [a], Jerzy Ciesiołka [a], Jan Wrzesiński [a, *]

[a] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland
[b] Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland
[c] Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznań, Poland

ARTICLE INFO

Article history:
Received 18 September 2019
Received in revised form
4 October 2019
Accepted 5 October 2019
Available online 11 October 2019

Keywords:
Mammalian tRNA fragments
Tissue specificity
Northern blot
PIWIL4 protein
Small RNA binding

ABSTRACT

The recently discovered group of noncoding RNAs, which are fragments of tRNA molecules (tRFs), has not been fully characterized and its potential functions still require investigation. Porcine tRFs were characterized and compared to mouse and human tRFs. Two tRFs, 5′ 32–33 nt and 3′ 41–42 nt that are derived from the mature tRNAVal(CAC) and tRNAGly(GCC) were detected with the use of bioinformatics and the Northern blot method. The abundance of these tRFs in the case of *Sus scrofa* is restricted to the ovary and the kidney. The same tRFs were found in human cancer cells and in mouse sperm, circulating blood and its serum. The binding of selected sncRNAs (piRNA, 5'tRFVal(CAC) and miRNA) to the overexpressed PAZ domain of the PIWIL4 protein was also studied. It is noteworthy that porcine 5'tRFVal(-CAC) and human 5'tRFVal(CAC)as well as 5'tRFGly(GCC) are bound to the PIWIL4 protein. The potential role of the analyzed tRFs in the development of mammals is also discussed.

1. Introduction

The family of sncRNAs detected in mammalian tissues contains several RNA groups, including miRNA, siRNA, piRNA and tRF [1]. These sncRNAs are characterized by variable sizes, different biogeneses and an array of mechanisms which control gene expression.

miRNAs are coded in the genome and are involved in gene regulation [2]. One of the strands of 21–23 nt long miRNAs binds to mRNAs, mainly to the 5' UTR regions. The miRNA-mRNA duplexes are then recognized by the RISC complex, which contains 1–4 AGO proteins, and degraded.

piRNAs are the largest group of sncRNAs and are mainly present in germ cells [3]. In mammalian organisms, repetitive elements of the genome, such as transposons, are the main source of piRNA. The piRNAs, which are obtained by the "ping-pong mechanism", bind to PIWI proteins. piRNAs bound to PIWI proteins interact with their target mRNAs through base pairing and the targeted mRNAs are then cleaved by the endonuclease activity of PIWI proteins.

tRF RNAs, a newly discovered group of sncRNAs, are essentially tRNA fragments [4]. Although the role tRNA plays in translation has been known since the late 1950s, tRFs were first discovered in human cells in 2009 [5]. To date, tRFs have been detected in many organisms [4,6,7].

tRFs are primarily formed by the cleavage of tRNA molecules in the single-stranded loop region, such as in the D, anticodon and TψC loops, by specific nucleases with different biogenesis pathways. Half-tRNA fragments (30–35 nt) are generated by angiogenin nuclease [4,8]. Shorter 23 nt tRFs are products of Dicer tRNA cleavage [9]. Moreover, tRF halves generated from some tRNAs under stress conditions regulate translation repression [4,6]. The transfection of human cells with oligomers mimicking 5'tRFAla(AGC) and 5'tRFCys(GCA) strongly impacts translation [10]. In addition, the tRF family has been found to be involved in tumor suppression [11].

RNA usually acts in the cell in complex with proteins [2,3]. Small RNAs interact with Argonaute family proteins, which can be divided into two clades: AGO and PIWI proteins. miRNA and siRNA interact with AGO proteins to form RISC complex in order to silence gene expression, either by cleaving the mRNA or by translational repression.

* Corresponding author. Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14 str. 61-704, Poznan, Poland.
E-mail address: wrzesinj@ibch.poznan.pl (J. Wrzesiński).

https://doi.org/10.1016/j.bbrc.2019.10.062
0006-291X/

The second clade of Argonaute proteins contains the PIWI proteins; these differ from AGO proteins because their expression is primarily restricted to germ cells [3]. To date, 2–4 different PIWI protein have been identified in various organisms, ranging from *Drosophila* to humans. In the case of *S. scrofa*, three PIWI proteins, PIWIL1, PIWIL2 and PIWIL4 have been identified [12]. Due to their large size of approximately 100 kDa, the PIWI proteins have a domain structure comprising the PIWI, MID and PAZ domains. The N-terminal PAZ domain is associated with RNA binding. The C-terminal PIWI domain has nucleolytic properties that degrade the target RNA.

2. Materials and methods

2.1. RNA isolation

Presented studies were conducted with the permission of the National Animal Experimentation Ethics Committee, Local Committee in Poznan.

RNA isolation from 8-month-old female ovaries was carried out using the TRIzol reagent as previously described [7,12]. Briefly, 150 mg of the tissue, which was obtained postmortem, was covered with 1.5 ml of TRIzol reagent. After homogenization, the solution was mixed with 300 μl of chloroform and kept for 20 min at RT. After incubation, the solution was centrifuged at 12,000 rpm for 15 min at 4 °C to allow the two phases to separate. The RNA-containing upper phase was transferred to a new tube and mixed with isopropanol. The solution was incubated for 10 min at RT and then centrifuged at 12,000 rpm for 10 min at 4 °C. The supernatant was removed, and the resulting precipitate was washed with 1.5 ml of 75% ethanol. After centrifugation, the ethanol was removed and the pellet was dried and dissolved in 30 μl of water. The concentration of the resulting RNA solution was measured spectrophotometrically.

2.2. In vitro transcription

All the small RNAs used for the analysis of their interactions with the PAZ domain of the PIWIL4 protein in the EMSA test were obtained using the T7 *in vitro* transcription method, as previously described [13]. The double-stranded DNA template required for transcription was obtained through the PCR technique using the appropriate oligomers (Table S2).

2.3. DNA or RNA labeling

The RNA labeling reaction mixture contained the RNA, buffer (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM of DTT, 0.1 mM spermidine), 10 U T4 polynucleotide kinase, 20 μCi of [γ-32ATP] with an activity of 4000–5000 Ci/mmol (Hartmann Analytic, Braunschweig, Germany) and 20 pmol of DNA or RNA. After an hour long incubation at 37 °C, the products were purified using a G25 column (Sigma-Aldrich, St.Louis, MO, USA), precipitated, centrifuged and dried. The level of radioactivity in the labeled molecules was measured using a scintillation counter.

2.4. Northern blotting

The total RNA isolated from the different *S. scrofa* tissues was separated on 12% (w/v) polyacrylamide gels, electrotransferred onto Hybond-Nnylon membranes (GE Healthcare, Chicago, IL, USA), crosslinked with UV light (120 mJ/cm^2), and hybridized to ^{32}P labeled DNA probes at 42 °C in perfectHyb^{TM}Plus solution. Two washing cycles (10 min each) were performed at 42 °C in 2xSSC, followed by one cycle (30 min) at 42 °C in a mixture of 2xSSC and 0.1% sodium dodecyl sulfate.

2.5. Sequencing data analysis

sncRNA sequencing data deposited in the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under accession no. GSE57414 and GSE76156 were applied to characterize the *S. scrofa* ovary and kidney tRFs [7,14]. The bioinformatics analysis was carried out as previously described [7].

2.6. Overexpression of the PAZ domain of the PIWIL4 protein

cDNA was obtained with the Transcriptor High Fidelity cDNA Synthesis Kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. 20 mg of ovary RNA, 2.5 mM of oligo dT primer and 60 mM of hexamers were denatured for 10 min at 65 °C. Next, Roche reaction buffer, 1 mM of dNTP, 5 mM of DTT, 20 U of RNase inhibitor and 10 U of reverse transcriptase were added, and the reaction was allowed to proceed for 30 min at 50 °C. The reaction was stopped by incubation at 85 °C for 5 min. The coding sequence for the PAZ domain of the PIWIL4 protein was amplified using Pfu DNA polymerase. The PCRs were carried out using forward (CACCATGAGTGGACGGGCCCGCGTG) and reverse (TCACAGGTAGAAGAGA GAATTGGCTAATTCC) primers.

The PAZ domain insert and the pET151/D-TOPO vector were mixed at a 2:1 M ratio, incubated at RT for 5 min and used to transform chemically competent TOP10 *E. coli* cells (Novagen, Madison, WI, USA). The resulting sequence - verified pET151/D/PAZ plasmid was used to transform the *E. coli* expression strain Rosetta (DE3) pLysS (Novagen). The bacteria were grown in LB medium with ampicillin and chloramphenicol at 37 °C to OD 0.8. The temperature was reduced to 18 °C, protein production was induced with 0.5 mM of IPTG and continued for 5 h. Next, the bacteria were centrifuged at 6700×*g* for 10 min at 4 °C and the resulting pellet was suspended in lysis buffer (500 mM NaCl, 50 mM Hepes pH 8.0, 20 mM imidazole, 10% glycerol, 5 mM TCEP, 2% sarcosyl). The bacteria were disrupted by sonication for 4 min (2 s pulse, 20 s intervals), after which they were centrifuged for 30 min at 31,000×*g* and 4 °C. The cleared supernatant was loaded onto a Ni-NTA column that had been equilibrated in washing/binding buffer (500 mM NaCl, 50 mM Hepes pH 8.0, 20 mM imidazole, 10% glycerol, 2 mM TCEP). After extensive washing the His-tagged protein was eluted (500 mM NaCl, 50 mM Hepes pH 8.0, 300 mM imidazole, 10% glycerol, 2 mM TCEP) and the His tag was removed by applying TEV protease. The protein was concentrated with an Amicon 10 kDa cutoff concentrator (Merck, Darmstadt, Germany), filtered, and injected onto a Superdex HiLoad 200 (GE Healthcare, Chicago, IL, USA) size exclusion column that had been equilibrated with buffer (20 mM Hepes pH 7.4, 100 mM KCl, 10% glycerol, 1 mM TCEP). Each step of the purification process was analyzed on 14% SDS-PAGE.

2.7. Analysis of PAZ-RNA protein domain complexes by delayed migration on polyacrylamide gel

RNA radiolabeled on its 5′ end (60000–170000 c.p.m.) was suspended in glycerol-free protein storage buffer, denatured for 2 min at 90 °C and cooled for 15 min at room temperature. The mixture was supplemented with $MgCl_2$ to a final concentration of 1 mM and incubated for 10 min at 37 °C to ensure proper RNA folding.

The PAZ domain protein was diluted to a concentration range of 0.17–180 mM in glycerol-free protein storage buffer. The RNA and the PAZ domain protein samples were combined and incubated for 20 min at room temperature. Next, 50% glycerol was added to the

samples at a ratio of 10:1 and applied to the polyacrylamide gel at RT, with a constant current of 10 mA. Following electrophoresis, the gel was transferred to Whatman 3 M paper, dried and the electrophoresis results were visualized using a Fuji BAS-5000 radioactivity scanner.

3. Results and discussion

3.1. Characterization of mammalian models

Mouse and cell culture models are frequently used to study the role of mammalian tRFs. The advantage of using them is the short breeding time of mice or cell cultures and their well-known genetics. The disadvantages are that the mouse physiology and biochemistry differs from that of humans. Recently, *S. scrofa* has been recognized as an excellent animal model for biological and biochemical studies. Pigs are more closely related to humans in terms of pathology, pharmacology, biochemistry, and physiology than fruit flies (*D. melanogaster*), nematodes (*C. elegants*), amphibians (*X. leavis*) or rodents (*M. muscullus, R. norvegicus)*, all of which are frequently used as research models [15,16]. For this reason, pigs have been used in medical research for over 30 years, and are what's known as translational research models.

3.2. The presence of tRFs in porcine, human and mouse tissues/cells

The previously applied high-throughput sequencing and bioinformatic analysis of the sncRNA libraries from *S. scrofa* gonads enabled the identification of three RNA families: piRNA, miRNA and tRF [7]. In gonads, the tRF family was limited to tRNA halves, such as 5'tRFGly(GCC), 5'tRFVal(CAC), 5'tRFHis(GUG), 5'tRFLys(CUU) and 5'tRFGlu(UUC). Two tRFs, tRFGly(GCC) and tRFVal(CAC) were prevalent, with over 20,000 counts, while other tRFs were significantly less numerous (500–1.000 counts). In this paper, the bioinformatic analysis results were supplemented with experimental data. As had been previously predicted, ovary-specific tRFs were confirmed by Northern blot analysis and their abundance was the greatest in 5'tRFGly(GCC), 5'tRFVal(CAC), and 5'tRFHis(GUG) (Fig. 1A).

The question still remained as to whether these tRFs are present in other tissues. Two probes that bound to the 5′- or 3′-region of the tRNAVal(CAC) target were synthesized to evaluate the nature of the potential species resulting from tRNA processing (e.g., tRFs) (Fig. S1, Table S1). First, the ^{32}P-labeled 31-nucleotide probe was hybridized to the 5′ end of the tRNAVal(CAC) sequence from G1 to C31 to detect the 3′ tRNA cleavage (probe 1, Fig. S1). In the RNA pools isolated from the stomach, the liver, the lung, the kidney, the brain, and the muscles, the hybridization signal was restricted only to the kidney-specific tRFVal(CAC) (Fig. 1B). The size of the detected 5'tRFVal(CAC) was 34 and 35 nt, indicating that the tRNA molecule was cut in the anticodon loop (Fig. 1D, Fig. S1). In addition, the length of the detected 5'tRFVal(CAC) was similar to that of the tRF in the ovary (32 nt) according to bioinformatic analysis [17]. Small discrepancies in terms of size occurred between the polyacrylamide gel separation, Northern blot detection (34 and 35 nt) and bioinformatics analyses (32 nt), but this difference can be explained by the shifted mobility of the tRNA fragments during electrophoresis.

Subsequently, to analyze the 5′ part of tRFVal(CAC), a 31-nucleotide ^{32}P-labeled probe that hybridized to the tRNA molecules between nucleotides G41-C72 was designed (probe 6, Fig. S1). There are two cleavage sites at positions 43 and 44 of tRNAVal(CAC) (Fig. 1C). The cleavage sites precisely coincide with those identified using the previous 5' probe (Fig. 1B).

In the next step, the analysis was extended to other tRFs. The previous *S. scrofa* ovary and kidney sequencing data deposited in GEO was assessed [7,14). The bioinformatics analysis revealed the presence of tRNA fragments in both tissues with 34–35% abundance (Fig. 1E). Further analysis using Northern blotting with appropriate probes (Data T1) clearly showed the presence of tRF molecules in the kidney tissue that were also found in the *S. scrofa* ovary tissue (Fig. 1F). The relative ratios of tRF/tRNA cleavage for 5'tRFVal(CAC), 5'tRFHis(GUG), 5'tRFLys(CUU) and 5'tRFGlu(UUC) were very similar and ranged from 6% to 10%. However, the cleavage ratio for tRFGly(GCC), which was calculated using the Northern blotting results, was twice as high.

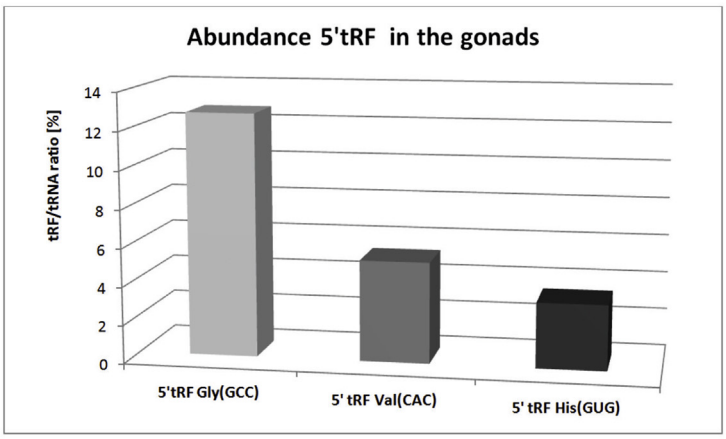
In summary, in the *S. scrofa* ovary and kidney in stress-free conditions, only two tRNA halves generated from the same tRNA were detected (Fig. 1D). The first tRNA half was 5'tRFVal(CAC) with a length of 32 and 33 nucleotides and the second tRNA half was 3'tRFVal(CAC) with a size of 43 and 44 nucleotides (Fig. 1B–D). The size of the 3′ tRF indicated that it came from a mature tRNA molecule containing the CCA sequence from the 3′ end. The CCA sequence is not template-coded in the gene but is added post-transcriptionally in the cytoplasm by the tRNA-nucleotidyl transferase enzyme [17].

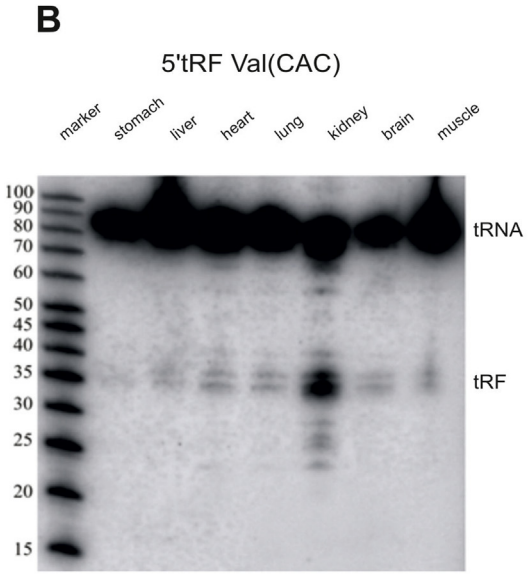
In the human genome, 635 tRNA genes and 321 coded tRNAs involved in tRF biogenesis have been identified [http://trna.bioinf.uni-leipzig.de]. In addition, 235 tRNA genes generate 5′ tRFs, and 243 tRNA genes 3′ tRFs. In this study, we found *S. scrofa* 5′ and 3′ tRFs which were restricted to the gonads and kidney. The bioinformatic analysis of the deep sequencing of the library obtained after the immunoprecipitation of PIWI/piRNA complexes from human breast cancer MDAMB231 cells and from mouse sperm cells has shown the same set of tRFs as the ones found in porcine gonads and kidney with similar sizes of 27 and 28 nucleotides [18]. In mice, 90% of the tRFs in the circulating blood and serum consisted of 5′tRFGly(GCC) and 5′tRFVal(CAC) [19]. However, in other organisms and tissues different tRFs were observed [4,6,7].

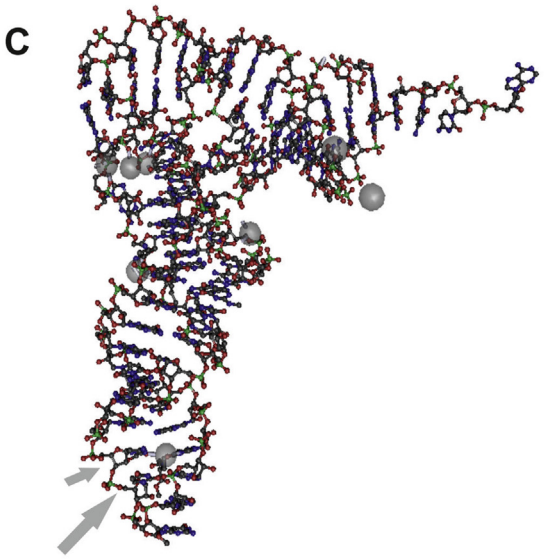
Thus, the question arises as to what is responsible for the occurrence of a specific set of tRFs in different tissues. It is known that tRNA halves are formed through the cleavage of mature tRNAs by angiogenin in the anticodon loop (Fig. 1D), a small nuclease (14 kDa) present in the nucleus and cytoplasm [8]. Additionally, rat angiogenin mRNA has been detected predominantly in the adult liver, but it was also detectable at low levels in other tissues [20]. Another possibility is the different distribution of selected tRNAs in tissues. Dittmar *el al.* analyzed the abundance of human tRNA molecules in different tissues and observed that human tRNA expression varies up to ten fold among human tissues, with the highest concentration in the liver and the lowest in the testis [21].

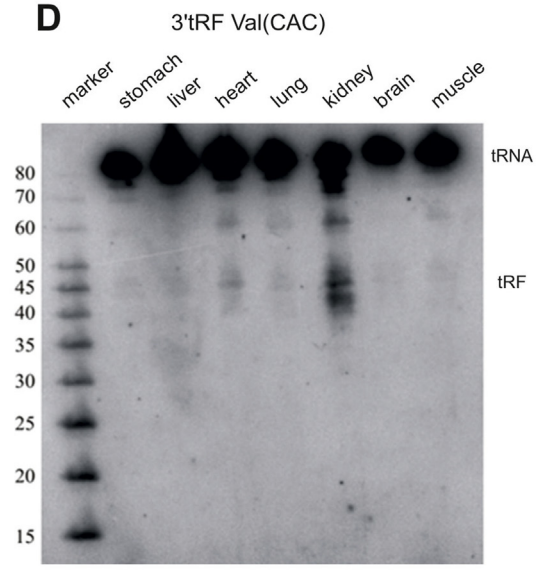
3.3. Interaction of sncRNAs with the PIWIL4 protein

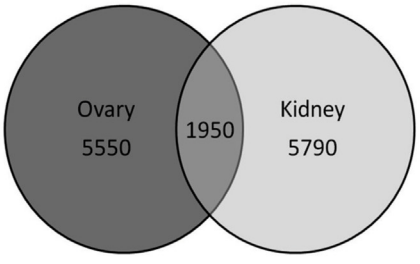
It has previously been found that some tRFs from human breast cancer MDAMB231 cells and from mouse testis tissue can immunoprecipitate together with the PIWIL4 protein [18]. Do *S. scrofa* kidney- and ovary-specific 5'tRFVal(CAC) also interact with PIWI proteins similarly to human tissue? To answer this question, we tested the overexpression of PIWI proteins in *E. coli* cells. For this experiment the pET151/D-TOPO vector containing the ORF sequence of *Piwil1*, *Piwil2* or *Piwil4* genes was constructed. However, the overexpressed PIWI proteins formed inclusion bodies and were insoluble even in 7 M of urea or guanidine. To improve solubility, the *S. scrofa* PIWIL4 protein, containing 842 amino acids, was shortened to the PAZ domain. The cDNA sequence was cloned in the pET151/D-TOPO vector, expressed in *E. coli* and purified. The 140 amino acid PAZ domain of the PIWIL4 protein is an independent domain that is responsible for piRNA binding. Fig. 2A shows the resolved crystal structure of the silkworm SIWI protein [22]. The PAZ domain of the SIWI protein is circled, clearly showing that











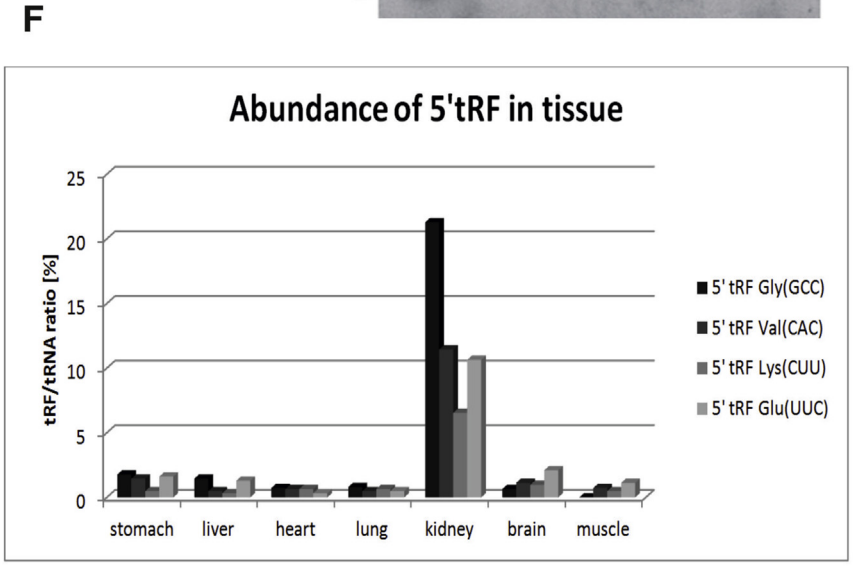


Fig. 1. Characteristics of the tRFs present in *S. scrofa* ovary and kidney. **A** – Confirmation of the abundance of selected tRFs in the ovary using Northern blotting, **B** - Detection of 5'tRFVal(CAC) and **C** - 3'tRFVal(CAC) in different tissues using Northern blotting. Total isolated RNAs were separated on a 12% polyacrylamide gel, transferred to a membrane, and detected by a ^{32}P-labeled DNA probe that hybridized to the 5′ or 3′ part of the tRNAs (Fig. S1 and Table S1). M indicates the RNA marker 10–100 nt. **D** – On the tertiary structure of tRNA (yeast tRNA^{Phe}) cleavage sites in the anticodon loop generating 5′ and 3'tRFs are marked by arrows. **E** – Venn diagram showing the tRFs present in both the ovary and the kidney. **F** – Occurrence of selected tRFs in the tissues detected by Northern blotting.

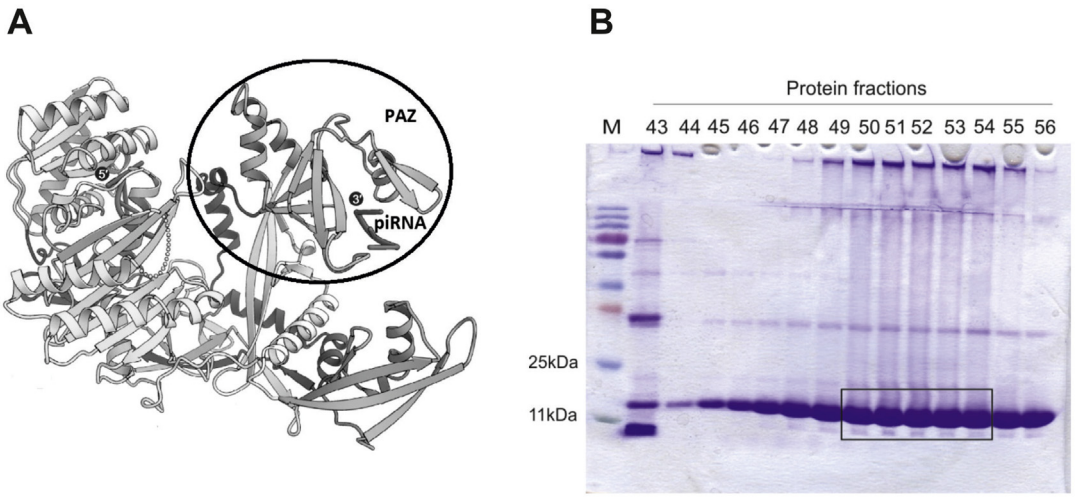


Fig. 2. Characteristics of the sncRNA-binding protein. **A** – The tertiary structure of the silkworm SIWI protein was adopted from a previously resolved crystal structure [22]. The PAZ domain is circled. **B** – Purification of the PAZ domain PIWIL4 protein via a Superdex column. Fractions collected for EMSA analysis are boxed.

it is an independent domain that weakly interacts with the other domains of this protein.

The EMSA method was used to test the binding of the sncRNA to the purified PAZ domain of the PIWIL4 protein (Fig. 2B). The sncRNAs present in the *S. scrofa* ovary (i.e., piRNA 6/92, tRFVal(CAC) and miR 10b) [7] were investigated to see whether they formed complexes with the PAZ domain of the PIWIL4 protein (Fig. 3). It was assumed that only piRNA would interact with the PIWI proteins [3]. However, we found that one of the most abundant sncRNAs in *S. scrofa* gonad and kidney tissues, 5'tRFVal(CAC), also interacts with the PAZ domain. In spite of that, the mir10 oligomer did not bind to the tested protein. The determined K_d values were as follows: 1.70 ± 0.46 μM for binding the piRNA oligomer and 6.69 ± 1.36 μM M for the tRFVal(CAC) oligomer; these values were similar to the values reported in other studies. For example, the K_d for the mouse PIWI-PAZ domain interacting with piRNA was calculated to be 2.0 μM [23].

In germ cells, PIWI proteins bind differently sized piRNAs (25–31 nt). The mechanism of action of piRNA-PIWI complexes in germ cells involves piRNA recognizing the retrotransposon sequences and silencing them through degradation, using the nuclease activity of some PIWI proteins [3]. Less is known about the role of RNA-PIWI complexes in somatic cells. As mentioned above, immunoprecipitation of PIWIL4/RNA complexes isolated from human cancer cells revealed the presence of 5′tRF molecules, of which tRFVal(CAC) and tRFGly(GCC) were the most prevalent. In this short communication we showed that the tRFVal(CAC) found in the gonads and kidney can bind to the PIWIL4 protein with a K_d value comparable to that of piRNA. Recently, PIWIL4 protein expression was detected in mouse tissues and was found to be the highest in the gonads and kidney [24]. This high expression is unlike that of the PIWIL1 protein, which has the lowest expression levels in the kidney. Thus, in *S. scrofa*, the PIWL4 protein may also be present in the kidneys and interact with tRFs. In addition, our current investigations indicate that the expression of PIWI proteins depends on the stage of renal cell development, from stem cells to adult cells. However, the actual biological importance of these observations requires further research.

3.4. Possible tRFs functions in model mammalian organisms

The role of the 31, 32 and 41, 42 nt RNA fragments derived from mature tRNAGly(GCC) and tRNAVal(CAC) and restricted to the ovaries and kidneys has not yet been clarified. However, during embryonic development both secretory and sexual organs evolve from the intermediate mesoderm [25]. The undifferentiated gonads and kidneys are located in the mammalian dorsal–ventral axis, from which these organs differentiate. Furthermore, in the embryonic stage, these organs are located next to each other and use the same leading cord. Thus, a similar *S. scrofa* tRF profile in the

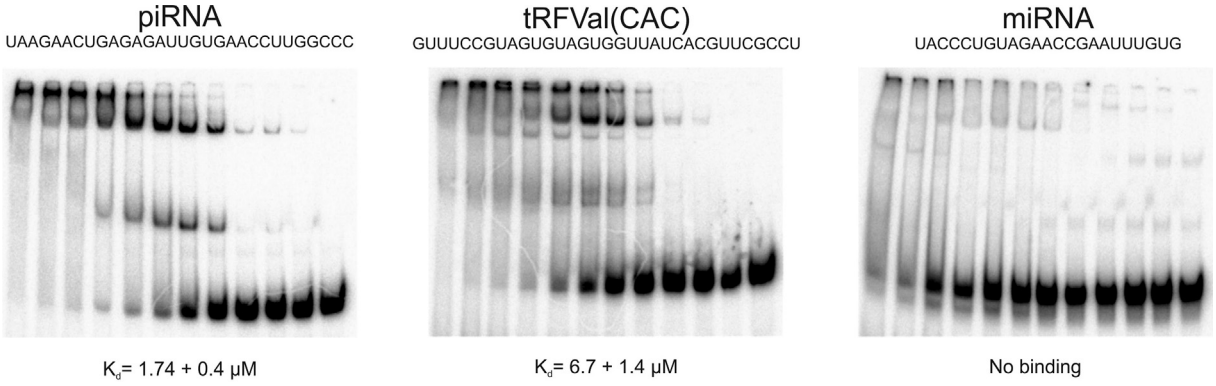


Fig. 3. EMSA analysis interaction of the PAZ domain of the PIWIL4 protein with sncRNA oligomers. The PAZ domain of PIWIl44 protein was overexpressed in *E. coli*. piRNA 6/92, 5'tRFGly(GCC) and miR 10b were obtained by *in vitro* transcription (Table S2). The assay was carried out on 10% polyacrylamide gels with ^{32}P-labeled sncRNA and the PAZ protein.

gonads and kidneys may be a remnant of a common developmental path.

However, it is postulated that the 5'tRFGly(GCC) and tRFVal(-CAC) PIWIL4 complexes present in MDAMB231 human cancer cell line are involved in translation regulation because they are associated with ribosomes [18]. The mechanism suggested for human cells involves the displacement of the translation initiation factors by tRFs and the inhibition of translation [26]. Recent studies confirmed the folding of the tRFs structure into a G-quarduplex motif, which is necessary to decrease translation efficiency [27].

In mouse circulating blood and serum 5'tRFGly(GCC) and 5'tRFVal(CAC) are the most abundant, with the amounts varying with age and diet [19]. However in contrast to porcine tRFs, the 3'tRFs in mouse serum are at a very low level. These observations suggest that these 5'tRFs can be treated as organismal signaling molecules. Sharma et al. presented another mechanism in which the diet dependent 5'tRFGly(GCC) represses mouse genes that are associated with the endogenous MERVL retroelements, both in embryonic stem cells and embryos [28].

Acknowledgements

This paper was supported by the National Science Center grant no. UMO-2016/21/B/NZ2/03468.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrc.2019.10.062.

References

[1] T.R. Cech, J.A. Steitz, The Noncoding RNA Revolution-Trashing Old Rules to Forge New Ones Cell, vol. 157, 2014, pp. 77–94.
[2] J. Król, I. Loedige, W. Filipowicz, The widespread regulation of microRNA biogenesis function and decay, Nat. Rev. Genet. 11 (2010) 597–610.
[3] Y.W. Iwasaki, M.C. Siomi, H. Siomi, Piwi-interacting RNA: its biogenesis and function, Annu. Rev. Biochem. 84 (2015) 405–433.
[4] J. Gebetsberger, N. Polacek, Slicing tRNAs to boost functional ncRNA diversity, RNA Biol. 10 (2013) 1798–1806.
[5] Y.S. Lee, Y. Shibata, A. Malhotra, A. Dutta, A novel class of small RNAs: tRNA-derived RNA fragments (tRFs), Genes Dev. 23 (2009) 2639–2649.
[6] V. Oberbauer, R.S. Schaefer, tRNA-derived small RNAs: biogenesis, modification, function and potential impact on human disease development Genes 9 (2018) 607.
[7] D. Kowalczykiewicz, A. Świercz, L. Handschuh, K. Leśniak, M. Figlerowicz, J. Wrzesinski, Characterization of *Sus scrofa* small non-coding RNAs present in both female and male gonads, PLoS One 9 (11) (2014) e113249.
[8] L. Zhu, J. Ge, T. Li, Y. Shen, J. Guo, tRNA-derived fragments and tRNA halves: the new players in cancers, Cancer Lett. 452 (2019) 31–37.
[9] C. Cole, A. Sobala, C. Lu, S.R. Thatcher, et al., Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs, RNA 25 (2019) 2147–2160.
[10] C. Kuscu, P. Kumar, M. Kiran, Z. Su, et al., RNA fragments (tRFs) guide Ago to regulate gene expression post-transcriptionally in a Dicer-independent manner, RNA 24 (2018) 1093–1105.
[11] H. Goodarzi, X. Liu, H.C.B. Nguyen, S. Zhang, et al., Endogenous tRNA-Derived fragments suppress breast cancer progression via YBX1 displacement, Cell 161 (2015) 790–802.
[12] D. Kowalczykiewicz, P. Pawlak, D. Lechniak, J. Wrzesinski, Altered expression of porcine *Piwi* genes and piRNA during development, PLoS One 7 (8) (2012) e43816.
[13] J. Wrzesinski, S.K. Jóźwiakowski, Structural basis for recognition of Co^{2+} by RNA aptamers, FEBS J. 275 (2008) 1651–1662.
[14] X. Liu, S. Liao, Z. Xu, L. Zhu, et al., Identification and analysis of the porcine microRNA in porcine cytomegalovirus-infected macrophages using deep sequencing, PLoS One 11 (3) (2016) e0150971.
[15] R.S. Prather, Pig genomics medicine Nature Biotech 31 (2013) 122–124.
[16] H. Marx, H. Hahne, S.E. Ulbrich, A. Schnieke, et al., Annotation of the Domestic Pig Genome by Quantitative Proteogenomics, vol. 16, 2017, pp. 2887–2898.
[17] Y.-M. Hou, CCA Addition to tRNA: implications for tRNA quality control, IUMB Life 62 (2010) 251–260.
[18] S.P. Keam, P.E. Young, A.L. McCorkindale, T.H. Dang, et al., The human Piwi protein Hiwi2 associates with tRNA-derived piRNAs in somatic cells, Nucleic Acids Res. 42 (2014) 8984–8995.
[19] J.M. Dhahbi, S.R. Spindler, H. Atamna, A. Yamakawa, et al., 5′ tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and calorie restriction, BMC Genomics 14 (2013) 198.
[20] H.L. Weiner, L.H. Weiner, J.L. Swain, Tissue Distribution and Developmental Expression of the Messenger RNA Encoding Angiogenin Science, vol. 237, 1987, pp. 280–282.
[21] K.A. Dittmar, J.M. Goodenbour, T. Pan, Tissue-specific differences in human transfer RNA expression, PLoS Genet. 2 (12) (2006) e221.
[22] N. Matsumoto, H. Nishimasu, K. Sakakibara, K.M. Nishida, et al., Crystal Structure of Silkworm PIWI-Clade Argonaute Siwi Bound to piRNA Cell, vol. 167, 2016, pp. 484–497.
[23] B. Simon, J.P. Kirkpatrick, S. Eckhardt, M. Reuter, et al., Recognition of 2'-O-Methylated 3'-end of piRNA by the PAZ Domain of a Piwi Protein Structure, vol. 19, 2011, pp. 172–180.
[24] B.P.U. Perera, Z.T. Tsai, M.L. Colwell, T.R. Jones, et al., Somatic expression of piRNA and associated machinery in the mouse identifies short, tissue-specific piRNA, Epigenetics 14 (2019) 504–521.
[25] S.Y. Park, J.L. Jameson, Minireview: Transcriptional Regulation of Gonadal Development and Differentiation Endocrinology, vol. 146, 2005, pp. 1035–1042.
[26] P. Ivanov, M.M. Emara, J. Villen, S.P. Gygi, P. Anderson, Angiogenin-induced tRNA fragments inhibit translation initiation Mol, Cell 43 (2011) 613–623.
[27] P. Jackowiak, A. Hojka-Osinska, K. Gasiorek, M. Stelmaszczuk, et al., Effects of G-quadruplex topology on translational inhibition by tRNA fragments in mammalian and plant systems in vitro, Int. J. Biochem. Cell Biol. 92 (2017) 148–154.
[28] U. Sharma, C.C. Conine, C.C. Shea, A. Boskovic, et al., Biogenesis and function of tRNA fragments during sperm maturation Science 351 (2016) 391–396.

International Journal of
Molecular Sciences

MDPI

Article

Characteristics of Transfer RNA-Derived Fragments Expressed during Human Renal Cell Development: The Role of Dicer in tRF Biogenesis

Marek Kazimierczyk, Marta Wojnicka, Ewa Biała, Paulina Żydowicz-Machtel, Barbara Imiołczyk, Tomasz Ostrowski, Anna Kurzyńska-Kokorniak and Jan Wrzesinski *

Institute of Bioorganic Chemistry, Polish Academy of Sciences, 61-704 Poznań, Poland; mkazimierczyk@ibch.poznan.pl (M.K.); mwojnicka@ibch.poznan.pl (M.W.); ewabiaa@gmail.com (E.B.); pzydowicz@ibch.poznan.pl (P.Ż.-M.); barim@ibch.poznan.pl (B.I.); tostr@ibch.poznan.pl (T.O.); akurzyns@ibch.poznan.pl (A.K.-K.)
* Correspondence: wrzesinj@ibch.poznan.pl

Abstract: tRNA-derived fragments participate in the regulation of many processes, such as gene silencing, splicing and translation in many organisms, ranging from bacteria to humans. We were interested to know how tRF abundance changes during the different stages of renal cell development. The research model used here consisted of the following human renal cells: hESCs, HEK-293T, HK-2 and A-489 kidney tumor cells, which, together, mimic the different stages of kidney development. The characteristics of the most abundant tRFs, tRFGly(CCC), tRFVal(AAC) and tRFArg(CCU), were presented. It was found that these parental tRNAs present in cells are the source of many tRFs, thus increasing the pool of potential regulatory RNAs. Indeed, a bioinformatic analysis showed the possibility that tRFGly(CCC) and tRRFVal(AAC) could regulate the activity of a range of kidney proteins. Moreover, the distribution of tRFs and the efficiency of their expression is similar in adult and embryonic stem cells. During the formation of tRFs, HK-2 cells resemble A-498 cancer cells more than other cells. Additionally, we postulate the involvement of Dicer nuclease in the formation of tRF-5b in all the analyzed tRNAs. To confirm this, 293T NoDice cells, which in the absence of Dicer activity do not generate tRF-5b, were used.

Keywords: renal cells; tRNA-derived fragment; tRF; Dicer; Northern blot

Citation: Kazimierczyk, M.; Wojnicka, M.; Biała, E.; Żydowicz-Machtel, P.; Imiołczyk, B.; Ostrowski, T.; Kurzyńska-Kokorniak, A.; Wrzesinski, J. Characteristics of Transfer RNA-Derived Fragments Expressed during Human Renal Cell Development: The Role of Dicer in tRF Biogenesis. *Int. J. Mol. Sci.* **2022**, *23*, 3644. https://doi.org/10.3390/ijms23073644

Academic Editors: Anna Cieślińska and Alfred King-Yin Lam

Received: 9 February 2022
Accepted: 24 March 2022
Published: 26 March 2022

1. Introduction

According to current analyses, nearly 80% of the human genome has been transcribed [1]. However, only a negligible part, comprising some 20,000 genes, of which about 2% are in the form of mRNA, contain protein-coding sequences [2]. The remaining non-mRNA RNAs belong to a huge group of non-coding RNAs (ncRNAs). Due to their function, ncRNAs can be classified as housekeeping RNAs or regulatory RNAs [3,4]. Housekeeping ncRNAs, which include transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and ribosomal RNAs (rRNAs), are commonly expressed constitutively [5]. However, regulatory RNAs are a type of ncRNA with a strong regulatory impact on the expression of protein-coding genes. Based on their size, regulatory RNAs can be divided into two groups: small noncoding RNAs (sncRNAs), which consist of miRNAs as well as piRNAs, and long noncoding RNAs (lncRNAs) [6,7].

The main role played by tRNAs is their participation in the protein biosynthesis process. tRNAs transfer amino acids attached to their 3′ ends by a specific enzyme, aminoacyl-tRNA synthetase, to the ribosome. They then interact with the appropriate codon, embedded in mRNA, complementing it with a tRNA anticodon sequence. This guarantees that the protein has the appropriate structure and function [8]. Moreover, the translation process is dependent on regulatory RNAs, miRNAs. These are 22–23 nucleotide small

RNAs obtained by cleaving longer pri-miRNAs. miRNAs bind to mRNA and influence translation mechanisms in the cell [9,10]. The role of miRNAs has been well recognized, and Fire and Mello were awarded the Nobel Prize in 2006 for their research concerning small interfering RNAs [11]. Apart from the above-mentioned roles tRNAs play in protein translation, they also have other noncanonical functions. tRNAs participate in cellular processes, such as transcription, splicing, immune responses and apoptosis, as well as acting as primers for reverse transcriptase in the retroviral genome [12,13].

In 2009, it was discovered that a second level of gene regulation involving tRNA-derived fragments (tRFs) exists [14]. The solved tRNA crystal structure showed two single-stranded regions exposed to the solution—the first, an anticodon loop, and the second, a tRNA elbow, i.e., a junction between the D and TΨC loops [15,16]. Both sites are preferred in the biogenesis of tRNA-derived fragments by specific nucleases, which generate the appropriate sets of fragments. It was suggested [17,18] that the tRFs be divided into several groups, one of which was tRF-5, which starts at the 5′ end of the mature tRNA and extends into the D loop and, at a maximum, into the anticodon stem. This variation in size allows for the further classification of tRF-5s: (i) tRF-5a extends into the D loop; (ii) tRF-5b consists of a 3′ end part of tRNA, which reaches beyond the D loop and into the D arm; (iii) tRF-5c represents the longest tRF-5 class, extending into the anticodon stem and ranging from 27 to 30 nucleotides; (iv) 5′ tRNA half; (v) 3′ tRNA half; and (vi) internal tRF. Additionally, tRFs derived from the 3′ CCA end of a mature tRNA are known as tRF-3s. These tRFs have two major size classifications: tRF-3a, which is about 18 nucleotides long, and tRF-3b, which is about 22 nucleotides long.

There are several groups of enzymes involved in tRF biogenesis. The 5′ leader sequence is a substrate for RNase P, producing the mature 5′ end of tRNA, while nuclease Z cleaves the 3′ trailer [19,20]. Moreover, it was suggested that the Dicer nuclease produces tRF-5 and tRF-3 fragments of a size similar to that of miRNA, i.e., 20–25-nucleotide long oligomers, which are cut by Dicer at the tRNA elbow end [18]. Furthermore, nuclease T2 is involved in tRF formation by cutting the D and TΨC loops in various tRNAs. [21]. Members of the RNase T2 family are found in all groups of living organisms, namely, viruses, bacteria, fungi, plants, animals, as well as humans. However, it is the tRNA anticodon loop that is the main substrate for known RNases, such as angiogenin, which is responsible for the formation of tRNA halves in mammalian cells [22].

It is interesting to note that the composition of tRFs depends on the type of organism, or even the cells present in different tissues. Moreover, the abundance of tRFs is affected by environmental factors, diet, diseases or external stimulation [23]. Therefore, a network of small non-coding RNAs—tRFs—in a cell can be an excellent regulator of cellular processes. Indeed, tRFs are known to play a significant role in various physiological and pathological processes [18,23]. Goodarzi et al. reported that in breast cancer cells, under hypoxic conditions, several tRFs were upregulated, which suppressed the stability of the oncogenic transcript by displacing its 3′ UTR from the YBX1 protein [24]. These hypoxic stress-induced tRFs can competitively bind to the YBX1 protein and block its interaction with oncogenic mRNAs. Another 3′ half tRF derived from mature tRNAGlu(UUC) has been found to be able to bind and displace the RBP nucleolin in breast cancer [25].

tRFs are involved in regulating translation during stress by competing with mRNA for ribosome binding [26]. Another way tRFs function in the cell is through their interaction with Argonaute (AGO) proteins, by forming complexes and inhibiting gene expression in the same manner as miRNAs [27]. Moreover, tRFs interact with PIWI proteins, forming stable complexes [28,29]. The immunoprecipitation of PIWIl4 (Hiwi2) from MDAMB231 cancer cells revealed the presence of the following tRFs: 5′ tRFGlu(CUC), tRFLys(TTT) and tRFVal(ACC), which are predominantly derived from processed tRNAs [28]. Similarly, the *S. scrofa* PAZ domain of the Piwil4 protein bound to the 5′ tRFVal(CAC) half [29]. Previous studies on the role of tRFs have emphasized their importance in a variety of cellular processes, including mRNA stabilization, miRNA-mediated gene silencing, as well as the regulation of cap-dependent and cap-independent translation [30]. In addition,

these molecules are usually dysregulated in cancers, which is why a significant amount of research is currently focused on the role of tRFs in development and cancer [31–33].

2. Results

2.1. Research Model

Figure 1A shows the research model used to study kidney cell development. It consists of a human renal cell culture which mimics the different stages of kidney development. The first stage reflects human embryonic stem cells (hESCs), which are derived from the inner cell mass of blastocyst-stage embryos and have the ability to differentiate into various cell types. The next stage of kidney development is represented by embryonic HEK-293T cells. The original 293 cells were derived from the human embryo kidney and transformed through being exposed to fragments of sheared adenovirus 5 DNA. The HK-2 are adult cells, from an immortalized human kidney cell line. In contrast, A-498 cells, derived from a kidney carcinoma whose development was disturbed, were chosen as a model for renal cancer cells. The cells (ATCG®) were cultured in standard conditions, according to the manufacturer's recommendations listed in the Material and Methods section. Figure 1, panel A presents cell culture physiology. A comparison with the manufacturer's data shows that the cell cultures grew correctly. RNA isolated using the Trizol procedure was purified, deprived of ribosomal RNAs and fragmented. A small fraction of the obtained RNA was used for library construction. Illumina sequencing was carried out by the Macrogene company (Seoul South Korea). Bioinformatic analysis of the Illumina sequencing results revealed 30–40 million reads for each library.

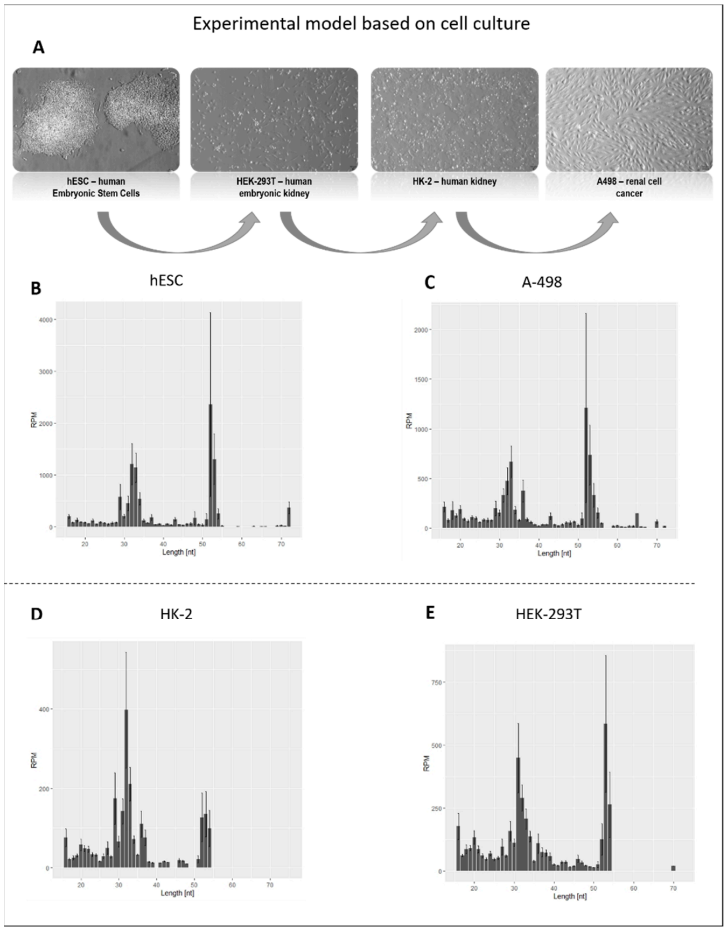


Figure 1. The size distribution (in nucleotide length) of tRNA-derived fragments isolated from renal cells. (**A**) The physiology of model cell cultures. (**B**–**E**) The names of the analyzed cells are shown at the top of each panel. RPM—reads per million.

2.2. Bioinformatic Analysis of tRNA-Derived Fragments in the Cells

Figure 1B–E display the expression of tRFs in the analyzed cells, depending on their length. It should be emphasized that there was a wide variety of tRNA fragments, with different levels of expression and length, in the model cells. tRNA oligomers, ranging from 31–35 nucleotides to 52–54 nucleotides were the most abundant. This size corresponds to cleavages in the anticodon and the TΨC loops. Moreover, there is a lower abundance of tRFs, due to the cleavage of the D loop.

It was found that the abundance of tRNA-derived fragments is dependent on the cell type. The greatest number of tRNA fragments was found in regular HK2 cells and the smallest number in hESCs (Figure 1B–E). Since multiple codons are present for many amino acids, multiple tRNAs with distinct anticodons (i.e., isoacceptors) are needed to read these codons. In humans, tRNA isoacceptors are present for 12 amino acids, ranging from two for Glu, Lys and Gln, three for Ile, Val, Thr, Ala, Gly and Pro, four for Ser and five for Leu and Arg.

The analysis of the expression of tRFs in a selection of renal cells (stem cells, embryonic cells, adult cells and cancer cells) revealed the highest abundance of the following tRNA fragments: tRFGly, tRFVal and tRFArg (Figure 2A).

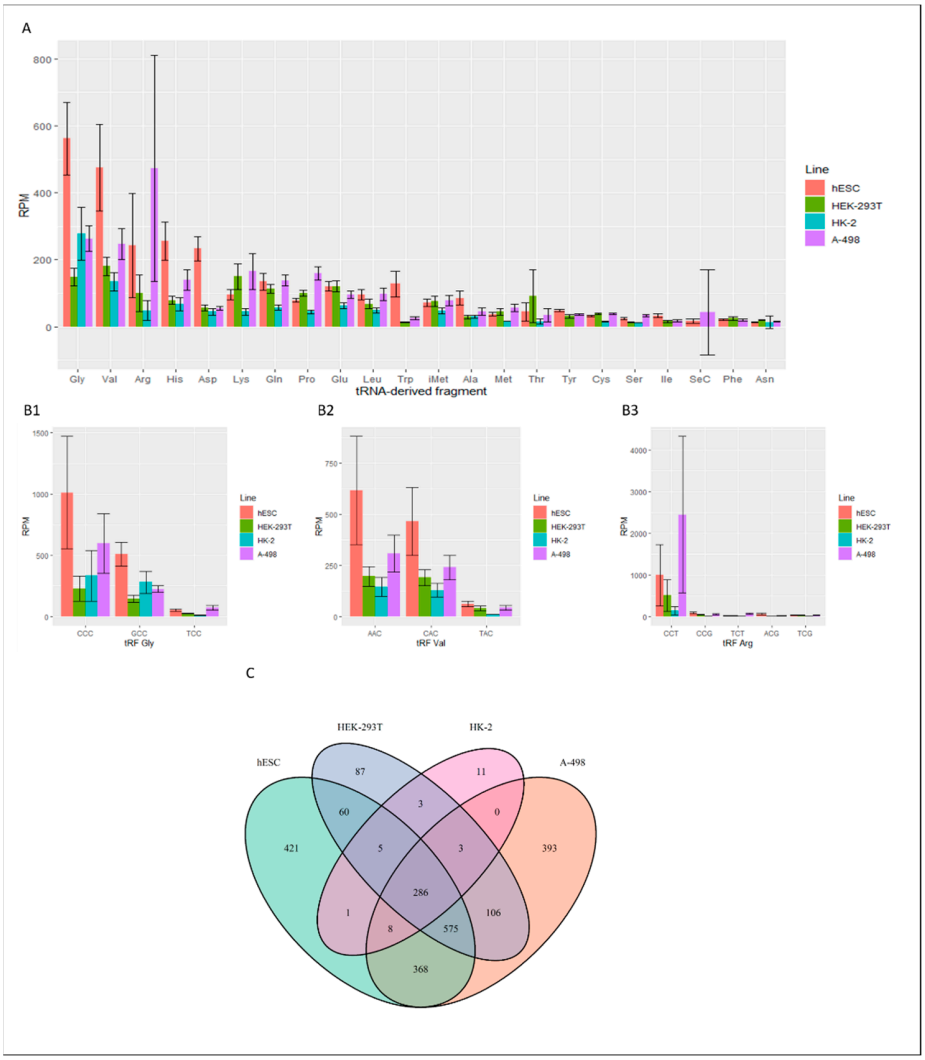


Figure 2. Bioinformatic characterization of tRNA-derived fragments in renal cells. (**A**) tRF occurrence in different renal cells. (**B1**–**B3**) Analysis of the most abundant tRFGly, tRFVal and tRFArg isoacceptors. (**C**) Venn diagram formed using the number of specifically expressed tRFs and their presence in different cells.

In the case of tRFHis, tRFAsp and tRFLys, their expression levels were much lower. However, in tRNAPhe and tRNAAsn-derived fragments, they were hardly expressed. Two fragments, tRFGly(CCC) and tRFGly(GCC), were the most abundant in human renal cells (Figure 2(B1)). In the bioinformatic analysis results, the highest expression of tRFGly(CCC) and tRFGly(GCC) occurred in hESC and cancer cells. However, much lower levels of expression of these tRNA fragments were found in HEK-293T and HK-2 cells during the actual experiment. The highest expression of tRFVal(AAC) and tRFVal(CAC) was noted in hESCs; by contrast, it was two times lower in cancer cells (Figure 2(B2)). The expression of these tRFs was lowest in HEK-293T and HK-2. Unlike the results obtained during the analysis, tRFArg(CCT) was the most abundant in A-498 cancer cells, while its levels were three times lower in hESCs (Figure 2(B3)). The expression of this tRF in HEK-293T and HK-2 cells was almost negligible. In Figure 2C, the Venn diagram based on the Illumina sequencing data shows the division of the identified tRF fragments into those that are specific to one type of model kidney cell and those that are found in more than one type of model kidney cell. The largest number of the identified tRNA fragments was found in hESC and A-498 cancer cells, 421 and 393, respectively. In HEK-293T and HK-2 cells, the number of tRFs present is several times lower. Interestingly, hESC, HEK-293T and cancer cells have a large number of identical tRFs—575.

2.3. Mapping tRFs in Renal Cells

The results of the bioinformatic analysis were experimentally verified, using the Northern blot technique (Figure 3). Despite the fact that tRNA molecules constitute 15% of the total RNA in the cell, less than 1% of tRNA is involved in the biogenesis of tRNA fragments [34]. To detect tRFs, two 20-nucleotide long ^{32}P-labeled probes were hybridized to the parental tRNA from the 5′ and 3′ ends (Table S1). For tRNAGly(CCC), three dominant cleavages in the anticodon loop, at U33, C35, generating 5′ tRNA half, and C54, generating tRF3b, were detected (Figure 3A,C). Analysis of 3′ tRF fragments detected with the use of probe 2 showed three main cleavage sites, namely, C33, C38 (3′ tRNA half) and C54 (tRF3b), with an additional one in the D loop (Figure 3B,C). Such cleavage sites indicate that all of the single-stranded regions, the D, anticodon and TΨC loops, are nuclease substrates (Figure 3A–C). When comparing the tRFs detected by both probes, both 5′ and 3′ fragments were present. There were also many fragments that were exclusively detected only by a 5′ or a 3′ probe. The first group includes tRFs, whose biogenesis takes place through the cleavage of the phosphodiester bond in the C35–C36 and T52–Ψ53 nucleotides. The second group includes 5′ tRNA halves and 3′ tRNA halves, obtained by cleavage at C31 and U34, respectively. Other tRF-5s were generated by cleavage at G10, D16 and G42. What, then, is the reason for the observed effects of the presence of one 5′ tRNA half and the absence of the other 3′ tRNA half? RNA surveillance in the cell may explain this observation, as redundant RNA is often degraded.

In the case of tRNAVal(AAC), the following cleavages (Figure 4A–C) were detected during tRF biogenesis. The 5′ probe 3 revealed cuts at U20 and U22 in the D loop (tRF-5b). Similar cleavages in the D loop were identified using the 3′ probe. In the anticodon loop, the most prominent cleavage occurred between A35 and A36 (5′ and 3′ tRNA halves). Additionally, cleavages at A44 and U47 (tRF-3b) were found in the variable loop (not marked), mainly in hESCs. Several cleavages were detected in the TΨC loop (Figure 4A–C). Cleavages at A58 and A60 (tRF-3b) were detected using both probes. Moreover, cleavages at T54 were most visible in HEK-293T cells.

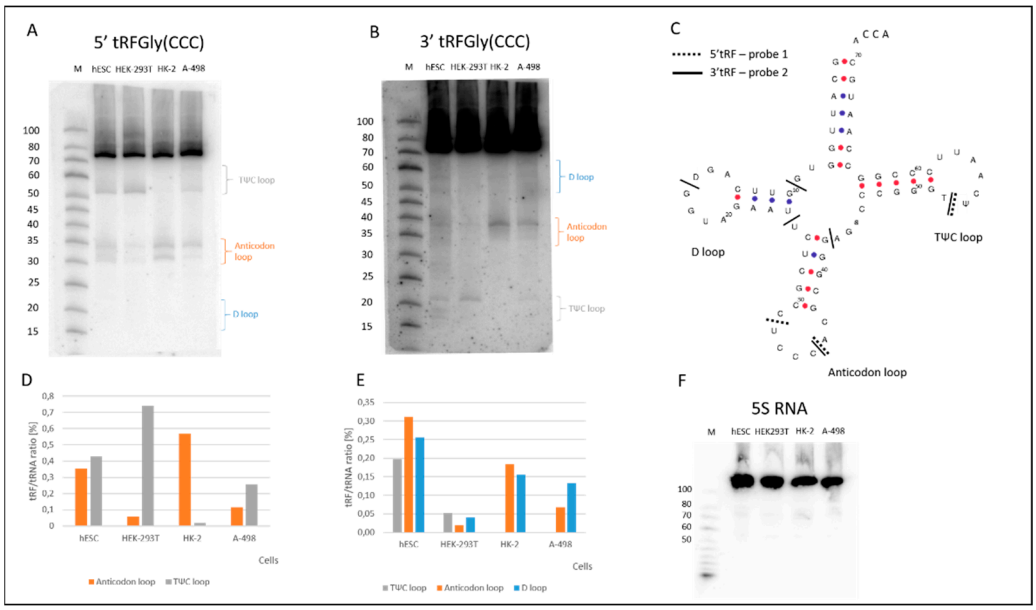


Figure 3. Characterization of the tRNAGly(CCC)-derived fragments. Northern blot (**A**) 5′ and (**B**) 3′ tRF detection with the use of two ^{32}P labelled probes 1 and 2 respectively (Table S1). (**C**) Secondary structure of human tRNAGly(CCC), with marked 5′ (—) and 3′ (—) cleavage sites. The locations of some of the modified D (dihydrouridine), T (ribothymidine) and Ψ (pseudouridine) nucleotides are marked in the tRNA structure. The loops in the tRNA structure have been named after these modifications. (**D**,**E**) The efficiency of the tRNAGly(CCC) cleavage in loops of the analyzed renal cells. (**F**) Loading control with a hybridization probe, which matches 5S rRNA.

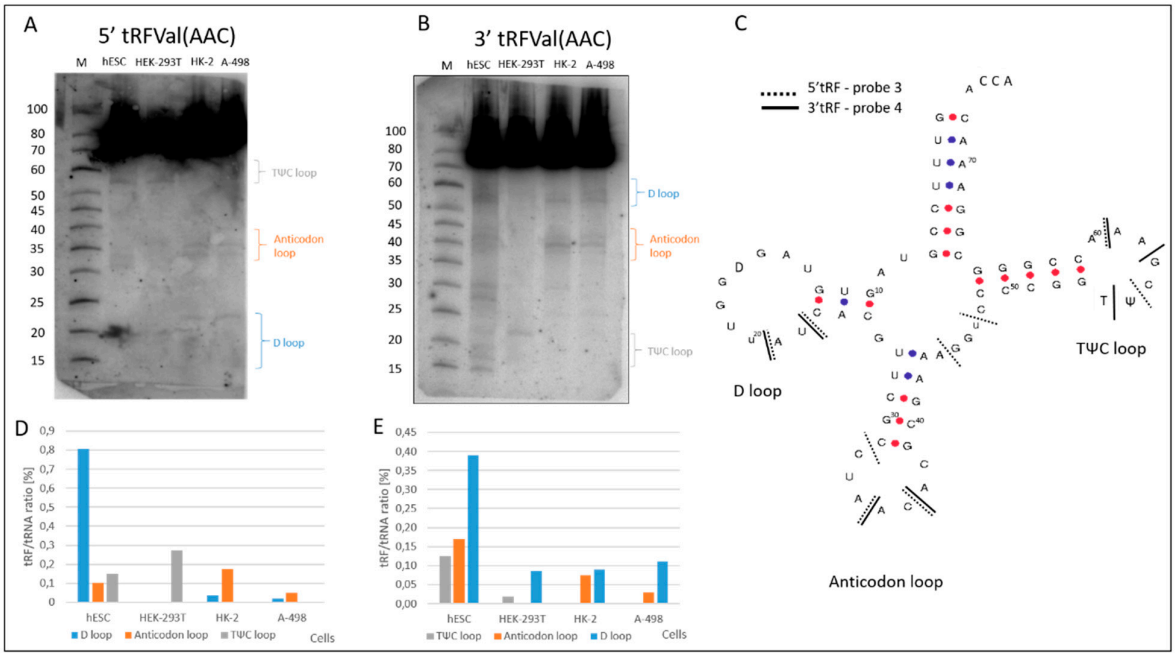


Figure 4. Characterization of the tRNAVal(AAC)-derived fragments. Northern blot (**A**) 5′ and (**B**) 3′ tRF detection with the application of the ^{32}P labelled probes 3 and 4 respectively (Table S1). (**C**) Secondary structure of human tRNAVal(AAC) with marked 5′ (—) and 3′ (—) cleavage sites. The locations of some of the modified D (dihydrouridine), T (ribothymidine) and Ψ (pseudouridine) nucleotides are marked in the tRNA structure. The loops in the tRNA structure have been named after these modifications. (**D**,**E**) tRF occurrence and efficiency of tRNAVal(AAC) cleavage in the analyzed renal cells detected by 3 and 4 probes.

tRNAArg(CCU) is the source of a limited number of fragments (Figure 4). Three different cut sites were found using probe number 5 and four sites using probe number 6 (Figure 5A–E, Table S1). It is interesting to note that there are sites that are specific to a certain cell type, such as U20 cleavage in the D loop (tRF-3b), which is specific to hESCs, just like C35 cleavage in the anticodon loop or the G51 cleavage site in the TΨC loop, which occurs only in HEK-293T cells. There are also certain kinds of cleavages which are more widespread, such as C34 cleavage (5′ tRF half), which is present in all the cells with a preference for hESCs, or G17 cleavage, which occurs in all of the model cells.

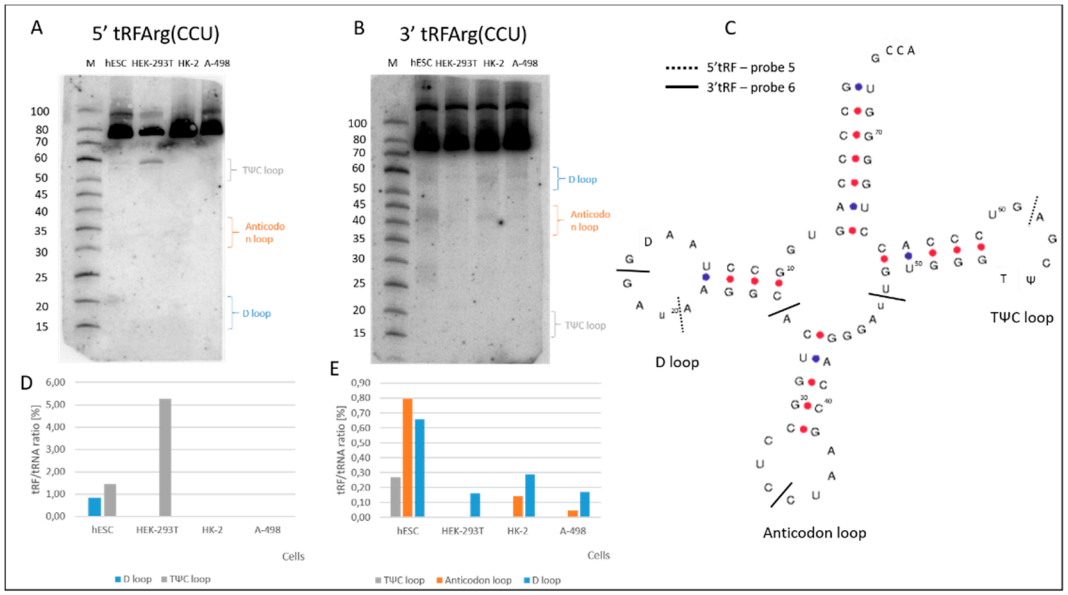


Figure 5. Characterization of the tRNAArg(CCU)-derived fragments. Northern blot (**A**) 5′ and (**B**) 3′ tRF detection with the application of ^{32}P labelled probes 5 and 6, respectively (Table S1). (**C**) Secondary structure of human tRNAArg(CCU) with marked 5′ (—) and 3′ (—) cleavage sites. The locations of some of the modified D (dihydrouridine), T (ribothymidine) and Ψ (pseudouridine) nucleotides are marked in the tRNA structure. The loops in the tRNA structure have been named after these modifications. (**D**,**E**) tRF occurrence and efficiency of tRNAArg(CCU) cleavage in the analyzed renal cells detected by 5 and 6 probes.

Analysis of tRF patterns, detected by 5′ probes and 3′ probes showed that in all of the tRNAs tested here, the D, anticodon or TΨC loops were cut during tRF biogenesis, although with differing efficiency, depending on the type of cell line (Figures 3 and 4). tRF-3b, formed by the cleavage of tRNAGly(CCC) in the TΨC loop at position T54, is predominantly present in hESC and HEK-293T cells (Figure 3).

Moreover, 5′ and 3′ tRNAGly(CCC) halves were almost absent in HEK-293T cells and present in other model cell cultures. In addition, a different expression of the tRF-5b fragment, obtained by cleavage of the D loop in other model cells was observed. In the case of tRNAVal(CAC), this tRF-5b occurred only in hESCs. Generally, the highest cleavage efficiency of three analyzed tRFs was detected for hESCs and the lowest for cancer A-498 cells.

The main reason why cultured A-498 cancer cells were added was to analyze the differences in the occurrence of tRFs compared to normal cells. However, the occurrence of the individual tRFs, as well as the level of their expression, was almost identical. A 50% difference was observed in the tRF/tRNA ratio for tRNAArg(CCU).

2.4. Is the Expression of Kidney Proteins Regulated by tRNA-Derived Fragments?

In order to answer this question, the tRFTars tool was used. [35]. In the first step, the tRF sequences detected using Northern blot analysis were entered into the tRF database

and sequences with similar IDs were found. By applying the tRFTars database, potential kidney protein targets for tRF with the highest probability were found.

The presented analysis showed that the expression of different proteins can be regulated by 5′ tRFGly(CCC), 5004b (tRF-5b) and 5004c (5′ tRF halves) (Table S2A). The expression of MEIS2, FMN1 and CTDSPL2 proteins is associated with impaired renal function and can be regulated by tRF-5b, which can replace miR-458 during tumor suppression. Bioinformatic analysis of 5′ tRFGly(CCC) halves binding to protein mRNA shows their ability to regulate cellular processes, such as glucose hemostasis, the import of membrane proteins or cell energy management, while 3′ tRFGly(CCC) 3027a and 3027b can regulate the activity of a number of proteins specifically found in the kidneys (Table S2B). Seventeen-nucleotide tRF-3b can regulate Ca2+ ion transport by interacting with mRNA ORAI2 or CACNG proteins and modulating the activity of CYP20A1 and GAS7, proteins associated with kidney diseases, such as kidney nephropathy, proteinuria and chronic kidney disease.

The tRFVal(AAC)s which affect the expression of selected proteins in the kidney are limited to tRF 5017b and 3008. In silico analysis showed that tRF-5b can regulate the activity of the NFIC, GNAO1 and HIPK2 proteins, which are associated with renal tumors (Table S3A). In contrast, tRF-3b regulates the activity of the FOXD1, ARHGAP45, TOX 23 and TPCN2 proteins, which perform various functions in the development of the kidneys (Table S3B).

The role of small non-coding RNAs, mainly miRNAs, in regulating gene expression is being intensively studied. In most cases, miRNAs interact with the 3′ untranslated region (3′ UTR) of the target mRNAs to induce mRNA degradation and translational repression [8]. However, miRNAs can interact with other mRNA regions, including 5′ UTR, affecting gene expression. Does a similar mechanism of protein expression regulation exist for tRFs in the kidneys? First, a 5–7 nucleotide tRFGly(CCC) and tRFVal(ACC) seed sequence interacts with the mRNAs, presented in Tables S2 and S3, respectively. The interaction between tRF and mRNA through the seed sequence may be similar to the interaction between mRNA and miRNA. Secondly, tRFs form complexes with Ago proteins, just like miRNAs [26]. The RISC complex, consisting of miRNAs and Ago proteins, is necessary to silence protein expression [36]. However, the presented computer analysis results should be treated as preliminary, requiring experimental verification.

2.5. The Involvement of the Dicer Nuclease in the Generation of tRNA-Derived Fragments

Experiments demonstrating the impact of the transfection of HEK-293T and 293T (NoDice) cells with a plasmid producing full-length wild-type human nuclease Dicer (hDicer) on the efficiency of the cleavage of tRNA were carried out (Figure 6). Unlike natural HEK-293T cells, 293T NoDice cells are entirely devoid of the hDicer nuclease (the Dicer knock-out HEK-293T cells) [37]. To detect hDicer a Western blot with an anti-Dicer antibody was applied. As expected, the hDicer protein was present in the wild-type HEK-293T cells and in the cells transfected with the plasmid producing hDicer. However, the (hDicer) nuclease was not present in 293T NoDice cells. hDicer expression was only detected in these cells after they were transfected with a plasmid producing hDicer. Interestingly, the amounts of the protein produced by wild-type cells before and after they were transfected with the plasmid producing hDicer was the same (Figure 5A). This result suggests a possible mechanism that regulates the level of hDicer nuclease in the cell.

The expression of the most abundant 3′ tRFGly(CCC), tRFVal(AAC) and tRFArg(CCU) in HEK-293T, 293T NoDice and HEK-293T + hDicer, 293T NoDice + hDicer cells was analyzed, using the Northern blot approach. The efficiency of the cleavages in the anticodon loop was very similar for all of the tested tRNAs, regardless of the cell line (Figure 5B–D). However, the cleavages of these tRNAs in the D loop were strongly dependent on the cell line. For wild-type and transfected HEK-293T cells, the cleavages were almost identical, but in 293T NoDice cells, tRF biogenesis did not occur. It only occurred in 293T NoDice cells, which were transfected with a plasmid encoding the wild-type full-length hDicer.

As a control, the biogenesis of miRNA 92 was used. Mature miR-92 was detected in the wild-type and transfected HEK-293Tcells but was absent in 293T NoDice cells.

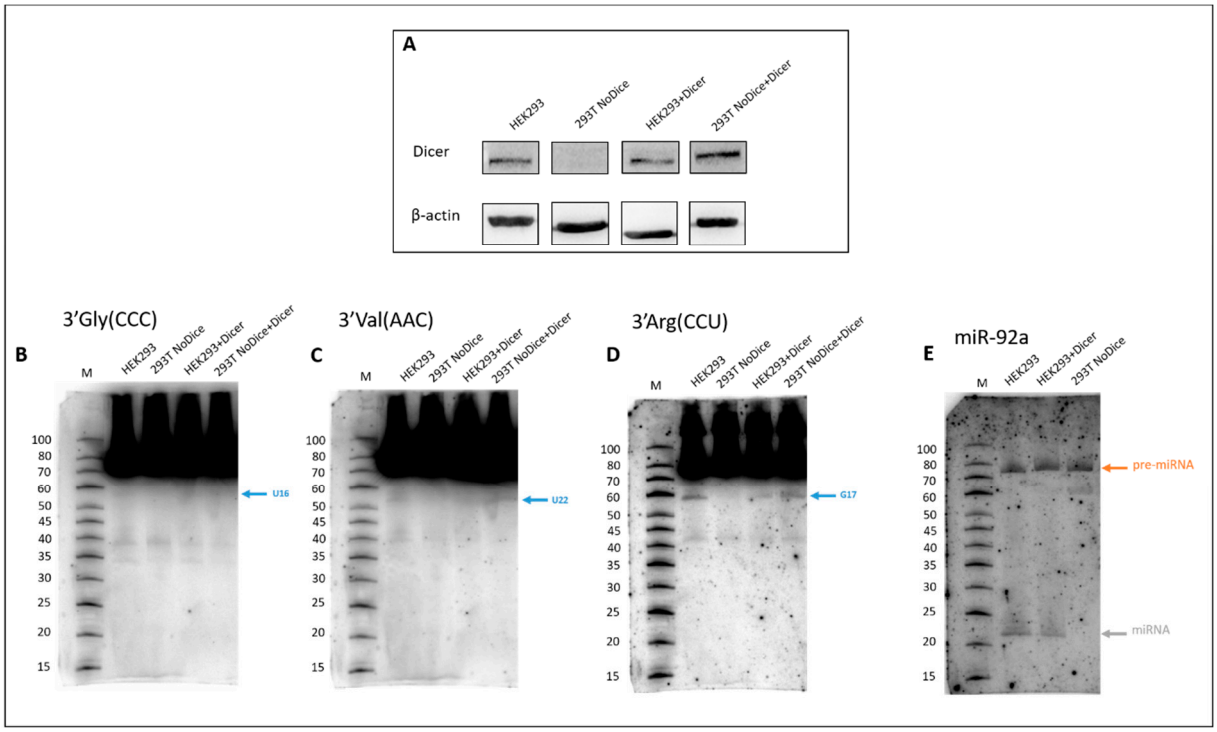


Figure 6. The impact of Dicer overexpression on the biogenesis of tRNA-derived fragments. (**A**) Western blot detection of hDicer nuclease expression in wild-type HEK-293T and 293T NoDice cells, before and after transfection with a plasmid producing hDicer. Cells were harvested 48 h after transfection with the plasmid producing the wild-type full-length hDicer and subsequently lysed and analyzed by Western blotting with anti-Dicer antibodies. β-actin was used as a loading control. Dicer's influence on the cleavage of tRFGly(CCC) (**B**), tRF(AAC) (**C**) and tRFArg(CCU) (**D**). Cleavage sites in the D loop have been marked with a blue arrow. (**E**) miR-92A biogenesis in wild-type HEK-293T and 293T NoDice cells and HEK-293T cells after transfection with the plasmid producing hDicer. Pre-miRNA and mature miRNA are marked with orange and grey arrows, respectively.

3. Discussion

In the tested model kidney cells, many RNA oligomers are formed during the biogenesis of tRNA-derived fragments. So, then, what are the decisive factors in the biogenesis of specific tRNA fragments? One of them appears to be the expression of the parental tRNA. Indeed, it has been estimated that there is an excess of 500 tRNA genes in human cells. About half of them have been proven to be actively expressed genes [38]. Environmental conditions, such as diet and stress, influence the expression of tRFs [39,40]. Depending on the biological context, not all tRNAs are processed into tRFs, suggesting some specificity or selectivity in their biogenesis. In the human tRF database, 270 of the tRNA fragments collected had a similar number of tRF-3s and tRF-5s (tRF-3: 62, tRF-3a: 32, tRF-3b: 30, tRF-5: 63 tRF-5a: 22, tRF-5b: 23, tRF-5c: 18, 3′ and 5′ tRNA halves: 113) [41]. tRNA epigenetics is another factor, which influences tRF biogenesis. For example, the presence of 5-methyl cytidine in the tRNA chain protects it from angiogenin cleavage [42]. Recently, Guzzi et al. have observed that pseudouridine synthase 7 (PUS7)-mediated pseudouridylation is a critical regulator of the biogenesis of tRFs in hESCs [43].

The presence of the same tRNA-derived fragments in various organisms, tissues and cells has been confirmed in many reports. In epididymal spermatozoa, the expression of the three most abundant tRFs was as follows: 5-tRFGlu (34.5%), 5-tRFGly (22.0%) and 5-tRFVal (18.1%) [39]. In addition, the expression levels of 5-tRFGly, 5-tRFVal, 5-tRFMet and 5-tRFArg

were significantly increased upon nutritional starvation, while 5-tRFTyr was not induced, thereby confirming the idea of specific 5-tRF biogenesis in different conditions [21]. In a recent study, tRNA halves have been identified in mouse serum using RNA-Seq [44]. The results revealed that 5-tRFGly and 5-tRFVal together account for around 90% of circulating 5-tRFs, while the majority of other 5-tRFs were below the detection limit.

The mechanism of tRF formation has not been fully elucidated. Previous studies carried out by several research groups predicted that, under stress conditions, angiogenin (ANG) cleaves tRNAs into 5′ and 3′ halves [17,45–47]. During normal cell homeostasis, ANG is retained in an inhibited state through its interactions with RNH1 (ribonuclease/angiogenin inhibitor 1). When the interaction is disrupted under stress conditions, it leads to ANG-mediated cleavage of the tRNAs. This observation is supported by a previous observation, that when RNH1 is knocked down, ANG-induced tRNA cleavage increases under oxidative stress conditions, indicating the importance of the ANG/RNH1 interaction in controlling the cleavage process [48]. ANG is likely not to be the only RNase to produce tRNA halves in response to oxidative stress, since they are still present in ANG knockout cells. However, additional RNase(s) await identification. ANG overexpression selectively cleaves a subset of tRNAs, tRNAGly(CCC) and tRNAVal(AAC) among them, to produce tRNA halves and tRF-5s that are 26–30 bases long. While ANG knockout has revealed that a majority of stress-induced tRNA halves, except for the 5′ half from tRNAHis(GTG) and the 3′ half from tRNAAsp(GTC), are ANG-independent, other RNases may be involved in producing tRNA halves.

The experimental results (Figures 3, S1 and S2) revealed the presence of 5′ and 3′ tRNA halves in the unstressed cells. However, at the present stage of research, it is difficult to determine the role of angiogenin nuclease in cutting the anticodon loop. Angiogenin cleaves C–A and U–A sequences [49,50]. In the analyzed tRNAs, the C–A base pair occurs twice, in tRNAGly(CCC) and tRNAVal(AAC), and in both cases it is cleaved. In the remaining cases, the pyrimidine base pairs are cut. Most likely, other nucleases, such as RNase T2, are involved in their cleavage [21]. Together, the results of the present and the previous studies on human cells, yeast and plants suggest that the biogenesis of half of the tRNA requires the presence of angiogenin and/or endoribonucleases from the RNase T2 family.

Additionally, some controversy remains about the role played by Dicer proteins. Early studies of human HeLa cells have shown that the abundance of tRFGln (CUG) decreased in Dicer knockdown cells, suggesting that Dicer is necessary for tRF biogenesis [50]. However, sequencing analysis of the Dicer knockout cells in mice, *Drosophila* and yeast revealed that tRF biogenesis is independent of Dicer [46]. The results presented here show the participation of Dicer in the biogenesis of small RNAs, namely, tRFs and miRNA [51]. The abundance level of tRF-5b RNA fragments is constant in the wild-type and the transfected HEK-293T cells. Although Dicer proteins are unlikely to be major players in small RNA biogenesis, the RNase T2 family has been implicated in tRF-5 biogenesis [21]. This is suggested by the presence of cleavage sites that are not Dicer-specific. An example of this is the additional cleavage at U20 in the D loop, which was present in all of the analyzed tRNAs.

In summary, the role of Dicer in tRF biogenesis is so far unclear (see the Introduction). Our findings showed that knockdown of the dicer gene in the case of NoDicer cells stopped the biogenesis of a 22-nucleotide long tRNA fragment of tRF-5b. This tRF appeared after the transfection of the cells with a plasmid containing the *dicer* gene. This effect is most pronounced for tRFArg(CCU) and less so for tRFGly(CCC). Perhaps Dicer is not responsible for tRF biogenesis on a global scale.

In conclusion, our research revealed that parental tRNA is a source of several tRNA-derived fragments, increasing the pool of small regulatory molecules. For example, during biogenesis, parental tRNAGly(CCC) produced 7 tRNA-derived fragments with different nucleotide lengths, tRFVal(AAC) produced as many as 11 tRFs and tRNAArg(CCU) produced 6 tRFs. The role of these tRNA fragments is unknown. Bioinformatic analysis suggests their participation in the gene expression of a number of proteins found in the

kidneys by binding to mRNA in a manner similar to that of another small RNA, miRNA. However, this hypothesis requires experimental verification.

4. Materials and Methods

4.1. Cell Cultures

All the cell cultures were from ATCC (American Type Culture Collection, Manassas, VA, USA). A-498 and HEK-293T cells were cultured in Minimum Essential Medium Eagle (MEM) supplemented with 15% fetal bovine serum, non-essential amino acids (Gibco Thermo Fisher Scientific, Waltham, MA, USA), 100 U/mL of penicillin G and 0.1 mg/mL of streptomycin sulphate (Sigma-Aldrich St. Louis, Missouri, USA). The cells were cultured at 37 °C, in an atmosphere of 5% CO_2 and 95% air, with a humidity of approximately 95%. The HK-2 cells were cultured in Keratinocyte SFM (serum-free medium) supplemented with the human recombinant epidermal growth factor (rEGF), bovine pituitary extract (BPE), non-essential amino acids (Gibco Thermo Fisher Scientific), 100 U/mL of penicillin G and 0.1 mg/mL of streptomycin sulphate (Sigma-Aldrich). The hESCs were cultured on Geltrex, in supplemented Essential 8 medium. Geltrex was mixed with cold DMEM/F12 at a 1:1 ratio, then poured onto a 60 mm plate. The prepared plate was incubated for 1 h at 37 °C. Essential 8 was supplemented with RevitaCell at a 100:1 ratio. The medium was changed every day. Passage occurred as the colonies began to grow upwards. For passage, DPBS + EDTA in a ratio of 100:1 was used instead of trypsin to detach the cells, then wash medium (DMEM/F12 and BSA in a ratio of 100:1) was used to suspend the detached cells. The cells were cultured at 37 °C, in an atmosphere of 5% CO_2 and 95% air, with the humidity level at approximately 95%.

The 293T NoDice cells (the Dicer knock-out HEK-293T cells) [37] were cultured in DMEM (Gibco Thermo Fisher Scientific) supplemented with 10% FBS (Gibco Thermo Fisher Scientific) and penicillin–streptomycin (100 U/mL of penicillin and 100 µg/mL of streptomycin respectively, and 1 mM sodium pyruvate (Gibco Thermo Fisher Scientific), as described in Bogerd et al. [45]. For the transfection of 293T NoDice cells, the expression plasmid, including the full-length cDNA encoding human transcript variant 2 of DICER1 (NM_030621.4), was used as described previously [51]. Transfection was carried out using DharmaFECT kb DNA Transfection Reagent (Dharmacon, Lafayette, CO, USA), according to the manufacturer's instructions. The 293T NoDice cell line was kindly provided by Prof. Bryan R. Cullen.

4.2. Cell Lysis and Western Blot Analysis

Cells were collected 48 h after transfection, precipitated and resuspended in lysis buffer (30 mM Hepes pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 10% glicerol, 0.5 mM DTT and 0.2% Tergitol) containing 1× protease inhibitor without EDTA (Sigma-Aldrich) and broken by passing through a 0.9 × 40 mm needle. Lysates were centrifuged at 13,000 rpm for 5 min at 4 °C. Cell extracts were separated on 10% SDS-PAGE and electrotransferred onto a PVDF membrane (Gibco Thermo Fisher Scientific). For hDicer detection, the blots were probed with a mouse monoclonal primary anti-Dicer antibody mapping at the C-terminus of hDicer (1:300, Santa Cruz Biotechnology, Dallas, TX, USA); for β-Actin, the blots were probed with a rabbit monoclonal primary anti-β-actin antibody (1:1100, Cell Signaling Technology, Danvers, MS, USA) and subsequently with HRP-conjugated secondary antibody, anti-mouse or anti-rabbit (1:5000, Jackson ImmunoResearch Laboratories, Inc., Cambridgeshire, UK). The immunoreactions were detected using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific, Waltham, MA, USA).

4.3. RNA Isolation

Two different ways of isolating RNA were used. In the first, the total RNA was extracted from the culture cells using the TRIzol LS reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. The concentration and the purity of the total RNA was determined using a NanoDrop™ ND-1000 spectrophotometer (Nan-

oDrop, Thermo Fisher Scientific, Inc., Wilmington, DE, USA). To eliminate redundant modifications that could interfere with the small RNA sequencing library preparation in the subsequent steps, the total RNA was pretreated using the rtStarTM tRF and tiRNA Pretreatment Kit (cat. no. AS-FS-005, Arraystar, Rockville, MD, USA), in accordance with the manufacturer's instructions.

The second small RNA isolation technique involved the use of the mirVana miRNA Isolation Kit, according to the manufacturer's instructions (Invitrogen). The total RNA isolated with the Trizol method was additionally cleared using the mirVana Isolation Kit to discard long RNA and rRNA. The sRNA fraction present in the filtrate was then mixed with 100% ethanol, in the amount of 2/3 of the volume of the filtrate, and purified using a second filter cartridge. The concentration of the eluted sRNAs was measured using a NanoDrop spectrophotometer.

4.4. Library Construction and Illumina Sequencing

The small RNA (sRNA) fractions extracted from the cell culture were subjected to 12% (*w*/*v*) denaturing PAGE (polyacrylamide gel electrophoresis) and the sRNA fragments (18–40 nucleotides) were isolated. In the next step, the RNAs were analyzed using a Bioanalyzer 2100 with a small RNA assay (Agilent, Santa Clara CA, USA). A sequencing library was prepared with 1 mg of each sRNA sample using the TruSeq sRNA Sample Prep Kit (Illumina), according to the TruSeq Small RNA Sample Preparation Guide. The sncRNAs were subsequently ligated with RNA 39 and 59 adapters (each ligation was carried out at 28 °C for 1 h) and reverse-transcribed using SuperScript III Reverse Transcriptase (Invitrogen) at 50 °C for 1 h. The synthesized cDNA was then amplified using Phusion DNA Polymerase, indexed primers and 11 PCR cycles (98 °C 10 s, 60 °C 30 s, 72 °C 15 s) and validated with a High Sensitivity DNA Chip (Agilent). The amplified cDNA library was size-selected using electrophoresis (60 min, 145 V) in 6% Novex TBE PAGE Gel (Invitrogen). Fragments ranging in length from 140 to 190 nt, corresponding to small RNAs (15–70 nt) with both adapters, were extracted from the gel using Gel Breaker tubes (IST Engineering) and incubated in water for 3 h (RT, Rotator Mixer RM-Multi 1 (Starlab)). Then, the libraries were purified on 5 mm filter tubes (IST Engineering) and concentrated by ethanol (100%, 3.25 vol.) and sodium acetate (3 M, 1/10 vol.) precipitation. After a second validation with a High Sensitivity DNA Chip (Agilent), the libraries were quantified using a Qubit fluorometer (Invitrogen Waltham, MA, USA). All the libraries were prepared in duplicate. Sequencing was performed by Macrogen, Inc. (Seoul, Korea).

4.5. Bioinformatic Analysis

Quality reports were generated with FastQC v0.11.8. The reads were then subject to adapter trimming with Cutadapt. Then, they were filtered for quality using fastq_quality_filter from FASTX-Toolkit, with –q 20 –p 95 parameters, i.e., a minimum 95% of bases in a read are required to achieve a Phred quality score of 20 or higher. Individual sequence counts were determined using DESeq. Sequence mapping was performed with the bowtie tool comparing the GtRNAdb database with the RNA-Seq results. The result list was restricted to sequences of more than 10 counts using alignment2.py. Sequences were compared using the mean of the repetitions. Big data analysis was carried out using the "dplyr" library. The "VennDiagram" library was used to create the Venn diagram; the remaining graphs were made with the "ggplot" library.

4.6. DNA Labeling and Northern Blotting

The DNA labeling reaction mixture contained a buffer (50 mM Tris–HCl pH 7.6, 10 mM $MgCl_2$, 5 mM of DTT, 0.1 mM spermidine), 10 U T4 polynucleotide kinase, 20 mCi of (γ-^{32}P)ATP with an activity of 5000 Ci/mmol (Hartmann Analytic, Braunschweig, Germany) and 20 pmol of a DNA oligomer (Table S1). After an hour-long incubation at 37 °C, the ^{32}P labelled products were purified using a G25 column (Sigma-Aldrich, St. Louis, MO, USA). The radioactivity level of the labeled molecules was measured using a scintilla-

tion counter. The total RNA isolated from the different cultures was separated on 12% (*w*/*v*) polyacrylamide gels, electrotransferred onto Hybond-Nylon membranes (GE Healthcare, Chicago, IL, USA), crosslinked with UV light (120 mJ/cm^2) and prehybridized in a 2xSSC, 1X Denhard solution for 1 h at 37 °C. Hybridization of ^{32}P labelled DNA probes (20×10^6 cpm) was carried out at 42 °C overnight in PerfectHyb^{TM} Plus solution (Sigma). Then, the hybridization mixture was discarded, and the blot was washed in the same solution several times, until the radioactivity in the solution disappeared. Finally, the blots were analyzed using phosphor imaging screens and a FLA-5100 image analyzer with MultiGauge software (FujiFilm).

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms23073644/s1.

Author Contributions: M.K., M.W., A.K.-K., M.W., B.I., T.O. and J.W. planned the experiments and wrote the manuscript, prepared the figures and reviewed the manuscript. M.K., E.B., M.W. and P.Ż.-M. carried out all of the experiments. J.W. supervised and accepted the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Polish National Centre of Science, with grant 2016/21/NZ2/03468 to J.W. and grant 2016/22/E/NZ1/00422 to A.K.-K.

Acknowledgments: We would like to thank Bryan R. Cullen for providing the Dicer-deficient 293T NoDice human cell lines and Aga Imiołczyk for proofreading the text. We also thank the reviewers for comments on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

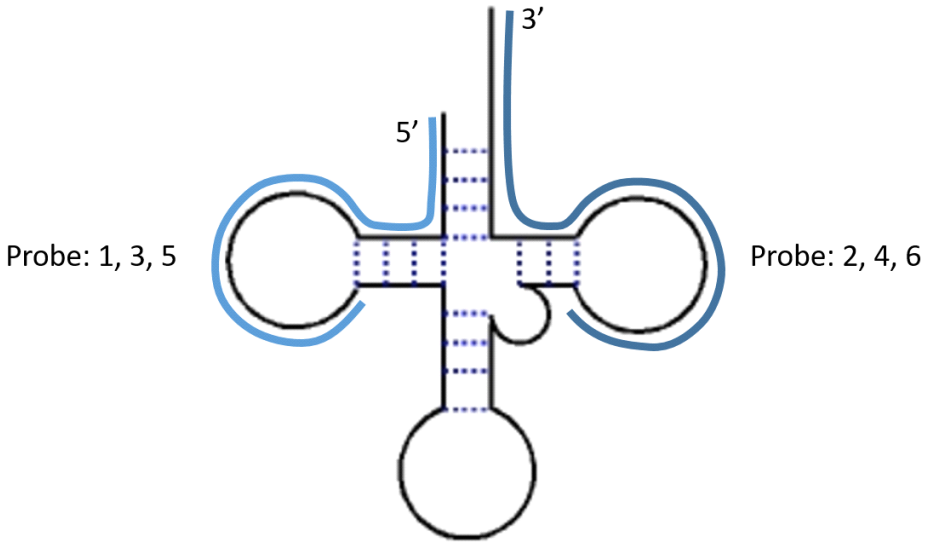
References

1. Li, J.; Liu, C. Coding or noncoding, the converging concepts of RNAs. *Front. Genet.* **2019**, *10*, 496. [CrossRef] [PubMed]
2. Kapranov, P.; Willingham, A.T.; Gingeras, T.R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **2007**, *8*, 413–423. [CrossRef] [PubMed]
3. Iwasaki, Y.W.; Siomi, M.C.; Siomi, H. PIWI-Interacting RNA: Its Biogenesis and Functions. *Ann. Rev. Biochem.* **2015**, *84*, 405–433. [CrossRef] [PubMed]
4. Uchida, S.V.; Adams, J.C. Physiological roles of non-coding RNAs. *Am. J. Physiol. Cell* **2019**, *317*, C1–C2. [CrossRef] [PubMed]
5. Verma, P.S.; Agarwal, V.K. *Cell Biology, Genetics, Molecular Biology, Evolution and Ecology: Evolution*; S Chand & Co., Ltd.: New Delhi, India, 2004.
6. Layton, E.; Fairhurst, A.-M.; Griffiths-Jones, S.; Grencis, R.K.; Roberts, I.S. Regulatory RNAs: A universal language for inter-domain communication. *Int. J. Mol. Sci.* **2020**, *21*, 8919. [CrossRef] [PubMed]
7. Kazimierczyk, M.; Kasprowicz, M.K.; Kasprzyk, M.E.; Wrzesinski, J. Human long noncoding RNA interactome: Detection, characterization and function. *Int. J. Mol. Sci.* **2020**, *21*, 1027. [CrossRef] [PubMed]
8. Berg, M.D.; Brandl, C.J. Transfer RNAs: Diversity in form and function. *RNA Biol.* **2021**, *18*, 316–339. [CrossRef] [PubMed]
9. Krol, J.; Loedige, I.; Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **2010**, *11*, 597–610. [CrossRef] [PubMed]
10. O'Brien, J.; Hayder, H.; Zayed, Y.; Peng, C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Front. Endocrinol.* **2018**, *9*, 402. [CrossRef]
11. Fire, A.; Xu, S.Q.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **1998**, *391*, 806–811. [CrossRef] [PubMed]
12. Yu, X.; Xie, Y.; Zhang, S.; Song, X.; Xiao, B.; Yan, Z. tRNA-derived fragments: Mechanisms underlying their regulation of gene expression and potential applications as therapeutic targets in cancers and virus infections. *Theranostics* **2021**, *11*, 461–469. [CrossRef] [PubMed]
13. Avcilar-Kucukgoze, I.; Kashina, A. Hijacking tRNAs from translation: Regulatory functions of tRNAs in mammalian cell physiology. *Front. Mol. Biosci.* **2020**, *7*, 610–617. [CrossRef] [PubMed]
14. Lee, Y.S.; Shibata, Y.; Malhotra, A.; Dutta, A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev.* **2009**, *23*, 2639–2649. [CrossRef]
15. Zhang, J.; Ferré-D'Amaré, A.R. The tRNA elbow in etructure, recognition and evolution. *Life* **2016**, *6*, 3. [CrossRef] [PubMed]
16. Chen, Q.; Zhang, X.; Shi, J.; Yan, M.; Zhou, T. Origins and evolving functionalities of tRNA-derived small RNAs. *Trends Biochem. Sci.* **2021**, *46*, 790–804. [CrossRef] [PubMed]

17. Su, Z.; Wilson, B.; Kumar, P.; Dutta, A. Noncanonical roles of tRNAs: tRNA fragments and beyond. *Annu. Rev. Genet.* **2020**, *54*, 47–69. [CrossRef] [PubMed]
18. Xie, Y.; Yao, L.; Yu, X.; Ruan, Y.; Li, Z.; Guo, J. Action mechanisms and research methods of tRNA-derived small RNAs. *Signal Transduct. Target Ther.* **2020**, *5*, 109. [CrossRef] [PubMed]
19. Jarrous, N.; Mani, D.; Ramanathan, A. Coordination of transcription and processing of tRNA. *FEBS J.* **2021**. [CrossRef] [PubMed]
20. Rossetti, G.; Richman, T.R.; Perks, K.; Ermer, J.A.; Kuznetsova, I.; Hughes, L.; Shearwood, A.J.; Viola, H.M.; Hool, L.C.; Rackham, O.; et al. Concerted regulation of mitochondrial and nuclear non-coding RNAs by a dual-targeted RNase Z. *EMBO Rep.* **2018**, *19*, e46198.
21. Magee, R.; Rigoutsos, I. On the expanding roles of tRNA fragments in modulating cell behavior. *Nucleic Acids Res.* **2020**, *48*, 9433–9448. [CrossRef] [PubMed]
22. Krishna, S.; Raghavan, S.; DasGupta, R.; Palakodeti, D. tRNA-derived fragments (tRFs): Establishing their turf in post-transcriptional gene regulation. *Cell Mol. Life Sci.* **2021**, *78*, 2607–2619. [CrossRef] [PubMed]
23. Oberbauer, V.; Schaefer, M.R. tRNA-derived small RNAs: Biogenesis, modification, function and potential impact on human disease development. *Genes* **2018**, *9*, 607. [CrossRef]
24. Goodarzi, H.; Liu, X.; Nguyen, H.C.; Zhang, S.; Fish, L.; Tavazoie, S.F. Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell* **2015**, *161*, 790–802. [CrossRef] [PubMed]
25. Falconi, M.; Giangrossi, M.; Zabaleta, M.E.; Wang, J.; Gambini, V.; Tilio, M.; Bencardino, D.; Occhipinti, S.; Belletti, B.; Laudadio, E.; et al. A novel 3′-tRNA Glu -derived fragment acts as a tumor-suppressor in breast cancer by targeting nucleolin. *FASEB J.* **2019**, *33*, 13228–13240. [CrossRef]
26. Gebetsberger, J.; Wyss, L.; Mleczko, A.M.; Reuther, J.; Polacek, N.A. tRNA-derived fragment competes with mRNA for ribosome binding and regulates translation during stress. *RNA Biol.* **2017**, *14*, 1364–1373. [CrossRef] [PubMed]
27. Guan, L.; Karaiskos, S.; Grigoriev, A. Inferring targeting modes of Argonaute-loaded tRNA fragments. *RNA Biol.* **2020**, *17*, 1070–1080. [CrossRef] [PubMed]
28. Keam, S.P.; Young, P.E.; McCorkindale, A.L.; Dang, T.H.; Clancy, J.L.; Humphreys, D.T.; Preiss, T.; Hutvagner, G.; Martin, D.I.; Cropley, J.E.; et al. The human Piwi protein Hiwi2 associates with tRNA-derived piRNAs in somatic cells. *Nucleic Acids Res.* **2014**, *42*, 8984–8995. [CrossRef] [PubMed]
29. Kazimierczyk, M.; Jędroszkowiak, A.; Kowalczykiewicz, D.; Szymanski, M.; Imiołczyk, B.; Ciesiołka, J.; Wrzesinski, J. tRNA-derived fragments from the Sus scrofa tissues provide evidence of their conserved role in mammalian development. *Biochem. Biophys. Res. Commun.* **2019**, *520*, 514–519. [CrossRef] [PubMed]
30. Pan, Q.; Han, T.; Li, G. Novel insights into the roles of tRNA-derived small RNAs. *RNA Biol.* **2021**, *17*, 2157–2167. [CrossRef]
31. Guzzi, N.; Bellodi, C. Novel insights into the emerging roles of tRNA-derived fragments in mammalian development. *RNA Biol.* **2020**, *17*, 1214–1222. [CrossRef]
32. Sun, X.; Yang, J.; Yu, M.; Yao, D.; Zhou, L.; Li, X.; Qiu, Q.; Lin, W.; Lu, B.; Chen, E.; et al. Global identification and characterization of tRNA-derived RNA fragment landscapes across human. *NAR Cancer* **2020**, *2*, zcaa031. [CrossRef]
33. Chen, L.; Wang, F. An emerging role of the 5′ termini of mature tRNAs in human diseases: Current situation and prospects. *Biochim. Biophys. Acta Mol. Basis Dis.* **2022**, *1868*, 166314.
34. Turowski, T.W.; Tollervey, D. Transcription by RNA polymerase III: Insights into mechanism and regulation. *Biochem. Soc. Trans.* **2016**, *44*, 1367–1375. [CrossRef] [PubMed]
35. Xiao, Q.; Gao, P.; Huang, X.; Chen, X.; Chen, Q.; Lv, X.; Fu, Y.; Song, Y.; Wang, Z. tRFTars: Predicting the targets of tRNA-derived fragments. *J. Transl. Med.* **2021**, *19*, 88. [CrossRef]
36. Iwakawa, H.O.; Tomari, Y. Life of RISC: Formation, action, and degradation of RNA-induced silencing complex. *Mol. Cell* **2022**, *82*, 30–43. [CrossRef]
37. Bogerd, H.P.; Whisnant, A.W.; Kennedy, E.M.; Flores, O.; Cullen, B.R. Derivation and characterization of Dicer- and microRNA-deficient human cells. *RNA* **2014**, *20*, 923–937. [CrossRef]
38. Li, N.; Shan, N.; Lu, L.; Wang, Z. tRFtarget: A database for transfer RNA-derived fragment targets. *Nucleic Acids Res.* **2021**, *49*, D254–D260. [CrossRef]
39. Sharma, U.; Conine, C.C.; Shea, J.M.; Boskovic, A.; Derr, A.G.; Bing, X.Y.; Belleannee, C.; Kucukural, A.; Serra, R.W.; Sun, F.; et al. Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* **2016**, *351*, 391–396. [CrossRef] [PubMed]
40. Dhahbi, J.M.; Spindler, S.R.; Atamna, H.; Yamakawa, A.; Boffelli, D.; Mote, P.; Martin, D.I. 5′ tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and calorie restriction. *BMC Genom.* **2013**, *14*, 298. [CrossRef]
41. Saikia, M.; Krokowski, D.; Guan, B.J.; Ivanov, P.; Parisien, M.; Hu, G.f.; Anderson, P.; Pan, T.; Hatzoglou, M. Genome-wide identification and quantitative analysis of cleaved tRNA fragments induced by cellular stress. *J. Biol. Chem.* **2012**, *287*, 42708–42725. [CrossRef] [PubMed]
42. Schaefer, M.; Pollex, T.; Hanna, K.; Tuorto, F.; Meusburger, M.; Helm, M.; Lyko, F. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev.* **2010**, *24*, 1590–1595. [CrossRef] [PubMed]

43. Guzzi, N.; Ciesla, M.; Ngoc, P.C.T.; Lang, S.; Arora, S.; Dimitriou, M.; Pimková, K.; Sommarin, M.N.; Munita, R.; Lubas, M.; et al. Pseudouridylation of tRNA. Derived Fragments Steers Translational Control in Stem Cells. *Cell* **2018**, *173*, 1204–1216.e26. [CrossRef]
44. Cole, C.; Sobala, A.; Lu, C.; Thatcher, S.R.; Bowman, A.; Brown, J.W.; Green, P.J.; Barton, G.J.; Hutvagner, G. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* **2009**, *15*, 2147–2160. [CrossRef] [PubMed]
45. Lyons, S.M.; Fay, M.M.; Akiyama, Y.; Anderson, P.J.; Ivanov, P. RNA biology of angiogenin: Current state and perspectives. *RNA Biol.* **2017**, *14*, 171–178. [CrossRef]
46. Su, Z.; Kuscu, C.; Malik, A.; Shibata, E.; Dutta, A. Angiogenin generates specific stress-induced tRNA halves and is not involved in tRF-3–mediated gene silencing. *J. Biol. Chem.* **2019**, *294*, 16930–16941. [CrossRef]
47. Shen, Y.; Yu, X.; Zhu, L.; Li, T.; Yan, Z.; Guo, J. Transfer RNA-derived fragments and tRNA halves: Biogenesis, biological functions and their roles in diseases. *J. Mol. Med.* **2018**, *96*, 1167–1176. [CrossRef]
48. Sarangdhar, M.A.; Allam, R. Angiogenin (ANG)—Ribonuclease Inhibitor (RNH1) system in protein synthesis and disease. *Int. J. Mol. Sci.* **2021**, *22*, 1287. [CrossRef] [PubMed]
49. Giantazopoulou, V.I.; Georgiou, S.; Kakoulidis, P.; Giannakopoulou, S.N.; Tseleni, S.; Stravopodis, D.J.; Anastasiadou, E. From the Argonauts Mythological Sailors to the Argonautes RNA-Silencing Navigators: Their Emerging Roles in Human-Cell Pathologies. *Int. J. Mol. Sci.* **2020**, *21*, 4007. [CrossRef]
50. Garcia-Silva, M.R.; Cabrera-Cabrera, F.; Güida, M.C.; Cayota, A. Hints of tRNA-Derived Small RNAs Role in RNA Silencing Mechanisms. *Genes* **2012**, *3*, 603–614. [CrossRef]
51. Wojnicka, M.; Szczepanska, A.; Kurzynska-Kokorniak, A. Unknown areas of activity of human ribonuclease Dicer: A putative deoxyribonuclease activity. *Molecules* **2020**, *25*, 1414. [CrossRef]

Supplementary material



No.	Probe name	Sequence (5’ –> 3’)
1.	5'tRF Gly CCC	CGTAACCACCAAGTTACCAT
2.	3'tRF Gly CCC	AAGTTAAGGGCCGGTTACGT
3.	5'tRF Val AAC	CAAAGGCATCACATCACCAA
4.	3'tRF Val AAC	AAGCTTTGGCCCGCCTTTGT
5.	5'tRF Arg CCT	CGGGGTCACCGGATTACCTA
6.	3'tRF Arg CCT	CAAGCTCAGGGTGGACCCCAC
7.	5S RNA	AGGGTGGTATGGCCGTAGAC

Table S1 Hybridization sites of DNA probes and their sequences.

Table S2A. Computer analysis of the possibility of kidney protein regulation by 5' tRFGly(CCC)

5'tRF	tRFdb ID	Length	Positive probability	Gene	Protein	Function in the kidneys
GlyCCC	5004b	22 nt	0,8696	MEIS2	Homeobox protein Meis2	• Overexpression during aging and kidney damage [1] • Participates in nephrogenesis [2] • Meis1 is required for correct vascular network formation in the embryo [3]
			0,8685	SH3TC	SH3 domain and tetratricopeptide repeat-containing protein 2	• miR-584, which is located within the SH3TC intron, is a tumor suppressor, directly targeting the Rock-1 oncogene and reducing the invasion capacity in human clear cell renal cell carcinoma [4]
			0,7550	FMN1	Formin-1	• Genomic rearrangements of the FMN1 locus cause kidney defects [5] • Associated with acro-renal disorders [6]
			0,7099	CTDSPL2	CTD small phosphatase-like protein 2	• Dephosphorylation of the FoxOs protein by stimulating proteolysis during nephritis [7]
	5004c	31 nt	0,8077	ZBTB20	Zinc finger and BTB domain-containing protein 20	• Glucose homeostasis [8]
			0,7730	KSR2	Kinase suppressor of Ras 2	• Cell energy homeostasis [9]
			0,7018	SEC62	Translocation protein SEC62	• Part of the ER transmembrane proteins responsible for the post-translational import

						of small presecretory proteins [10]

Table S2B. Computer analysis of the possibility of kidney protein regulation by 3' tRFGly(CCC)

3'tRF	tRFdb ID	Length	Probability	Gene	Protein	Function in the kidneys
GlyCCC	3027a	tRF3b 17nt	1,0000	ORAI2	Protein orai-2	• Transporting the Ca^{2+} ion[11]
			0,9868	CCDC9B0	Coiled-coil domain-containing protein 9B	• Cell morphology, cellular assembly and organization, cellular function and maintenance [12,13]
			0,9854	CYP20A1	Cytochrome P450 20A1	• Involved in diabetic nephropathy [14]
			0,9841	CACNG8	Voltage-dependent calcium channel gamma-8 subunit	• Involved in transporting calcium ions [15]
			0,9806	GAS7	Growth arrest-specific protein 7	• Involved in proteinuria and chronic kidney disease [16]
	3027b	22	0,9404	PLD6	Phospholipase D Family Member 6	• Member of a family of proteins correlated with Kidney Fibrosis [17]
			0,8715	LY6G6C	Lymphocyte antigen 6 complex locus protein G6c	• LY6G6C belongs to a cluster of leukocyte antigen-6 (LY6) genes located in the major histocompatibility complex (MHC) class III region [18]

1. Chang-Panesso, M; Kadyrov, F.F; Machado, F.G; Kumar, A; Humphreys, B.D. Meis1 is specifically upregulated in kidney myofibroblasts during aging and injury but is not required

for kidney homeostasis or fibrotic response. *Am J Physiol Renal Physiol*. **2018**, *315* (2), F275–F290.

2. Dekel, B; Metsuyanim, S; Schmidt-Ott, K.M; Fridman, E; Jacob-Hirsch, J; Simon, A; Pinthus, J; Mor, Y; Barasch, J; Amariglio, N; et al. Multiple imprinted and stemness genes provide a link between normal and tumor progenitor cells of the developing human kidney. *Cancer Res*. **2006**, *66* (12), 6040 – 6049.
3. Azcoitia, V; Aracil, M; Martínez-A, C; Torres, M. The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. *Dev Biol*. **2005**, *280* (2), 307-320.
4. Ueno, K; Hirata, H; Shahryari, V; Chen, Y; Zaman, M.Z; Singh, K; Tabatabai, Z.L; Hinoda, Y; Dahiya, R. Tumour suppressor microRNA-584 directly targets oncogene Rock-1 and decreases invasion ability in human clear cell renal cell carcinoma. *British Journal of Cancer*. **2011**, *104*, 308–315.
5. Dimitrov, B.I; Voet, T; De Smet, L; Vermeesch, J.R; Devriendt, K; Fryns, J.P; Debeer, P. Genomic rearrangements of the GREM1-FMN1 locus cause oligosyndactyly, radio-ulnar synostosis, hearing loss, renal defects syndrome and Cenani-Lenz-like non-syndromic oligosyndactyly. J Med Genet. **2010**, 47 (8), 569-574.
6. Natarajan, G; Jeyachandran, D; Subramaniyan, B; Thanigachalam, D; Rajagopalan, A. Congenital anomalies of kidney and hand: a review. Clin Kidney J. **2013**, 6 (2), 144-149.
7. Liu, X; Yu,R; Sun, L; Garibotto, G; Lin, X; Wang, Y; Thomas, S.S; Li, R; Hu, Z. The nuclear phosphatase SCP4 regulates FoxO transcription factors during muscle wasting in chronic kidney disease. Kidney Int. **2017**, 92 (2), 336-348.
8. Zhang, Y; Xie, Z; Zhou, L; Li, L; Zhang, H; Zhou, G; Ma, X; Herrera, P.L; Liu, Z; Grusby, M.J; Zhang, W.J. The zinc finger protein ZBTB20 regulates transcription of fructose-1,6-bisphosphatase 1 and β cell function in mice. Gastroenterology. **2012**, 142 (7), 1571-1580.
9. Guo, L; Costanzo-Garvey, D.L; Smith, D.R; Neilsen, B.K; MacDonald, R.G; Lewis, R.E. Kinase Suppressor of Ras 2 (KSR2) expression in the brain regulates energy balance and glucose homeostasis. Mol Metab. **2016**, 6 (2), 194-205.
10. Lang, S; Benedix, J; Fedeles, S.V; Schorr, S; Schirra, C; Schäuble, N; Jalal, C; Greiner, M; Hassdenteufel, S; et al. Different effects of Sec61α, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. J Cell Sci. **2012**, 125 (Pt 8), 1958-1969.
11. Mercer, J.C; Dehaven, W.I; Smyth, J.T; Wedel, B; Boyles, R.R; Bird, G.S; Putney Jr, J.W. Large store-operated calcium selective currents due to co-expression of Orai1 or Orai2 with the intracellular calcium sensor, Stim1. *J Biol Chem*. **2006**, *281*, 24979-24990.
12. Sassano, A; Mavrommatis, E; Arslan, A.D; Kroczynska, B; Beauchamp, E.M; Khuon, S; Chew, T.L; Green, K.J; Munshi, H.G; et al. Human Schlafen 5 (SLFN5) Is a Regulator of Motility and Invasiveness of Renal Cell Carcinoma Cells. *Molecular and Cellular Biology*. **2015**, *35*, 2684 – 2698.
13. Wang, L; Ma, S; Cui, Y; Yang, Y; Wang,D. Genes And Pathways Associated With Apparent Treatment-resistant Hypertension In Patients With Chronic Kidney Disease: An Analysis Of The Cric Study. *Hypertension*. **2020**, *76*, AP125.
14. Kiyanpour, F; Abedi, M; Gheisari, Y. A systematic integrative approach reveals novel microRNAs in diabetic nephropathy. *J Res Med Sci*. **2020**, *25*, 1.
15. Sharma, J; Deb, B; George, I.A; Kapil, S; Coral, K; Kakkar, N; Pattanaik, S; Mandal, A.K; Mavuduru, R.S; Kumar, P. Somatic Mutations Profile of a Young Patient With Metastatic

Urothelial Carcinoma Reveals Mutations in Genes Involved in Ion Channels. *Front Oncol*. **2019**, *9*, 435

16. Wuttke, M; Wong, C.S; Wühl, E; Epting, D; Luo, L; Hoppmann, A; Doyon, A; Li, Y; et al. Genetic loci associated with renal function measures and chronic kidney disease in children: the Pediatric Investigation for Genetic Factors Linked with Renal Progression Consortium. *Nephrol Dial Transplant*. **2016**, *31*, 262-269.
17. Trivedi, P; Kumar, R.K; Iyer, A; Boswell, S; Gerarduzzi, C; Dadhania, V.P; Herbert, Z; Joshi, N; Luyendyk, J.P; et al. Targeting Phospholipase D4 Attenuates Kidney Fibrosis. *J Am Soc Nephrol*. **2017**. *28*, 3579-3589.
18. Mallya, M; Campbell, R.D; Aguado, B. Transcriptional analysis of a novel cluster of LY-6 family members in the human and mouse major histocompatibility complex: five genes with many splice forms. *Genomics*. **2002**, *80*, 113-123.

Table S3A. Kidney protein regulation by 5'tRFVal(AAC)

5'tRF	tRFdb ID	Length	Probability	Gene	Protein	Function in the kidneys
Val(AAC)	5017b	tRF-5b 20nt	1,0000	NFIC[1]	Nuclear transcription factor	• NFIC may regulate FABP7 expression in RCC [1]
			0,9985	GNAO1	Guanine nucleotide-binding protein G(o) subunit alpha	• GNAO1 are likely to be involved in the regulation of RCC through G protein signal transduction [2]
			0,9984	HIPK2	Homeodomain interacting protein kinase 2	• Regulator of transcription factors, modulates growth, development, morphogenesis, and cell death in renal cells [3]

Table S3B. Kidney protein regulation by 3'tRFVal(AAC)

3'tRF	tRFdb ID	Length	Probability	Gene	Protein	Function in the kidneys
Val(AAC)	3008	tRF-3b 17nt	0,9946	FOXD1	Forkhead box protein K1	• Foxd1 is required during kidney development [4] • Inactivation results in the failure of nephron progenitor cell differentiation
			0,9821	ARHGAP45	Rho GTPase-activating protein 45	• Expression of ARHGAP24 results in inhibited cell proliferation [5] • Reduces cell invasion ability • Induces apoptosis in renal cancer cells.

			0,9394	TOX23	TOX high mobility group box family member 2	• The genes are linked to biological pathways involving solute transport and renal physiology [6]
			0,9393	TPCN2	Two pore calcium channel protein 2	• Tpcn2 mRNA expression was the highest in kidneys, liver and adipose tissue [7]

1. Takaoka, N. Analysis of the regulation of fatty acid binding protein 7 expression in human renal carcinoma cell lines. *BMC Mol. Biol*. **2021**, 12, 31.
2. Liu X, Wang J, Sun G. Identification of key genes and pathways in renal cell carcinoma through expression profiling data. Kidney Blood Press Res. 2015;40(3):288-97.
3. Fan Y, Wang N, Chuang P, He JC. Role of HIPK2 in kidney fibrosis. Kidney Int Suppl. 2014, 4, 97-101.
4. Fetting JL, Guay JA, Karolak MJ, Iozzo RV, Adams DC, Maridas DE, Brown AC, Oxburgh L. FOXD1 promotes nephron progenitor differentiation by repressing decorin in the embryonic kidney. Development. 2014 Jan;141(1):17-27.
5. Xu G, Lu X, Huang T, Fan J. ARHGAP24 inhibits cell cycle progression, induces apoptosis and suppresses invasion in renal cell carcinoma. *Oncotarget*. 2016;7(32):51829-51839.
6. Sima, C., Iordache, P., Poenaru, E., Manolescu, A., Poenaru, C., Jinga, V. Genome-wide association study of nephrolithiasis in an Eastern European population. *Int Urol Nephrol.* 53, 309–313 (2021).
7. Tunn RE Expression of two-pore channels in mammalian primary cells and tissues, and their role in adipose tissue formation and function. eTHOS 2012

International Journal of
Molecular Sciences

MDPI

Review

Human Long Noncoding RNA Interactome: Detection, Characterization and Function

Marek Kazimierczyk, Marta K. Kasprowicz, Marta E. Kasprzyk and Jan Wrzesinski *

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland; mkazimierczyk@man.poznan.pl (M.K.); mkasprowicz@ibch.poznan.pl (M.K.K.); martha.kasprzyk@gmail.com (M.E.K.)

* Correspondence: wrzesinj@ibch.poznan.pl

Received: 3 January 2020; Accepted: 2 February 2020; Published: 4 February 2020

Abstract: The application of a new generation of sequencing techniques has revealed that most of the genome has already been transcribed. However, only a small part of the genome codes proteins. The rest of the genome "dark matter" belongs to divergent groups of non-coding RNA (ncRNA), that is not translated into proteins. There are two groups of ncRNAs, which include small and long non-coding RNAs (sncRNA and lncRNA respectively). Over the last decade, there has been an increased interest in lncRNAs and their interaction with cellular components. In this review, we presented the newest information about the human lncRNA interactome. The term lncRNA interactome refers to cellular biomolecules, such as nucleic acids, proteins, and peptides that interact with lncRNA. The lncRNA interactome was characterized in the last decade, however, understanding what role the biomolecules associated with lncRNA play and the nature of these interactions will allow us to better understand lncRNA's biological functions in the cell. We also describe a set of methods currently used for the detection of lncRNA interactome components and the analysis of their interactions. We think that such a holistic and integrated analysis of the lncRNA interactome will help to better understand its potential role in the development of organisms and cancers.

Keywords: long noncoding RNA; lncRNA; lncRNA interactome; detection lncRNA interactome; lncRNA interactome function

1. Introduction

It was estimated from the data obtained during the ENCODE project that 70–80% of the human genome has been transcribed [1], but that only 2% of the genome codes proteins [2]. The function of the genome "dark matter", non-coding RNAs (ncRNA), has not yet been fully recognized [3]. During the last few years, the application of RNA-seq methods has made it possible to display a divergent repertoire of ncRNAs. ncRNA is classified based on the length of the transcripts [4]. Transcripts containing over 200 nucleotides are considered to be lncRNAs and RNAs smaller than 200 nucleotides were arbitrarily classified as sncRNAs, which encompass miRNA, siRNA, piRNA and tRF [5].

Micro RNAs (miRNAs) are small RNAs made up of 21 to 23 nucleotides with well-defined biogenesis and maturation pathways [6]. They play a pivotal role in controlling gene expression [7,8]. siRNAs perform a similar function, although they originate from long precursors, i.e., repetitive and transposon sequences of the genome [9]. piRNAs (Piwi-interacting RNAs) are the largest group of sncRNAs and are mainly present in germ cells [10–12]. The mechanisms of piRNA biogenesis involve several Piwi proteins, resulting in 30–35 nt single-stranded RNAs. This guards germ cells against transposon activity [12]. Moreover, abnormal expression of Piwi proteins and piRNAs results in the lack of fertility in males [10–12]. The newly discovered group of snRNA, tRF RNAs, are essentially

tRNA fragments [13]. The production of some tRFs under stress conditions is involved in translation repression [13,14].

Much less is known about the biology of lncRNA, which constitutes a significant part of the non-coding genome [4,15]. Recent results achieved using high-through sequencing technologies indicated a high level of diversity in lncRNA classes [16]. They are very heterogeneous in terms of size as lncRNAs number from several hundred to several thousand nucleotides. Furthermore, lncRNA transcripts are derived from different parts of the genome, and lncRNAs are localized in the nucleus or in the cytoplasm. A good example of this is the well-defined nuclear Xist [17] and cytoplasmic Uchl1 lncRNAs [18].

Usually, lncRNAs are transcribed by RNA polymerase II, and like mRNA, are 5′-capped, often spliced and polyadenylated [19]. Generally, compared to mRNA, lncRNAs display modest sequence conservation [20,21]. The reason for this may be that lncRNAs are free of codon preservation constraints [21]. Therefore, the sequence of lncRNAs appears to be less important than their secondary/tertiary structure, which plays a pivotal role [20,21].

In this paper, we present recently discovered information about the lncRNA interactome i.e., the cellular components that interact with lncRNA. We describe the interaction of the biomolecules, such as nucleic acids, sncRNA (miRNA), mRNA and DNA with lncRNA. A significant part of the review concerns the importance of lncRNA-protein complexes. Recently, a part of the lncRNA interactome has been found to be made out of short peptides, encoded inside lncRNA, and even small molecules. Their possible impact on lncRNA functions will be discussed. Recently, a part of the lncRNA interactome has been found to be made out of short peptides, encoded inside lncRNA, and even small molecules (Figure 1).

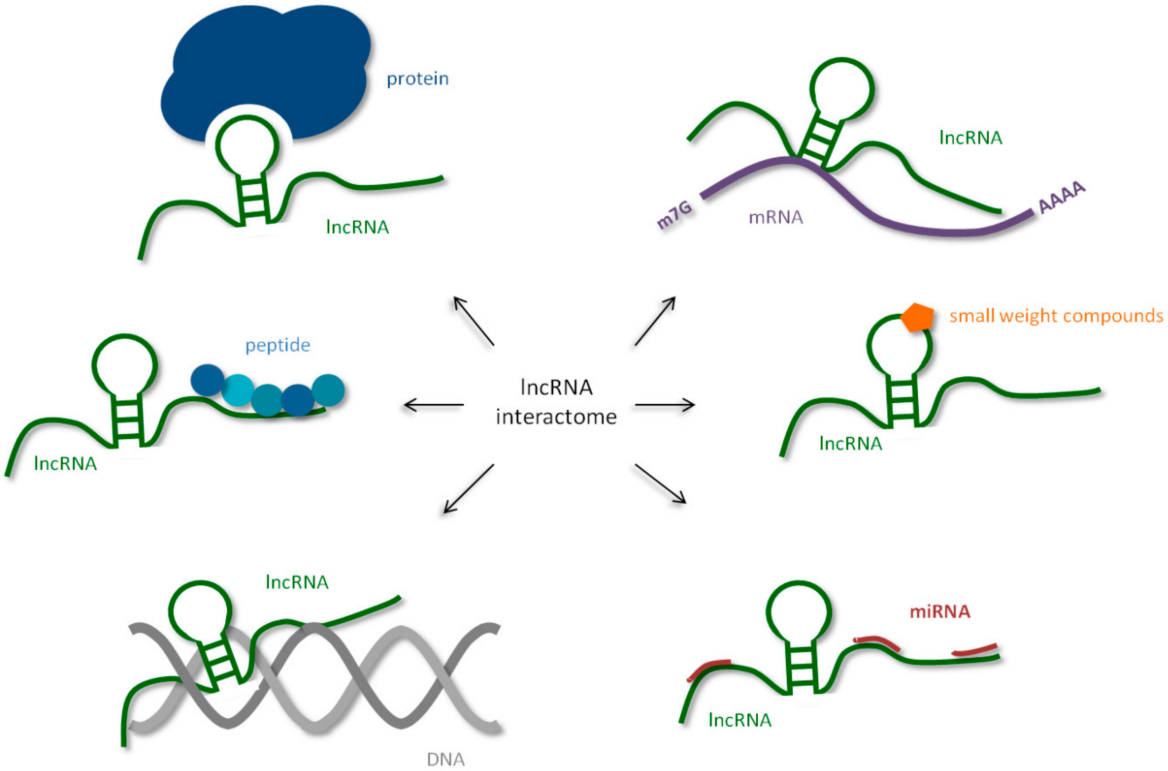


Figure 1. Human long non-coding RNAs (lncRNA) interactome, interaction of human lncRNA with cellular biomolecules.

lncRNAs have been discovered to be involved in divergent functions in the human organism, *inter alia*, during cell development and differentiation [15,22,23]. In addition, it appears that the mutation

and dysregulation of many lncRNAs may be connected to serious and complex human diseases (Table 1).

Table 1. Examples of pathogenesis related to lncRNA dysregulation.

Disease	LncRNA	Impact on Pathogenesis	Mechanism
Colon cancer	DACOR1 [24,25]	Downregulated in colon tumors	Interacts with and inhibits DNA methyltransferase DNMT1
Lung cancer	HOTAIR [26]	Impacts proliferation, survival, invasion, metastasis, and drug resistance in lung cancer cells	HOTAIR may promote dedifferentiation of lung epithelial cells through two distinct mechanisms, i.e., transcriptional repression of *Hox5* and ubiquitin-mediated proteolysis of Ataxin-
Prostate cancer	LINCRNA-p21 [27]	Decreases prostate cancer cell proliferation	LINCRNA-p21 inhibits many genes expression in a p53-dependent transcriptional response
Parkinson's disease	H19, LincRNA-p21, MALAT1, and SNHG1 [28]	H19 is significantly downregulated in Parkinson's disease while LincRNA-p21, Malat1 and SNHG1, are significantly upregulated.	Associated with synaptogenesis, proliferation, apoptosis, precedes Parkinson's disease
Leukemia	MALAT1 [29]	Inhibiting multiple myeloma growth	Involved in multiple myeloma DNA repair and cell death.
Cardiovascular diseases	GAS5 [30]	Promotes the development and progression of myocardial infarctions	Targeting of the miR-525-5p/CALM2 axis
Diabetes	HI-LNC901, PLUTO [31]	Implicated in pancreatic islet function	Regulates the transcription of *PDX1* gene of pancreatic β cell
AIDS	LINC00173 [32]	Regulates cytokines in T cells	Presumably involved in transcriptional regulation

Up to 70–80% of the human genome has been transcribed. However, only about 2% of the genome includes protein-coding genes (mRNA) [1,2]. ncRNAs account for the majority of the genome transcripts. The term ncRNA is usually used to refer to RNA that does not encode proteins, however this does not mean that such RNAs do not carry any information or have any functions [16,33,34]. ncRNAs are divided into housekeeping RNAs and regulatory RNAs, based on their functions. Housekeeping noncoding RNAs, including transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and ribosomal RNAs (rRNAs), are commonly expressed in a constitutive manner. Regulatory RNAs are ncRNAs with a strong regulatory impact on the expression of protein-coding genes. Based on their size, regulatory RNAs can be divided into two groups: sncRNAs (<200 nt) and lncRNAs (>200 nt). Small noncoding RNAs (sncRNAs) are a group which encompasses microRNAs (miRNAs), small interfering RNAs (siRNAs), Piwi associated RNAs (piRNAs), tRNA-derived fragments (tRFs) and circular RNAs (circRNAs). The second group, lncRNAs, which are larger than 200 nucleotides, as has been mentioned previously, do not have the ability to code protein.

Most lncRNAs have been found to be synthesized by RNA polymerase II under the control of the transcriptional activators of the SWI/SNF complex. However, some lncRNAs are transcribed by RNA polymerase III [19,35]. As has been already mentioned, the transcripts are capped, spliced and polyadenylated [36,37].

2. The LncRNA Interactome: The Nucleic Acid Story

2.1. Interactions of LncRNA with miRNAs, the ceRNA Hypothesis

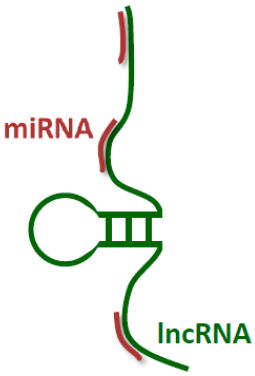
Despite over two decades of research, the role of non-coding RNAs in human development still remains a mystery. Most of the information available concerns the regulation of the activity of selected genes. There is a lot of evidence that suggests miRNAs contribute to this process, binding to specific mRNA 3′ UTR regions and regulating the expression of these genes [7]. In 2011, Salmena et al. proposed a competitive, endogenous RNA (ceRNA) hypothesis [38], which is supported by a significant amount

of experimental evidence [39–41]. According to this hypothesis, ncRNAs and miRNAs influence each other. When it comes to the mRNA pool, transcribed pseudogenes, lncRNA, circRNA and other RNAs, there is competition for the same pool of miRNA. When miRNA binds to mRNA, a "seed sequence" containing 2–8 nucleotides, which ensures efficient miRNA interaction, also referred to as the miRNA responsive element (MRE), is required [7,8,42]. It is already known that each mRNA may contain multiple MREs, and thus can be regulated by a number of miRNAs, while one miRNA can potentially regulate dozens of mRNAs. Many experimental findings support the idea that multiple ncRNAs, including sncRNAs, lncRNAs, and circRNAs, as well as pseudogenes, can act as so-called miRNA "sponges". By sharing identical MREs and competing for common miRNAs, they change miRNA's activity, which results in modified mRNA translation [41,43].

ceRNAs regulate each other through interactions with shared miRNAs, creating a large-scale regulatory network across the transcriptome, significantly expanding the functional genetic information in the genome (Table 2). In addition, the ceRNAs play an essential role in many biological processes, which is why the destruction of the balance between ceRNAs and miRNAs functions as a regulator. The lack of this balance plays a significant part in disease development and is found in many types of cancer [44].

Table 2. Selected lncRNAs interacting with microRNA (miRNA) and their functions.

	LncRNA	Interacting miRNA	Function in the Cell
	LINC-ROR	miR-138, miR-145, miR-204 [45,46]	competing endogenous RNAs
	LNCRNA-KRTAP5-AS1, LNCRNA-TUBB2A	specific miRNA [47]	
	CASC2	miR-21, miR-18a [48,49]	
	CDR1AS/CIRS-7	miR-671, miR-7 [50,51]	
	LINC-MD1	miR-133, miR-135 [52]	
	MDRL	miR-361 [53]	
	HULC	miR-372 [54]	
	LINC-223	miR-125-5p [55]	
	LNCARSR	mir-34, miR-449 [56]	
	LNCND	miR-143-3p [57]	
	UFC1	miR-34a [58]	cell cycle
	LINC00152	miR-138 [59]	
	MALAT1	miR-101, miR-217, miR-9, miR-125b [60–62]	controlling proliferation and senescence
	UCA1	miR-1 [63]	
	BACE1-AS	mir-485-5p [64]	modulation of mRNA stability
	H19	miR-106a, miR-17-5p, miR-20b, let-7 [65,66]	transcriptional regulation
	HOTAIR	miR-34a, miRNA-141, miR-130a, miRNA-let7 [67–69]	
	MEG3	miRNA-29 [70]	
	GAS5	miR-21 [71]	
	HOST2	let-7b [72]	
	PCAT-1	miR-3667-3p [73]	post-transcriptional regulation
	LINCRNA-P21	let-7 [74]	modulation of translation



One important tool used to confirm miRNA–ceRNA interactions is the in silico analysis of the MREs shared by mRNA and ceRNA, such as lncRNA or circRNA. Several computer analysis

approaches, such as MARIO, PARIS, LIGR etc, have already been successfully used in the past for performing such studies [75–77].

ncRNAs interact with each other during the biological processes that take place in the cell. In order to determine the interaction framework of different RNA molecules, and to determine which RNA molecules can anneal or hybridize to each other in the cell, the potential RNA–RNA interactions must be verified experimentally, in vivo. To achieve this goal, high-throughput sequencing of RNA, isolated by crosslinking and immunoprecipitation (HITS-CLIP) and the application of photoactivatable ribonucleoside enhanced crosslinking, as well as immunoprecipitation (PAR-CLIP) methods, were used [65,78,79]. PAR-CLIP, which is a modification of the HITS-CLIP method, utilizes the UV radiation of cell cultures, in the presence of psoralen derivatives [53,76,79]. These compounds specifically react with RNA, but not with proteins, and generate inter-strand cross-links between RNA's uridine bases. It has been shown that the integration of these crosslinking methods significantly enhances (over 20-fold) the search efficiency in terms of RNAs interacting with ceRNAs in liver cells, in comparison to *in silico* studies alone [80]. A more precise insight into miRNA–lncRNA interactions has recently been gained through the use of a combination of microarrays and the NGS method [81]. Using microarrays makes it possible for the expression of many sncRNA and lncRNAs to be analyzed in a single experiment. However, more information can be obtained through the analysis of NGS data. Although using both of these methods together is more informative, their high cost can be prohibitive.

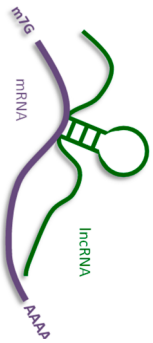
The question is whether the ceRNA hypothesis can explain ncRNA's role in the progression of cancer. It has been calculated that more than 30% of miRNAs are involved in cancer regulation [82–84]. Evidence demonstrates that lncRNAs may acts as miRNAs decoy molecules, and may regulate its activity influencing cellular processes, including those that are associated with cancer [39,85]. Comprehensive analyses of ceRNA networks involving lncRNA-associated miRNA have been conducted for many diseases, such as ovarian and prostate cancer, glioblastoma, thyroid carcinoma, as well as breast, lung, kidney and gut cancers [39,86,87]. Usually, lncRNAs were aberrantly expressed and significantly correlated with the cancer prognosis. Recently, it has been shown that the MALAT1 lncRNA can sponge miR-211 as a ceRNA, and potentially up-regulate the PHF19 protein, a component of the polycomb complex, thus facilitating the progression of ovarian cancer [88]. Another HOTTIP lncRNA sponges miR-216a-5p, promoting prostate cancer cell proliferation, migration and invasion [89].

Although there are well-documented examples confirming the ceRNA hypothesis, some scientists find it controversial due to the fact that the expression of an additional mir122 target did not affect its abundance in the hepatocytes and liver [90]. In addition, this observation was confirmed for other ceRNAs; blocking the miRNA binding using antimir oligonucleotides had no physiological influence on miRNA function [91].

2.2. Pairing LncRNAs with Messenger RNAs

However, there is another mechanism for regulating RNA activity through its direct base pairing with lncRNA (Table 3). LncRNA and mRNA and the pre-mRNA complementary hybrid may participate in regulating translation by affecting mRNA splicing and editing as well as its stability [92,93].

Table 3. Some examples of lncRNA–mRNA pairing.



LncRNA	mRNA	Impact on
SAF ZEB NAT MALAT	FAS [92] ZEB [92,93] CAMK2B, CDK7, SAT1, HMG2L1, ARHGEF1, B-MYB [94]	alternative splicing
PCA3	PRUNE [95]	A–I mRNA editing
LAST BACE1 LINCRNA p21	CCDI [96] BACE1 [97] JUNB, CTNNB1 [98]	mRNA stability

Most of the information concerning lncRNA–mRNA interactions come from bioinformatics analysis software such as: RNAplexn [99] and LncTar [100]. However, lncRNA–mRNA interactions need experimental verification *via* RNA probing or the aforementioned RNA-RNA crosslinking.

It has been predicted that lncRNA–pre-mRNA interactions may play an important role in alternative splicing. Since almost 90% of human genes are spliced alternatively, controlling this process is important for the development of organisms [92]. lncRNA may effectively regulate splicing; indeed, computer analyses indicate that out of about 24,500 genes, some 21,000 may be affected by the formation of lncRNA–mRNA duplexes. LncRNA may affect pre-mRNA splicing in two ways: by blocking spliceosome assembly involving intron-exon junction or by becoming the target for splicing factors [92]. For example, it has been suggested that MALAT1 lncRNA regulates the alternative splicing of pre-mRNA by controlling the functional levels of transcription factors [94].

The lncRNA-mRNA duplex is a good substrate for RNA adenosine deaminase, a double stranded specific enzyme, which converts adenine to inosine [101]. Inosine differs from adenine in that it possesses a carbonyl group instead of an amino group at position 6 of the purine ring. Such an RNA modification influences the base pairing. Adenine forms a base pair with uracil, unlike inosine, which pairs with cytosine. In the case of lncRNA A to I editing, most of the information comes from bioinformatics analyses [102]. It has been suggested that almost 200,000 editing sites occur in human lncRNAs, and the majority of them (65%) are located within the sites, which significantly changes their secondary structure. Editing may alter the target sites of the lncRNAs and, therefore, the edited and non-edited transcripts may differ in terms of their functions [95,101].

Additionally, lncRNAs influence mRNA stability. Recently Cao et al. reported that lncRNA-assisted stabilization of transcripts (LAST) binds to the 5′ UTR region of *CCND1*mRNA and protects it against possible nuclease targeting in cultured cells [94]. It appears that LAST interacts with other mRNAs. The overexpression of LAST lncRNA has also been observed in several cancer tissues.

2.3. LncRNA–DNA Interactions

Several mechanisms of lncRNA recruitment to genomic targets in lncRNA-DNA interactions have been proposed [103] (Table 4). One of them is the direct binding to DNA and the formation of a triple helix. The nature of the RNA-DNA-DNA triple helix formation is described in references [104,105]. Triplexes are formed by weaker, non-Watson-Crick base-pairs, Hoogsteen hydrogen bonds or reverse Hoogsteen hydrogen bonds between the Watson-Crick base-paired dsDNA and the third strand made up of RNA. Triplexes can be either parallel or antiparallel, based on the orientation of the third strand. The orientation of the third strand might be important for its functionality [106,107]. A couple of useful computer tools, which can be used to predict RNA-dsDNA triplex formation and lncRNA–DNA binding sites, such as GRIDseq [108], Triplexator [109], or LongTarget [110] are available. Several techniques, which can be used to search for lncRNA - DNA binding sites, such as Chromatin Isolation by RNA Purification (ChIRP), Chromatin Oligo Affinity Precipitation (ChOP), Capture Hybridisation

Analysis of RNA Target (CHART) and RNA Antisense Purification (RAP), have been reviewed by Chu et al. [111] and Vance and Ponting [112].

Table 4. Examples of lncRNA interacting with DNA and their postulated function.

LncRNA	Interacting DNA Region	Function
AIR	Slc22a3 promoter [113]	Epigenetic control of gene expression
FENDRR	Foxf1, Pitx2 promoters [114,115]	
TUNA	Nanog, Sox2 and Fgf4 promoters [116]	
MEG3	TGF-β pathway genes [117]	
PARTICLE	Upstream of MATSA promoter [118]	
NEAT1	Multiple binding sites [119,120]	Paraspecle formation
MALAT1	Multiple binding sites [121]	Alternative splicing regulation, promotes metastasis
LncRNA DHFR	DHFR promoter [122,123]	Transcription regulation
TERRA	Telomers [124–126]	Telomer replication control

There are several examples where lncRNA can form a triple helix with dsDNA, such as FENDRR, DHFR, Khps, PARTICLE or NEG3. FENDRR lncRNA forms a triplex structure with gene (*Foxf1, Pitx2*) promoters, creating binding sites for the polycomb repressive complex (PRC2) and regulating the expression of its target genes [114,115]. FENDRR plays an important role in carcinogenesis. The decrease of this lncRNA is associated with gastric cancer. LncRNA DHFR has been shown to inhibit the transcription of *Dfhr* mRNA by creating a triplex structure with the DHFR promoter [122,123]. This interaction results in lncRNA DHFR binding to the TFIIB transcription factor, preventing the formation of a transcription initiation complex. The human heart tissue-specific lncRNA, Khps1, interacts with a homopurine stretch, upstream of the promoter of sphingosine kinase SPHK1, and the recruitment of p300/CBP (histone acetyltransferase) [113]. p300/CBP changes chromatin's state to active, which allows transcription factor E2F1 to bind and enhance SPHK1 expression. Promoter of MAT2A-antisense radiation-induced circulating lncRNA (PARTICLE) is expressed in response to low-dose irradiation [127]. It has been shown to form a triplex upstream of the methionine adenosyltransferase promoter (MAT2A). PARTICLE leads to the methylation of MAT2A by G9a and PRC2 complexes, which represses transcription. MEG3 binding sites have been shown to contain higher levels of GA-rich sequences. These sites help to guide MEG3 to its targets, by forming RNA-DNA triplexes [117]. Moreover, MEG3 regulates the activity of the TGF-β genes cooperating with the PRC2 complex. AIR lncRNA, an antisense promoter transcript located in intron 2 of the insulin-like growth-factor type-2 receptor gene, has been found to occupy gene *Slc22a3* promoter of cation transporter and recruit H3K9 histone methyltransferase G9a to epigenetically repress transcription [113]. Numerous genomic binding sites, such as NEAT1 and MALAT1, have also been identified for lncRNAs [119–121]. Most of the sites include active chromatin at highly expressed genes. Telomeric repeat-containing RNAs (TERRA), also known as TelRNAs, has been associated with telomeric chromatin, thus its involvement in telomere shortening has also been proposed [124–126]. In vitro experiments revealed that TERRA inhibits telomerase activity and is evolutionary conserved.

3. LncRNA–Protein Interactome

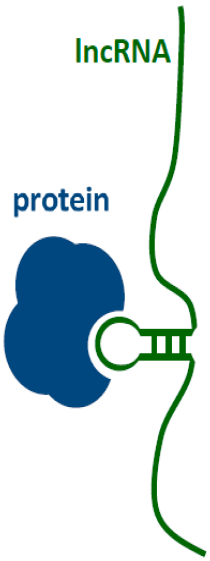
By defining the proteins that interact with lncRNA, it is possible to gain some insight into the molecular pathways, in which the lncRNA of interest might be involved [128]. Due to their significant size, reaching up to tens of thousands of nucleotides, lncRNAs have been shown to act as guides, signals, decoys and scaffolds for many different proteins [129,130]. Thus, it is important to detect which proteins

form complexes with specific lncRNAs. There are a couple of approaches to try out when looking for lncRNA–protein interactions. The method to be chosen depends, among other things, on whether we focus on RNAs (RNA-centric methods) or proteins (protein-centric methods) [131]. Electrophoretic Mobility Shift Assay [EMSA] was the first method used to analyze RNA–protein interactions [132]. A ^{32}P, fluorescent or chemiluminescent labeled lncRNA probe is incubated with cell lysate proteins and analyzed using non-denaturing, polyacrylamide gel electrophoresis [132–134]. The so-called pull down assay works in a similar fashion [131,135]. The biotin-labeled RNA is incubated with a cell lysis protein mixture to bind the interacting proteins. Subsequently, the complex is immobilized on streptavidin-agarose, purified, and detected using Western blotting [136]. Currently, there are three methods allowing for the detection of the lncRNA - protein interactome in use. The first method, immunoprecipitation, is usually used to prepare the RNA protein complexes. The second method combines crosslinking and immunoprecipitation (CLIP) [137,138]. This approach makes it possible to identify the proteins bound to the selected lncRNA, and to determine the protein binding sites in the lncRNA sequence. The last of the methods, Next Generation Sequencing (NGS), is used to analyze lncRNA - protein interactions, which significantly speeds up the analysis of RNA-protein complexes. There are several variants of the CLIP method, like HITS-CLIP, PAR-CLIP, iCLIP, etc. HITS-CLIP was developed as a genome-wide tool, designed to map protein-RNA binding sites in vivo [139]. The Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) method uses cells cultured in the presence of nucleotide analogs, such as 4-thiouridine (s^4U) or 6-thioguanosine (s^6G). Thio-nucleotide analogs are incorporated into the newly synthesized RNAs. The crosslinking of proteins to modified lncRNAs makes it possible to precisely determine the type of protein and its binding site in the lncRNA sequence [140].

Because sequencing is an inherent part of all of them, it is crucial to pick the right data analysis methods in order to obtain the necessary information. In silico analysis makes the prediction of the structure, function or interaction of lncRNA-protein complex possible through screening the annotated sequences or structural motifs, like RNA-binding sites. There are also many user-friendly lncRNA databases, which collect information about sequences, which have already been annotated, such as lncRNAdb [141], LNCipedia [142], or NONCODE 2018 [143].

Due to the wide range in lncRNA size, which span from several hundred to several thousand nucleotides, there are many different sites proteins may bind to. Indeed, in the cell lncRNAs exist mainly in the form of RNA-protein complexes (Table 5). It has been suggested that RNA molecules are the perfect scaffolds for protein binding [144]. An estimated 5% of lncRNAs may bind to about half of the interacting proteins. As has been shown in Table 3, there are lncRNAs that specifically bind one protein, and also lncRNAs which bind multiple proteins. For example, many proteins interact with HOTAIR lncRNA [34,145–147]. However, lncRNA, lnc-DC interacts only with the STAT 3 protein [148].

Table 5. lncRNA-protein complexes involved in the regulation of cellular functions.



LncRNA	LncRNA Length	Interacting Protein(s)	Function
KCNP1ot1	91.7 kb	PRC1, PRC2, G9a [149,150]	Epigenetic control of gene expression
XIST	19 kb	PRC2, hnRNP U, YY1 [151,152]	
AIR	4.3 kb	G9a [153,154]	
ANRIL	3.8 kb	PRC1 [155]	
HOTTIP	3.8 kb	WDR5, MLL [156]	
HOTAIR	2.4 kb	PRC2, LSD1/CoREST/REST [145–147]	
H19	2.3 kb	PRC2 [157]	
DEANR1	4.9 kb	SMAD2/3 [158]	Transcription regulation
LINCRNA-p21	2.7 kb	HnRNP [159,160]	
PANDA	1.5 kb	SAFA [161,162]	
GAS5	1.0 kp	GR (NR3C1) [163]	
LNC-DC	630 bp	STAT-3 [148]	
7SK	330 bp	P-TEFb [164,165]	
MIAT (Gomafu)	10 kb	SF1 [166]	Alternative splicing
MALAT (NEAT2)	8.7 kb	SR [167]	
PNKY	1.6 kb	hnRNP [168–170]	
BC1	152 bp	EIF4A/eIF4B, PABP [171]	Translation regulation
THRIL	2.9 kb	HnRNP [172]	Immune response

The results of the analyses presented in Table 3 clearly show that lncRNA-protein complexes regulate many cell processes: transcription and splicing, as well as gene expression. It should be emphasized that many human lncRNAs interact with the polycomb repressive protein complexes (PRC1 and PRC2) [173,174]. Both protein complexes interact with chromatin and are involved in chromatin remodeling as well as the epigenetic regulation of gene expression via DNA and histone modification [175]. A variety of molecular investigations authenticate the association between lncRNAs—such as HOTAIR, Kcnq1, Air and chromatin remodeling complexes, such as PRC1 and PRC2, which mediate ubiquitination and histone methylation, respectively (Table 4). At the *Kcnq1* locus, the lncRNA Kcnq1ot1 interacts with members of the PRC1 and PRC2 complex proteins [149,150]. At the *Igf2r* locus, Air lncRNA associates with the histone methyltransferase, G9a [154]. HOTAIR recruits PRC2 at the *HOXD* locus, to induce the silencing of the target genes [145,147]. Another example is the heterogeneous nuclear ribonucleoprotein complexes (hnRNP), which encompass several RNA binding proteins involved in gene expression, including pre-mRNA processing, mRNA stability, and translation [129,176]. However, a recent analysis of significant amounts of literature, concerning the interactions of PKC2 proteins with lncRNAs, showed that these interactions are either promiscuous or that the methods used to detect them have a lot of noise [177].

The interactions of the hnRNP protein with various lncRNAsplay a significant role in many cell functions [178]. As mentioned in Table 4, hnRNP interacts with LINC–p21 as well as PNKY lncRNA, thus regulating transcription and alternative splicing, respectively [159,160]. In addition, the THRIL lncRNA (TNFα and hnRNPL related immunoregulatory lncRNA) plays a key role in innate immune responses as well as in inflammatory diseases in humans [172]. To summarize, lncRNAs are targets of many proteins, whereas lncRNA-protein complexes perform many functions in the cell by participating in various cellular processes (Table 5).

4. Peptides—A New Factor in the LncRNA Interactome?

LncRNAs are RNAs defined as having a size exceeding 200 nucleotides and being a non-coding part of the genome. This means that lncRNAs do not encode proteins, i.e., they do not harbor an open reading frame. However, recent studies have revealed that a subset of lncRNAs that code small peptides, usually shorter than 100 amino acids, exists [179–181]. These lnRNAs are localized in the cytosol and contain only a single-exon-sequence coding peptide.

The detection of lncRNA-origin peptides is not easy, because computer analyses predict many Open Reading Frames (ORF), however, only a few of them are actively translated. It has been calculated that about 23% of the transcribed lncRNAs have been translated [182]. Moreover, the expression of many peptides is weak, and they are difficult to detect, especially in the case of peptides that lack sequence homologies to known proteins. Many methods for detecting peptide expression combine bioinformatic algorithms and experimental verification. An excellent overview of bioinformatics tools used to search for potential ORF in lncRNAs was put forth by Choi et al. [183]. The simplest way of verifying the predicted ORF involves in vitro translation methods. Constructing a template and performing translation with the use of rabbit reticulocyte, with ^{35}S methionine and SDS electrophoresis, makes it possible to determine the size of the predicted peptide [184]. This is important, because due to its size (>200 nucleotides), lncRNA contains many AUG start codons and the encoded peptides vary in length.

The identification of peptides in protein samples isolated from tissues or culture cells, is possible by applying the Western blot method with the appropriate antibodies, following peptide purification. Moreover, the development of mass spectrometry techniques has made this method perfect for identifying such peptides [185]. Another approach involves the use of the ribosome profiling, in connection with NGS methods [186]. In this approach, the RNAs bound to polysomes are purified using sucrose gradient centrifugation, and then, after nuclease digestion, the RNA fragment protected by ribosomes is recovered. These RNA fragments are sources for library preparation and sequencing and the use of bioinformatics tools in genome mapping, which makes the identification of the ORF region possible.

The fundamental issue associated with these bifunctional RNAs is the determination of the role of lncRNA and its coding peptide. The initial question is whether they act as a complex or as separate compounds. The next question is whether the maternal lncRNA, which contains ORF, interacts with its own peptide.

Analysis of available data indicates that the size of the peptide is not connected with the size of the maternal lncRNAs. An 8.7 kb long MALAT1 lncRNA codes a peptide containing 213 amino acids, while the four times shorter XGAT1 lncRNA [2.1 kb] codes a peptide containing 210 amino acids, which is identical in size. [183,187]. Dissection impact on the cell selected lncRNA and translated from its ORF peptide is difficult. Recently, the role some peptides play in cell functions has been explained. The 46 amino acid peptide, myoregulin, which is encoded by a 16.5 kb long skeletal, muscle-specific lncRNA, LINC00948, regulates the regeneration of skeletal muscles by interacting with sarcoplasmic reticulum Ca^{2+}-ATPase (SERCA) [188], lysosomal v-ATPase [189] or removing SERCA inhibitors [190].

Moreover, only the HOXB-AS3 peptide, not lncRNA, is critical for suppressing colon cancer growth, by blocking pyruvate kinase M [PKM] splicing, miR-18a processing, and the subsequent glucose metabolism reprogramming [191]. On the other hand, *HOXB-AS3* lncRNA could also regulate the cell cycle progression of OCI-AML3 cells in Npm1 mutated acute myeloid leukemia. Observations suggest that there might be an unknown interaction between the lncRNA and peptides, however, this hypothesis requires further investigation.

5. Do Small Compounds Influence LncRNA Activity?

Many RNAs form constrained structures containing "pockets" that bind small molecular weight compounds. This repertoire of RNAs includes long RNAs and small RNAs, such as rRNA and mRNA, as well as ribozymes, riboswitches, and aptamers, which were discovered during the last twenty

years [192,193]. The question is whether chemical compounds with a low molecular weight, which regulate lncRNA activity by structure-specific binding, really exist.

The recently determined 3.1 Å resolution MALAT1 crystal structure revealed a constrained structure, containing a bipartite triple helix at the 3′ end [194]. The disruption of the stability of the helix by a point mutation resulted in MALAT1 accumulation in the cell [195]. The recognition of the triple helix by a methyltransferase enzyme strongly supported its functional role [196]. Small molecules, which destabilize this MALAT1 structural element, make it possible to regulate lncRNA functions [197]. XIST is one of the best-characterized lncRNAs. It functions as the major effector of the X chromosome inactivation (XCI) process in mammals. Specific structural domains, namely six tandem hairpin repeats (A–F), are crucial to its functioning in the XCI process. Structural insight into the A and F repeat region of the hairpins revealed an intricate architecture within specific functional modules [198]. In cellulo chemical probing of the entire 18k.b. transcript enabled the discovery of an additional domain at the 3′ end, connected with XIST localization [199]. Decoding the structure should make it possible to identify small molecules, which have the ability to recognize lncRNA's structural elements. It will also elucidate the role they play in development and disease.

Research on the lncRNA structure using low molecular weight compounds (DMS, DEPC, Pb2+ ions, etc) has shown it is conserved [200]. This means that structural elements, such as loops, bulges and base-paired regions, which occur in small RNAs, are also found numerous times in lncRNA and may be places where various small molecules bind. Information concerning the interactions of small compounds with lncRNA is still limited. Fatemi et al. identified small molecules, which bind to the lncRNA-protein complex, using high-throughput compound screening methods, i.e., the Amplified Luminescent Proximity Homogeneous Assay [201]. They reported on the specific and quantifiable binding of the brain-derived neurotrophic factor antisense lncRNA to a component of the PCR2 complex, protein EZH2, and also identified a small-molecule inhibitor–ellipticine, that upregulates its downstream target genes. It was also reported that telRNAs, which form the G-quadruplex, target alkaloid quindoline derivatives [202]. The binding of this compound inhibits proliferation and causes G2/M phase arrest in osteosarcoma cancer cells as well as induces DNA damage response and apoptosis. Recently, it has been shown by Shi et al. [203] that the binding of fluorescent peptide derivative NP-C86 to GAS5 lncRNA stabilizes its structure.

6. Concluding Remarks and Perspectives

Although 70–80% of the genome has been transcribed, it has been discovered that only 2% of the genome encodes protein sequences. 80% of the remaining ncRNA pool is made up of lncRNA. As this review shows, lncRNAs are not just cellular junk but are involved in many processes such as transcription, translational regulation or in cell development in general. Long noncoding RNA has been shown to interact with a range of cell biomolecules such as other RNAs (miRNAs, mRNAs) and DNA to form the lncRNA interactome, which is involved in life processes. A significant part of the lncRNA interactome is associated with the formation of complexes with proteins or even, as has been recently discovered, peptides. Given the size of the entire lncRNA interactome, the examples that have been studied to date are just the tip of the iceberg. The fact that interference disorders in the lncRNA interactome are the cause of various pathogenesis pathways, including cancer, neurodegenerative, and immunological diseases, is also an impetus for further research.

In the future, one can expect new lncRNA annotation tools will be developed, such as new DNA and RNA sequencing methods (nanopore sequencing) and new bioinformatics methods [204,205]. The development of new methods of analysis, like mass spectroscopy makes it possible to detect biomolecules that bind to lncRNA, including proteins, peptides and low-molecular compounds. It is also important to determine how the lncRNA interactome influences cellular processes, as well as the impact it has on various stages of the organism's development. New DNA editing methods (CRISP/Cas9 and RNA (CRISP/Cas13 [206,207] could be helpful in achieving these goals. The destruction of genes encoding proteins or lncRNA and miRNA seems to be a powerful method of

testing the lncRNA interactome. It seems particularly important to determine the role of lncRNA in the pathogenesis of various diseases and as a biomarker of the disease state, for example, during the cancerogenesis processes. However, modulating lncRNA activity with low-molecular compounds (alkaloids, antibiotics, peptides, etc.) may be used for therapeutic purposes in the future.

To summarize, the lncRNA interactome is a large group of biomolecules, related to and interacting with lncRNA, that play a variety of roles in cell development and the pathogenesis of various diseases.

Author Contributions: M.K., M.K.K. and M.E.K. wrote the manuscript, prepared the figures and reviewed the manuscript. J.W. supervised and finally accepted the manuscript. All the authors have read and approved the manuscript submitted to the journal. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Polish National Centre of Science, grant number 2016/21/NZ2/03468 to J.W.

Acknowledgments: We thank B. Imiołczyk for critically reading the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Xue, C. Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108. [CrossRef] [PubMed]
2. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Funke, R. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [PubMed]
3. Palazzo, A.F.; Lee, E.S. Non-coding RNA: What is functional and what is junk. *Front. Genet.* **2015**, *6*, 2. [CrossRef]
4. Hombach, S.; Kretz, M. Non-coding RNAs: Classification, biology and functioning. *Adv. Exp. Med. Biol.* **2016**, *937*, 3–17.
5. Cech, T.R.; Steitz, J.A. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* **2014**, *157*, 77–94. [CrossRef] [PubMed]
6. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.; Mello, C.C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **1998**, *391*, 806–811. [CrossRef] [PubMed]
7. Krol, J.; Loedige, I.; Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **2010**, *11*, 597–610. [CrossRef]
8. Pu, M.; Chen, J.; Tao, Z.; Miao, L.; Qui, X.; Wang, Y.; Ren, J. Regulatory network of miRNA on its target: Coordination between transcriptional and post-transcriptional regulation of gene expression. *Cell. Mol. Life Sci.* **2019**, *76*, 441–451. [CrossRef]
9. Huyn, S. Small RNA pathways that protect the somatic genome. *Int. J. Mol. Sci.* **2017**, *18*, 912. [CrossRef]
10. Pillai, R.S.; Chuma, S. piRNAs and their involvement in male germline development in mice. *Dev. Growth Differ.* **2012**, *54*, 78–92. [CrossRef]
11. Weick, E.M.; Miska, E.A. piRNAs: From biogenesis to function. *Development* **2014**, *141*, 3458–3471. [CrossRef] [PubMed]
12. Iwasaki, Y.W.; Siomi, M.C.; Siomi, H. PIWI-interacting RNA: Its biogenesis and functions. *Annu. Rev. Biochem.* **2015**, *84*, 405–433. [CrossRef] [PubMed]
13. Gebetsberger, J.; Polacek, N. Slicing tRNAs to boost functional ncRNA diversity. *RNA Biol.* **2013**, *10*, 1798–1806. [CrossRef] [PubMed]
14. Oberbauer, V.; Schaefer, R.S. tRNA-Derived Small RNAs: Biogenesis, Modification, Function and Potential Impact on Human Disease Development. *Genes* **2018**, *9*, 607. [CrossRef]
15. Wu, H.; Yang, L.; Chen, L.L. The diversity of long noncoding RNAs and their generation. *Trends Genet.* **2017**, *33*, 540–552. [CrossRef]
16. Uszczyńska-Ratajczak, B.; Lagarde, J.; Frankish, A.; Guido, R.; Johnson, R. Towards a complete map of the human long non-coding RNA transcriptome. *Nature Rev. Genet.* **2018**, *18*, 535–548. [CrossRef]

17. Engreitz, J.M.; Pandya-Jones, A.; McDonel, P.; Shishkin, A.; Sirokman, K.; Surka, C.; Kadri, S.; Xing, J.; Goren, A.; Lander, E.S.; et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **2013**, *341*, 1237973. [CrossRef]
18. Carrieri, C.; Cimatti, L.; Biagioli, M.; Beugnet, A.; Zucchelli, S.; Fedele, S.; Pesce, E.; Ferrer, I.; Collavin, L.; Santoro, C.; et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **2012**, *491*, 454–457. [CrossRef]
19. Quinn, J.J.; Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **2017**, *17*, 47–62. [CrossRef]
20. Chodroff, R.A.; Goodstadt, L.; Sirey, T.M.; Oliver, P.L.; Davies, K.E.; Green, E.D.; Molnár, Z.; Ponting, C.P. Long noncoding RNA genes: Conservation of sequence and brain expression among diverse amniotes. *Genome Biol.* **2010**, *11*, R72. [CrossRef]
21. Necsulea, A.; Kaessmann, H. Evolutionary dynamics of coding and non-coding transcriptomes. *Nat. Rev. Genet.* **2014**, *15*, 734–748. [CrossRef] [PubMed]
22. Yan, P.; Luo, S.; Lu, J.Y.; Shen, X. Cis- and trans-acting lncRNAs in pluripotency and reprogramming. *Curr. Opin. Genet. Dev.* **2017**, *46*, 170–178. [CrossRef] [PubMed]
23. Sun, C.; Huang, L.; Li, Z.; Leng, K.; Xu, Y.; Jiang, X.; Cui, Y. Long non-coding RNA MIAT in development and disease: A new player in an old game. *J. Biomed. Sci.* **2017**, *25*, 23. [CrossRef] [PubMed]
24. Morlando, M.; Fatica, A. Alteration of Epigenetic Regulation by Long Noncoding RNAs in Cancer. *Int. J. Mol. Sci.* **2018**, *19*, 570. [CrossRef] [PubMed]
25. Jang, Y.Y.; Junjie, P.; Sanjun, C.; Ma, Y. Long non-coding RNAs in colorectal cancer: Progression and future directions. *J. Cancer* **2017**, *8*, 3212–3225.
26. Loewen, G.; Jayawickramarajah, J.; Zhuo, Y.; Shan, B. Functions of lncRNA HOTAIR in lung cancer. *J. Hematol Oncol.* **2014**, *7*, 90. [CrossRef] [PubMed]
27. Wang, X.; Ruan, Y.; Wang, X.; Zhao, W.; Jiang, Q.; Jiang, C. Long intragenic non-coding RNA lincRNA-p21 suppresses development of human prostate cancer. *Cell Prolif.* **2017**, *50*, e12318. [CrossRef]
28. Li, L.; Zhuang, Y.; Zhao, X.; Li, X. Long non-coding RNA in neuronal development and neurological disorders. *Front. Genet.* **2018**, *9*, 744. [CrossRef]
29. Hu, Y.; Lin, J.; Fang, H.; Fang, J.; Li, C.; Chen, W.; Liu, S.; Ondrejka, S.; Gong, Z.; Reu, F.; et al. Targeting the MALAT1/PARP1/LIG3 complex induces DNA damage and apoptosis in multiple myeloma. *Leukemia* **2018**, *32*, 2250–2262. [CrossRef]
30. Zhang, Y.; Hou, Y.M.; Gao, F.; Xiao, J.W.; Li, C.C.; Tang, Y. lncRNA GAS5 regulates myocardial infarction by targeting the miR-525-5p/CALM2 axis. *J. Cell Biochem.* **2019**, *120*, 18678–18688. [CrossRef]
31. Mirza, A.H.; Kaur, S.; Pociot, F. Long non-coding RNAs as novel players in β cell function and type 1 diabetes. *Hum. Genomics* **2017**, *11*, 17. [CrossRef] [PubMed]
32. Postler, T.S.; Pantry, S.N.; Desrosiers, R.C.; Ghosh, S. Identification and characterization of a long non-coding RNA up-regulated during HIV-1 infection. *Virology* **2017**, *511*, 30–39. [CrossRef] [PubMed]
33. Somarowthu, S.; Legiewicz, M.; Chillón, I.; Marcia, M.; Liu, F.; Pyle, A.M. HOTAIR forms an intricate and modular secondary structure. *Mol. Cell* **2015**, *58*, 353–361. [CrossRef] [PubMed]
34. Chen, L.L. Linking long noncoding RNA localization and function. *Trends Biochem. Sci.* **2016**, *41*, 761–772. [CrossRef] [PubMed]
35. Long, Y.; Wang, X.; Youmans, D.T.; Cech, T.R. How do lncRNAs regulate transcription? *Adv. Sci.* **2017**, *3*, eaao2110. [CrossRef] [PubMed]
36. St Laurent, G.; Wahlestedt, C.; Kapranov, P. The landscape of long non-coding RNA classification. *Trends Genet.* **2015**, *31*, 239–251. [CrossRef] [PubMed]
37. Schulz, D.; Schwalb, B.; Kiesel, A.; Baejen, C.; Torkler, P.; Gagneur, J.; Soeding, J.; Cramer, P. Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* **2013**, *155*, 1075–1087. [CrossRef]
38. Salmena, L.; Poliseno, L.; Tay, Y.; Kats, L.; Pandolfi, P.P. A ceRNA hypothesis: The Rosetta stone of a hidden RNA language? *Cell* **2011**, *146*, 353–358. [CrossRef]
39. Tay, Y.; Rinn, J.; Pandolfi, P.P. The multilayered complexity of ceRNA crosstalk and competition. *Nature* **2014**, *505*, 344–352. [CrossRef]
40. Smillie, C.L.; Sirey, T.; Ponting, C.P. Complexities of post-transcriptional regulation and the modeling of ceRNA crosstalk. *Crit. Rev. Biochem. Mol. Biol.* **2018**, *53*, 231–245. [CrossRef]

41. Ulitsky, I. Interactions between short and long noncoding RNAs. *FEBS Lett.* **2018**, *592*, 2874–2883. [CrossRef] [PubMed]
42. Stroynowska-Czerwinska, A.; Fiszer, A.; Krzyżosiak, W.J. The panorama of miRNA-mediated mechanisms in mammalian cells. *Cell Mol. Life Sci.* **2014**, *71*, 2253–2270. [CrossRef] [PubMed]
43. Chan, J.J.; Tay, Y. Noncoding RNA: RNA regulatory networks in cancer. *Int. J. Mol. Sci.* **2018**, *19*, 1310. [CrossRef] [PubMed]
44. Qi, X.; Zhang, D.H.; Wu, N.; Xiao, J.H.; Wang, X.; Ma, W. ceRNA in cancer: Possible functions and clinical implications. *J. Med. Genet.* **2015**, *52*, 710–718. [CrossRef] [PubMed]
45. Feng, L.; Shi, L.; Lu, Y.F.; Wang, B.; Tang, T.; Fu, W.M.; He, W.; Li, G.; Zhang, J.F. Linc-ROR promotes osteogenic differentiation of mesenchymal stem cells by functioning as a competing endogenous RNA for miR-138 and miR-145. *Mol. Ther. Nucleic Acids* **2018**, *11*, 345–353. [CrossRef] [PubMed]
46. Gao, H.; Wang, T.; Zhang, P.; Shang, M.; Gao, Z.; Yang, F.; Liu, R. Linc-ROR regulates apoptosis in esophageal squamous cell carcinoma via modulation of p53 ubiquitination by targeting miR-204-5p/MDM2J. *Cell Physiol.* 2019. [CrossRef]
47. Song, Y.X.; Sun, J.X.; Zhao, J.H.; Yang, Y.C.; Shi, J.X.; Wu, Z.H.; Chen, X.W.; Gao, P.; Miao, Z.F.; Wang, Z.N. Non-coding RNAs participate in the regulatory network of CLDN4 via ceRNA mediated miRNA evasion. *Nat. Commun.* **2017**, *8*, 289. [CrossRef]
48. Wang, P.; Liu, Y.H.; Yao, Y.L.; Li, Z.; Li, Z.Q.; Ma, J.; Xue, Y.X. Long non-coding RNA CASC2 suppresses malignancy in human gliomas by miR-21. *Cell Signal.* **2015**, *27*, 275–282. [CrossRef]
49. Huang, G.; Wu, X.; Li, S.; Xu, X.; Zhu, H.; Chen, X. The long noncoding RNA CASC2 functions as a competing endogenous RNA by sponging miR-18 a in colorectal cancer. *Sci. Rep.* **2016**, *6*, 26524. [CrossRef]
50. Hansen, T.B.; Wiklund, E.D.; Bramsen, J.B.; Villadsen, S.B.; Statham, A.L.; Clark, S.J.; Kjems, J. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **2011**, *30*, 4414–4422. [CrossRef]
51. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388. [CrossRef] [PubMed]
52. Cesana, M.; Cacchiarelli, D.; Legnini, I.; Santini, T.; Sthandier, O.; Chinappi, M.; Tramontano, A.; Bozzoni, I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **2011**, *147*, 358–369. [CrossRef]
53. Wang, K.; Sun, T.; Li, N.; Wang, Y.; Wang, J.X.; Zhou, L.Y.; Long, B.; Liu, C.Y.; Liu, F.; Li, P.F. MDRL lncRNA regulates the processing of miR-484 primary transcript by targeting miR-361. *PLoS Genet.* **2014**, *10*, e1004467. [CrossRef]
54. Wang, J.; Liu, X.; Wu, H.; Ni, P.; Gu, Z.; Qiao, Y.; Chen, N.; Sun, F.; Fan, Q. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res.* **2010**, *38*, 5366–5383. [CrossRef] [PubMed]
55. Mangiavacchi, A.; Sorci, M.; Masciarelli, S.; Larivera, S.; Legnini, I.; Iosue, I.; Bozzoni, I.; Fazi, F.; Fatica, A. The miR-223 host non-coding transcript linc-223 induces IRF4 expression in acute myeloid leukemia by acting as a competing endogenous RNA. *Oncotarget* **2016**, *7*, 60155–60168. [CrossRef] [PubMed]
56. Qu, L.; Ding, J.; Chen, C.; Wu, Z.J.; Liu, B.; Gao, Y.; Chen, W.; Liu, F.; Sun, W.; Li, X.F. Exosome-transmitted lncARSR promotes sunitinib resistance in renal cancer by acting as a competing endogenous RNA. *Cancer Cell* **2016**, *29*, 653–668. [CrossRef] [PubMed]
57. Rani, N.; Nowakowski, T.J.; Zhou, H.; Godshalk, S.E.; Lisi, V.; Kriegstein, A.R.; Kosik, K.S. A primate lncRNA mediates notch signaling during neuronal development by sequestering miRNA. *Neuron* **2016**, *90*, 1174–1188. [CrossRef] [PubMed]
58. Cao, C.; Sun, J.; Zhang, D.; Guo, X.; Xie, L.; Li, X.; Wu, D.; Liu, L. The long intergenic noncoding RNA UFC1, a target of microRNA 34a, interacts with the mRNA stabilizing protein HuR to increase levels of beta-catenin in HCC cells. *Gastroenterology* **2015**, *148*, 415–426. [CrossRef]
59. Cai, Q.; Wang, Z.; Wang, S.; Weng, M.; Zhou, D.; Li, C.; Wang, J.; Chen, E.; Quan, Z. Long non-coding RNA LINC00152 promotes gallbladder cancer metastasis and epithelial–mesenchymal transition by regulating HIF-1α via miR-138. *Open Biol.* **2017**, *7*, 160247. [CrossRef]
60. Wang, X.; Li, M.; Wang, Z.; Han, S.; Tang, X.; Ge, X.; Zhou, L.; Zhou, C.; Yuan, Q.; Yang, M. Silencing of long noncoding RNA MALAT1 by miR-101 and miR- 217 inhibits proliferation, migration and invasion of esophageal squamous cell carcinoma cells. *J. Biol. Chem.* **2015**, *290*, 3925–3935. [CrossRef]

61. Leucci, E.; Patella, F.; Waage, J.; Holmstrøm, K.; Lindow, M.; Porse, B.; Kauppinen, S.; Lund, A.H. MicroRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. *Sci. Rep.* **2013**, *3*, 2535. [CrossRef] [PubMed]
62. Han, Y.; Liu, Y.; Zhang, H.; Wang, T.; Diao, R.; Jiang, Z.; Gui, Y.; Cai, Z. Hsa-miR-125b suppresses bladder cancer development by down- regulating oncogene SIRT7 and oncogenic long noncoding RNA MALAT1. *FEBS Lett.* **2013**, *87*, 3875–3882. [CrossRef] [PubMed]
63. Wang, T.; Yuan, J.; Feng, N.; Li, Y.; Lin, Z.; Jiang, Z.; Gui, Y. Hsa-miR-1 downregulates long non-coding RNA urothelial cancer associated 1 in bladder cancer. *Tumor Biol.* **2014**, *35*, 10075–10084. [CrossRef] [PubMed]
64. Faghihi, M.A.; Zhang, M.; Huang, J.; Modarresi, F.; van der Brug, M.P.; Nalls, M.A.; Cookson, M.R.; St-Laurent, G.; Wahlestedt, C. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* **2011**, *11*, R56. [CrossRef]
65. Imig, J.; Brunschweiger, A.; Brümmer, A.; Guennewig, B.; Mittal, N.; Kishore, S.; Tsikrika, P.; Gerber, A.P.; Zavolan, M.; Hall, J. miR-CLIP captureof a miRNAtargetomeuncovers a lincRNA H19-miR- 106a interaction. *Nat. Chem. Biol.* **2015**, *11*, 107–114. [CrossRef]
66. Kallen, A.N.; Zhou, X.B.; Xu, J.; Qiao, C.; Ma, J.; Yan, L.; Lu, L.; Liu, C.; Yi, J.S.; Zhang, H.; et al. The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol. Cell* **2013**, *52*, 101–112. [CrossRef]
67. Chiyomaru, T.; Yamamura, S.; Fukuhara, S.; Yoshino, H.; Kinoshita, T.; Majid, S.; Saini, S.; Chang, I.; Tanaka, Y.; Enokida, H.; et al. Genistein inhibits prostate cancer cell growth by targeting miR- 34a and oncogenic HOTAIR. *PLoS ONE* **2013**, *8*, e70372. [CrossRef]
68. Ma, M.Z.; Li, C.X.; Zhang, Y.; Weng, M.Z.; Zhang, M.D.; Qin, Y.Y.; Gong, W.; Quan, Z.W. Long non-coding RNA HOTAIR, a c-Myc activated driver of malignancy, negatively regulates miRNA-130a in gallbladder cancer. *Mol. Cancer* **2014**, *13*, 156. [CrossRef]
69. Yoon, J.H.; Abdelmohsen, K.; Kim, J.; Yang, X.; Martindale, J.L.; Tominaga-Yamanaka, K.; White, E.J.; Orjalo, A.V.; Rinn, J.L.; Kreft, S.G.; et al. Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat. Commun.* **2013**, *4*, 2939. [CrossRef]
70. Braconi, C.; Kogure, T.; Valeri, N.; Huang, N.; Nuovo, G.; Costinean, S.; Negrini, M.; Miotto, E.; Croce, C.M.; Patel, T. microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer. *Oncogene* **2011**, *30*, 4750–4756. [CrossRef]
71. Zhang, Z.; Zhu, Z.; Watabe, K.; Zhang, X.; Bai, C.; Xu, M.; Wu, F.; Mo, Y.Y. Negative regulation of lncRNA GAS5 by miR-21. *Cell Death Differ.* **2013**, *20*, 1558–1568. [CrossRef] [PubMed]
72. Gao, Y.; Meng, H.; Liu, S.; Hu, J.; Zhang, Y.; Jiao, T.; Liu, Y.; Ou, J.; Wang, D.; Yao, L.; et al. LncRNA-HOST2 regulates cell biological behaviorsin epithelial ovarian cancer through a mechanism involving microRNA let-7b. *Hum. Mol. Genet.* **2015**, *24*, 841–852. [CrossRef] [PubMed]
73. Prensner, J.R.; Chen, W.; Han, S.; Iyer, M.K.; Cao, Q.; Kothari, V.; Evans, J.R.; Knudsen, K.E.; Paulsen, M.T.; Ljungman, M.; et al. The long non-coding RNA PCAT-1 promotes prostate cancer cell proliferation through cMyc. *Neoplasia* **2014**, *16*, 900–908. [CrossRef] [PubMed]
74. Yoon, J.H.; Abdelmohsen, K.; Srikantan, S.; Yang, X.; Martindale, J.L.; De, S.; Huarte, M.; Zhan, M.; Becker, K.G.; Gorospe, M. LincRNA-p21 suppresses target mRNA. *Mol. Cell* **2012**, *47*, 648–655. [CrossRef]
75. Nguyen, T.C.; Cao, X.; Yu, P.; Xiao, S.; Lu, J.; Biase, F.H.; Sridhar, B.; Huang, N.; Zhang, K.; Zhong, S. Mapping RNA–RNA interactome and RNA structure in vivo by MARIO. *Nat. Commun.* **2016**, *7*, 12023. [CrossRef]
76. Cao, H.; Wahlestedt, C.; Karranov, P. Strategies to Annotate and Characterize Long Noncoding RNAs: Advantages and Pitfalls. *Trends in Genet.* **2018**, *34*, 704–721. [CrossRef]
77. He, J.H.; Han, Z.P.; Zou, M.X.; Wang, L.; Lv, Y.B.; Zhou, J.B.; Cao, M.R.; Li, Y.G. Analyzing the lncRNA, miRNA, and mRNA regulatory network in prostate cancer with bioinformatics software. *J. Comput. Biol.* **2018**, *25*, 146–157. [CrossRef]
78. Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M., Jr.; Jungkamp, A.C.; Munschauer, M.; et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **2010**, *141*, 129–141. [CrossRef]
79. Li, J.H.; Liu, S.; Zhou, H.; Qu, L.H.; Yang, J.H. starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **2014**, *42*, D92–D97. [CrossRef]

80. Tang, J.; Zhuo, H.; Zhang, X.; Jiang, R.; Ji, J.; Deng, L.; Qian, X.; Zhang, F.; Sun, B. A novel biomarker Linc00974 interacting with KRT19 promotes proliferation and metastasis in hepatocellular carcinoma. *Cell Death Dis.* **2014**, *5*, e1549. [CrossRef]
81. Kashi, K.; Henderson, L.; Bonetti, A.; Carninci, P. Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. *Biochim. Biophys. Acta* **2016**, *1859*, 3–15. [CrossRef]
82. Kent, O.A.; Mendell, J.T. A small piece in the cancer puzzle: microRNAs as tumor suppressors. *Oncogene* **2006**, *25*, 6188–6196. [CrossRef]
83. Lu, J.; Andrew, G.; Clark, A.G. Impact of microRNA regulation on variation in human gene expression. *Genome Res.* **2012**, *22*, 1243–1254. [CrossRef]
84. Zhou, K.; Liu, M.; Cao, Y. New insight into microRNA: Functions in cancer: Oncogene–microRNA–tumor suppressor gene network. *Front. Mol. Biosci.* **2017**, *4*, 46. [CrossRef]
85. Wang, K.C.; Chang, H.Y. Molecular mechanism of long non-coding RNAs. *Mol. Cell* **2011**, *43*, 904–914. [CrossRef]
86. Anastasiadou, E.; Jacob, L.S.; Slack, F.J. Non-coding RNA networks in cancer. *Nat. Rev. Cancer* **2018**, *18*, 5–18. [CrossRef]
87. Cai, Y.; Wan, J. Competing endogenous RNA regulations in neurodegenerative disorders: Current challenges and emerging insights. *Front. Mol. Neurosci.* **2018**, *11*, 370. [CrossRef]
88. Tao, F.; Tian, X.; Ruan, S.; Shen, M.; Zhang, Z. miR-211 sponges lncRNA MALAT1 to suppress tumor growth and progression through inhibiting PHF19 in ovarian carcinoma. *FASEB J.* **2018**. [CrossRef]
89. Yang, B.; Gao, G.; Wang, Z.; Sun, D.; Wei, X.; Ma, Y.; Ding, Y. Long non-coding RNA HOTTIP promotes prostate cancer cells proliferation and migration by sponging *miR-216a-5p*. *Biosci. Rep.* **2018**, *38*, BSR20180566. [CrossRef]
90. Denzler, R.; Agarwal, V.; Stefano, J.; Bartel, D.P.; Stoffel, M. Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol. Cell* **2014**, *54*, 766–776. [CrossRef]
91. Thomson, D.W.; Marcel, E.; Dinger, M.E. Endogenous microRNA sponges: Evidence and controversy. *Nat. Rev. Genet.* **2016**, *17*, 272–283. [CrossRef] [PubMed]
92. Romero-Barrios, N.; Legascue, M.F.; Benhamed, M.; Federico Ariel, F.; Crespi, M. Splicing regulation by long non-coding RNAs. *Nucleic Acids Res.* **2018**, *46*, 2169–2184. [CrossRef] [PubMed]
93. Karapetyan, A.E.; Coen Buiting, C.; Kuiper, R.A.; Marcel, W.; Coolen, M.W. Regulatory roles for long ncRNA and mRNA. *Cancers* **2013**, *5*, 462–490. [CrossRef] [PubMed]
94. Tripathi, V.; Ellis, J.D.; Shen, Z.; Song, D.Y.; Pan, Q.; Watt, A.T.; Freier, S.M.; Bennett, C.F.; Sharma, A.; Bubulya, P.A.; et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **2010**, *39*, 925–938. [CrossRef]
95. Salameh, A.; Lee, A.K.; Cardo-Vila, M.; Nunes, D.N.; Efstathiou, E.; Staquicini, F.I.; Dobroff, A.S.; Marchio, S.; Navone, N.M.; Hosoya, H.; et al. PRUNE2 is a human prostate cancer suppressor regulated by the intronic long noncoding RNA PCA3. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 8403–8408. [CrossRef]
96. Cao, L.; Zhang, P.; Li, J.; Wu, M. LAST, a c-Myc-inducible long noncodingRNA, cooperates with CNBP to promote CCND1 mRNA stability in human cells. *eLife* **2017**, *6*, e30433. [CrossRef]
97. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; Laurent, G.S., III; Kenny, P.J.; Wahlestedt, C. Expression of a noncoding RNA is elevated in alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med.* **2008**, *14*, 723–730. [CrossRef]
98. Yoon, J.-H.; Abdelmohsen, K.; Gorospe, M. Post-transcriptional gene regulation by long noncoding RNA. *J. Mol. Biol.* **2013**, *19*, 3723–3730. [CrossRef]
99. Tafer, H.; Hofacker, I.L. RNAplex: A fast tool for RNA-RNA interaction search. *Bioinformatics* **2008**, *24*, 2657–2663. [CrossRef]
100. Li, J.; Ma, W.; Zeng, P.; Wang, J.; Geng, B.; Yang, J.; Cui, Q. LncTar: A tool for predicting the RNA targets of long noncoding RNAs. *Brief. Bioinform.* **2015**. [CrossRef]
101. Gott, M.J.; Emerson, R.B. Functions and mechanism of RNA editing. *Ann. Rev. Genet.* **2000**, *34*, 499–531. [CrossRef] [PubMed]
102. Szcześniak, M.W.; Makałowska, I. lncRNA-RNA Interactions across the human transcriptome. *PLoS ONE* **2016**, *11*, e0150353. [CrossRef] [PubMed]

103. Quinodoz, S.; Guttman, M. Longnoncoding RNAs: An emerging link between gene regulation and nuclear organization. *Cell Chem. Biol.* **2014**, *24*, 652–663.
104. Li, Y.; Syed, J.; Sugiyama, H. RNA-DNA triplex formation by longnoncoding RNAs. *Cell Chem. Biol.* **2016**, *23*, 1325–1333. [CrossRef]
105. Shaw, N.N.; Arya, D.P. Recognition of the unique structure of DNA: RNA hybrids. *Biochimie* **2008**, *90*, 1026–1039. [CrossRef] [PubMed]
106. Buske, F.A.; Mattick, J.S.; Bailey, T.L. Potential in vivo roles of nucleic acid triple-helices. *RNA Biol.* **2011**, *8*, 427–439. [CrossRef]
107. Bacolla, A.; Wang, G.; Vasquez, K.M. New Perspectives on DNA and RNA Triplexes As Effectors of Biological Activity. *PLoS Genet.* **2015**, *11*, e1005696. [CrossRef]
108. Zhou, B.; Li, X.; Luo, D.; Lim, D.H.; Zhou, Y.; Fu, X.D. GRID-seq for comprehensive analysis of global RNA-chromatin interactions. *Nat. Protoc.* **2019**, *14*, 2036–2068. [CrossRef]
109. Buske, F.A.; Bauer, D.C.; Mattick, J.S.; Bailey, T.L. Triplex-Inspector: An analysis tool for triplex-mediated targeting of genomic loci. *Bioinformatics* **2013**, *29*, 1895–1897. [CrossRef]
110. He, S.; Zhang, H.; Liu, H.; Zhu, H. LongTarget: A tool to predict lncRNA DNA-binding motifs and binding sites via Hoogsteen base-pairing analysis. *Bioinformatics* **2015**, *31*, 178–186. [CrossRef]
111. Chu, C.; Qu, K.; Zhong, F.L.; Artandi, S.E.; Chang, H.Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **2011**, *44*, 667–678. [CrossRef] [PubMed]
112. Vance, K.W.; Ponting, C.P. Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.* **2014**, *30*, 348–355. [CrossRef]
113. Nagano, T.; Mitchell, J.A.; Sanz, L.A.; Pauler, F.M.; Ferguson-Smith, A.C.; Feil, R.; Fraser, P. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **2008**, *322*, 1717–1720. [CrossRef]
114. Grote, P.; Herrmann, B.G. The long non-coding RNA Fendrr links epigenetic control mechanisms to gene regulatory networks in mammalian embryogenesis. *RNA Biol.* **2013**, *10*, 1579–1585. [CrossRef] [PubMed]
115. Xu, T.; Huang, M.; Xia, R.; Liu, X.; Sun, M.; Yin, L.; Chen, W.M.; Han, L.; Zhang, E.B.; Kong, R.; et al. Decreased expression of the long non-coding of FENDRR is associated with poor prognosis in gastric cancer and FENDRR regulates gastric cancer cell metastasis by affecting fibronectin1 expression. *J. Hematol. Oncol.* **2014**, *7*, 63. [CrossRef] [PubMed]
116. Lin, N.; Chang, K.Y.; Li, Z.; Gates, K.; Rana, Z.A.; Dang, J.; Zhang, D.; Han, T.; Yang, C.S.; Cunningham, T.J.; et al. An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol. Cell* **2014**, *53*, 1005–1019. [CrossRef] [PubMed]
117. Mondal, T.; Subhash, S.; Vaid, R.; Enroth, S.; Uday, S.; Reinius, B.; Mitra, S.; Mohammed, A.; James, A.R.; Hoberg, E.; et al. MEG3 long noncoding RNA regulates the TGF-b pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* **2015**, *6*, 7743. [CrossRef] [PubMed]
118. O'Leary, V.B.; Ovsepian, S.V.; Carrascosa, L.G.; Buske, F.A.; Radulovic, V.; Niyazi, M.; Moertl, S.; Trau, M.; Atkinson, M.J.; Anastasov, N. PARTICLE, a triplex-forming long ncRNA, regulates locus-specific methylation in response to low-dose irradiation. *Cell Rep.* **2015**, *11*, 474–485. [CrossRef]
119. West, J.A.; Davis, C.P.; Sunwoo, H.; Simon, M.D.; Sadreyev, R.I.; Wang, P.I.; Tolstorukov, M.Y.; Kingston, R.E. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* **2014**, *55*, 791–802. [CrossRef]
120. Bond, C.S.; Fox, A.H. Paraspeckles: Nuclear bodies built on long noncoding RNA. *J. Cell. Biol.* **2009**, *186*, 637–644. [CrossRef]
121. Martianov, I.; Ramadass, A.; Barros, A.; Chow, N.; Akoulitchev, A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **2007**, *445*, 666–670. [CrossRef] [PubMed]
122. Blume, S.W.; Meng, Z.; Shrestha, K.; Snyder, R.C.; Emanuel, P.D. The 5'-untranslated RNA of the human dhfr minor transcript alters transcription pre-initiation complex assembly at the major (core) promoter. *J. Cell. Biochem.* **2003**, *88*, 165–180. [CrossRef] [PubMed]
123. Deng, Z.; Norseen, J.; Wiedmer, A.; Riethman, H.; Lieberman, P.M. TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol. Cell.* **2009**, *35*, 403–413. [CrossRef] [PubMed]
124. Xu, Y.; Kimura, T.; Komiyama, M. Human telomere RNA and DNA form an intermolecular G-quadruplex. *Nucleic Acids Symp. Ser.* **2008**, *52*, 169–170. [CrossRef]

125. Toubiana, S.; Selig, S. DNA: RNA hybrids at telomeres—when it is better to be out of the R loop. *FEBS J.* **2018**, *285*, 2552–2566. [CrossRef]
126. Xu, Y.; Suzuki, Y.; Ito, K.; Komiyama, M. Telomeric repeat-containing RNA structure in living cells. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14579–14584. [CrossRef]
127. Postepska-Igielska, A.; Giwojna, A.; Gasri-Plotnitsky, L.; Schmitt, N.; Dold, A.; Ginsberg, D.; Grummt, I. LncRNA Khps1 Regulates Expression of the proto-oncogene SPHK1 via triplex-mediated changes in chromatin structure. *Mol. Cell* **2015**, *60*, 626–636. [CrossRef]
128. Yang, Y.; Wen, L.; Zhu, H. Unveiling the hidden function of long non-coding RNA by identifying its major partner-protein. *Cell Biosci.* **2015**, *5*, 59. [CrossRef]
129. Sun, X.; Haider Ali, M.S.S.; Moran, M. The role of interactions of long non-coding RNAs and heterogeneous nuclear ribonucleoproteins in regulating cellular functions. *Biochem. J.* **2017**, *474*, 2925–2935. [CrossRef]
130. Wang, K.C.; Chang, H.Y. Molecular mechanisms of longnoncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914. [CrossRef]
131. Barra, J.; Leucci, E. Probing long non-coding RNA-protein interactions. *Front. Mol. Biosci.* **2017**, *4*, 45. [CrossRef]
132. Hellman, L.M.; Fried, M.G. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* **2007**, *2*, 1849–1861. [CrossRef] [PubMed]
133. Lin, R.J. RNA protein interaction protocols. *Methods Mol. Biol.* **2008**, *488*, 5–7.
134. Konig, J.; Zarnack, K.; Luscombe, N.M.; Ule, J. Protein-RNA interactions: New genomic technologies and perspectives. *Nat. Rev. Genet.* **2011**, *13*, 77–83. [CrossRef]
135. Barnes, C.; Kanhere, A. Identification of RNA-protein interactions through *in vitro* RNA pull-down assays. *Methods Mol. Biol.* **2016**, *1480*, 99–113. [PubMed]
136. Zheng, X.; Cho, S.; Moon, H.; Loh, T.J.; Jang, H.N.; Shen, H. Detecting RNA-protein interaction using end-labeled biotinylated RNA oligonucleotides and immunoblotting. *Methods Mol. Biol.* **2016**, *1421*, 35–44. [PubMed]
137. Jalali, S.; Gandhi, S.; Scaria, V.S. Distinct and modular organization of protein interacting sites in long non-coding RNAs. *Front. Mol. Biosci.* **2018**, *5*, 27. [CrossRef]
138. Yoon, J.H.; Gorospe, M. Cross-Linking Immunoprecipitation and qPCR (CLIP-qPCR) Analysis to map interactions between longnoncoding RNAs and RNA-binding proteins. *Methods Mol. Biol.* **2016**, *1402*, 11–17.
139. Garzia, A.; Meyer, C.; Morozov, P.; Sajek, M.; Tuschl, T. Optimization of PAR-CLIP for transcriptome-wide identification of binding sites of RNA-binding proteins. *Methods* **2017**, *118–119*, 24–40. [CrossRef]
140. Kiefer, L.; Schofield, J.A.; Simon, M.D. Expanding the nucleoside recoding toolkit: Revealing RNA population dynamics with 6-thioguanosine. *J. Am. Chem. Soc.* **2018**, *140*, 14567–14570. [CrossRef]
141. Quek, X.C.; Thomson, D.W.; Maag, J.L.; Bartonicek, N.; Signal, B.; Clark, M.B.; Gloss, B.S.; Dinger, M.E. lncRNAdb v2.0: Expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res.* **2015**, *43*, D168–D173. [CrossRef] [PubMed]
142. Volders, P.J.; Anckaert, J.; Verheggen, K.; Nuytens, J.; Martens, L.; Mestdagh, P.; Vandesompele, J. LNCipedia 5: Towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.* **2019**, *47*, D135–D139. [CrossRef] [PubMed]
143. Frankish, A.; Diekhans, M.; Ferreira, A.M.; Johnson, R.; Jungreis, I.; Loveland, J.; Mudge, J.M.; Sisu, C.; Wright, J.; Armstrong, J.; et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **2019**, *47*, D766–D773. [CrossRef] [PubMed]
144. Ribeiro, D.M.; Zanzoni, A.; Cipriano, A.; Delli Ponti, R.; Spinelli, L.; Ballarino, M.; Bozzoni, I.; Tartaglia, G.G.; Brun, C. Protein complex scaffolding predicted as a prevalent function of long non-codingRNAs. *Nucleic Acids Res.* **2018**, *46*, 917–928. [CrossRef]
145. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076. [CrossRef]
146. Schorderet, P.; Duboule, D. Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genet.* **2011**, *7*, e1002071. [CrossRef]
147. Kogo, R.; Shimamura, T.; Mimori, K.; Kawahara, K.; Imoto, S.; Sudo, T.; Tanaka, F.; Shibata, K.; Suzuki, A.; Komune, S.; et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.* **2011**, *71*, 6320–6326. [CrossRef]

148. Wang, P.; Xue, Y.; Han, Y.; Lin, L.; Wu, C.; Xu, S.; Jiang, Z.; Xu, J.; Liu, Q.; Cao, X. The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. *Science* **2014**, *344*, 310–313. [CrossRef]
149. Pandey, R.R.; Mondal, T.; Mohammad, F.; Enroth, S.; Redrup, L.; Komorowski, J.; Nagano, T.; Mancini-Dinardo, D.; Kanduri, C. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **2008**, *32*, 232–246. [CrossRef]
150. Kanduri, C. Kcnq1ot1: A chromatin regulatory RNA. *Semin. Cell. Dev. Biol.* **2011**, *22*, 343–350. [CrossRef]
151. Schoeftner, S.; Sengupta, A.K.; Kubicek, S.; Mechtler, K.; Spahn, L.; Koseki, H.; Jenuwein, T.; Wutz, A. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* **2006**, *25*, 3110–3122. [CrossRef] [PubMed]
152. Moindrot, B.; Brockdorff, N. RNA binding proteins implicated in Xist-mediated chromosome silencing. *Semin. Cell Dev. Biol.* **2016**, *56*, 58–70. [CrossRef] [PubMed]
153. Seidl, C.I.; Stricker, S.H.; Barlow, D.P. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J.* **2006**, *25*, 3565–3575. [CrossRef] [PubMed]
154. Santoro, F.; Mayer, D.; Klement, R.M.; Warczok, K.E.; Stukalov, A.; Barlow, D.P.; Pauler, F.M. Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Development* **2013**, *140*, 1184–1195. [CrossRef] [PubMed]
155. Yap, K.L.; Li, S.; Muñoz-Cabello, A.M.; Raguz, S.; Zeng, L.; Mujtaba, S.; Gil, J.; Walsh, M.J.; Zhou, M.M. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **2010**, *38*, 662–674. [CrossRef] [PubMed]
156. Lian, Y.; Cai, Z.; Gong, H.; Xue, S.; Wu, D.; Wang, K. HOTTIP: A critical oncogenic long non-coding RNA in human cancers. *Mol. Biosyst.* **2016**, *12*, 3247–3253. [CrossRef]
157. Runge, S.; Nielsen, F.C.; Nielsen, J.; Lykke-Andersen, J.; Wewer, U.M.; Christiansen, J. H19 RNA binds four molecules of insulin-like growth factor II mRNA-binding protein. *J. Biol. Chem.* **2000**, *275*, 29562–29569. [CrossRef]
158. Jiang, W.; Liu, Y.; Liu, R.; Zhang, K.; Zhang, Y. The lncRNA DEANR1 facilitates human endoderm differentiation by activating FOXA2 expression. *Cell Rep.* **2015**, *11*, 137–148. [CrossRef]
159. Dimitrova, N.; Zamudio, J.R.; Jong, R.M.; Soukup, D.; Resnick, R.; Sarma, K.; Ward, A.J.; Raj, A.; Lee, J.T.; Sharp, P.A.; et al. LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. *Mol. Cell* **2014**, *54*, 777–790. [CrossRef]
160. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419. [CrossRef]
161. Puvvula, P.K.; Desetty, R.D.; Pineau, P.; Marchio, A.; Moon, A.; Dejean, A.; Bischof, O. Long noncoding RNA PANDA and scaffold-attachment-factor SAFA control senescence entry and exit. *Nat. Commun.* **2014**, *5*, 5323. [CrossRef] [PubMed]
162. Hung, T.; Wang, Y.; Lin, M.F.; Koegel, A.K.; Kotake, Y.; Grant, G.D.; Horlings, H.M.; Shah, N.; Umbricht, C.; Wang, P.; et al. Extensive and coordinatedtranscription of noncoding RNAs within cell-cycle promoters. *Nat. Genet.* **2011**, *43*, 621–629. [CrossRef] [PubMed]
163. Kino, T.; Hurt, D.E.; Ichijo, T.; Nader, N.; Chrousos, G.P. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal.* **2010**, *3*, ra8. [CrossRef] [PubMed]
164. Yang, Z.; Zhu, B.Q.; Luo, B.K.; Zhou, Q. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **2001**, *414*, 317–322. [CrossRef]
165. Barboric, M.; Lenasi, T.; Chen, H.; Johansen, E.B.; Guo, S.; Peterlin, B.M. 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7798–7803. [CrossRef]
166. Tsuiji, H.; Yoshimoto, R.; Hasegawa, Y.; Furuno, M.; Yoshida, M.; Nakagawa, S. Competition between a noncoding exon and introns: Gomafu contains tandem UACUAAC repeats and associates with splicing factor-1. *Genes Cells* **2011**, *16*, 479–490. [CrossRef]
167. Yoshimoto, R.; Mayeda, A.; Yoshida, M.; Nakagawa, S. MALAT1 long non-coding RNA in cancer. *Biochim. Biophys. Acta* **2016**, *1859*, 192–199. [CrossRef]
168. Ramos, A.D.; Andersen, R.E.; Liu, S.J.; Lim, D.A. The long noncoding RNA Pnky regulates neuronal differentiation of embryonic and postnatal neural stem cells. *Cell Stem Cell 2015 16*, 439–447. [CrossRef]

169. Antoniou, D.; Stergiopoulos, A.; Politis, P.K. Recent advances in the involvement of long non-coding RNAs in neural stem cell biology and brain pathophysiology. *Front. Physiol.* **2014**, *5*, 155. [CrossRef]
170. Cho, K.O.; Hsieh, J. The lncRNA Pnky in the brain. *Cell Stem Cell* **2015**, *16*, 344–345. [CrossRef]
171. Wang, H.; Iacoangeli, A.; Lin, D.; Williams, K.; Denman, R.B.; Hellen, C.U.; Tiedge, H. Dendritic BC1 RNA in translational control mechanisms. *J. Cell. Biol.* **2005**, *171*, 811–821. [CrossRef] [PubMed]
172. Li, Z.; Chao, T.C.; Chang, K.Y.; Lin, N.; Patil, V.S.; Shimizu, C.; Head, S.R.; Burns, J.C.; Rana, T.M. Thelong noncoding RNATHRIL regulates TNFα expression through its interaction with hnRNPL. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1002–1007. [CrossRef]
173. Davidovich, C.; Cech, T. The recruitment of chromatin modifiers by long noncoding RNAs: Lessons from PRC2. *RNA* **2015**, *1*, 2007–2022. [CrossRef] [PubMed]
174. Angrand, P.O.; Vennin, C.; Le Bourhis, X.; Adriaenssens, E. The role of long non-coding RNAs in genome formatting and expression. *Front. Genet.* **2015**, *6*, 165. [CrossRef] [PubMed]
175. Chittock, E.C.; Latwiel, S.; Miller, T.C.; Müller, C.W. Molecular architecture of polycomb repressive complexes. *Biochem. Soc. Trans.* **2017**, *45*, 193–205. [CrossRef] [PubMed]
176. Li, J.; Chen, Y.; Xu, X.; Jones, J.; Tiwari, M.; Ling, J.; Ling, J.; Wang, Y.; Harismedy, O.; Sen, G.L. HNRNPK maintains epidermal progenitor function through transcription of proliferation genes and degrading differentiation promoting mRNAs. *Nat. Commun.* **2019**, *10*, 1–14. [CrossRef]
177. McHugh, C.A.; Russell, P.; Guttman, M. Methods for comprehensive experimental identification of RNA-protein interactions. *Genome Biol.* **2014**, *15*, 203. [CrossRef]
178. Liu, A.; Liu, S. Non-coding RNAs in cancer biology. In *Advances in Experimental Medicine and Biology*; Son, E., Ed.; Springer: Berlin/Heidelberg, Germany, 2016; Volume 927, pp. 367–389.
179. Zhu, S.; Wang, J.; He, Y.; Meng, N.; Yan, G.R. Peptides/Proteins encoded by non-coding RNA: A novel resource bank for drug targets and biomarkers. *Front. Pharmacol.* **2018**, *9*, 1295. [CrossRef]
180. Chugunova, A.; Navalayeu, T.; Dontsova, O.; Sergiev, P. Mining for small translated ORFs. *J. Proteome Res.* **2018**, *17*, 1–11. [CrossRef]
181. Matsumoto, A.; Nakayama, K.I. Hidden peptides encoded by putative noncoding RNAs. *Cell Struct. Funct.* **2018**, *43*, 75–83. [CrossRef] [PubMed]
182. Ruiz-Orera, J.; Verdaguer-Grau, P.; Villanueva-Canas, J.L.; Messegue, X.; Alba, M.M. Functional and non-functional classes of peptides produced by long non-coding RNAs. *biorXiv* **2016**. [CrossRef]
183. Choi, S.W.; Kim, H.W.; Nam, J.W. The small peptide world in long noncoding RNAs. *Brief. Bioinform.* **2018**. [CrossRef] [PubMed]
184. Chong, S. Overview of cell-free protein synthesis: Historic landmarks, commercial systems, and expanding applications. *Curr. Protoc. Mol. Biol.* **2014**, *108*, 1–11. [CrossRef] [PubMed]
185. Guthals, A.; Bandeira, N. Peptide identification by tandem mass spectrometry with alternate fragmentation modes. *Mol. Cell. Proteomics* **2012**, *11*, 550–557. [CrossRef] [PubMed]
186. Ingolia, N.T. Ribosome profiling: New views of translation, from single codons to genome scale. *Nat. Rev. Genet.* **2014**, *15*, 205–213. [CrossRef]
187. Nam, J.W.; Choi, S.W.; You, B.H. Incredible RNA: Dual functions of coding and noncoding. *Mol. Cells* **2016**, *39*, 367–374.
188. Anderson, D.M.; Anderson, K.M.; Chang, C.L.; Makarewich, C.A.; Nelson, B.R.; McAnally, J.R.; Kasaragod, P.; Shelton, J.M.; Liou, J.; Bassel-Duby, R.; et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* **2015**, *160*, 595–606. [CrossRef]
189. Matsumoto, A.; Pasut, A.; Matsumoto, M.; Yamashita, R.; Fung, J.; Monteleone, E.; Saghatelian, A.; Nakayama, K.I.; Clohessy, J.G.; Pandolfi, P.P. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* **2017**, *541*, 228–232. [CrossRef]
190. Nelson, B.R.; Makarewich, C.A.; Anderson, D.M.; Winders, B.R.; Troupes, C.D.; Wu, F.; Reese, A.L.; McAnally, J.R.; Chen, X.; Kavalali, E.T.; et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* **2016**, *351*, 271–275. [CrossRef]
191. Huang, J.Z.; Chen, M.; Gao, X.C.; Zhu, S.; Huang, H.; Hu, M.; Zhu, H.; Yan, G.R. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. *Mol. Cell* **2017**, *68*, 171–184. [CrossRef]
192. Gestelanf, R.F.; Cech, T.R.; Atkins, J.F. (Eds.) *The RNA World*, 3rd ed.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2005.

193. Warner, K.D.; Hajdin, C.E.; Weeks, K.M. Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discov.* **2018**, *17*, 547–558. [CrossRef] [PubMed]
194. Brown, J.A.; Bulkley, D.; Wang, J.; Valenstein, M.L.; Yario, T.A.; Steitz, T.A.; Steitz, J.A. Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat. Struct. Mol. Biol.* **2014**, *21*, 633–640. [CrossRef] [PubMed]
195. Brown, J.A.; Kinzig, C.G.; DeGregorio, S.J.; Steitz, J.A. Hoogsteen-position pyrimidines promote the stability and function of the MALAT1 RNA triple helix. *RNA* **2016**, *22*, 743–749. [CrossRef]
196. Brown, J.A.; Kinzig, C.G.; DeGregorio, S.J.; Steitz, J.A. Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. *Proc. Nat. Acad. Sci. USA* **2016**, *113*, 14013–14018. [CrossRef]
197. Donlic, A.; Hargrove, A.E. Targeting RNA in mammalian systems with small molecules. *Wiley Interdiscip. Rev. RNA* **2018**, *9*, e1477. [CrossRef] [PubMed]
198. Liu, F.; Somarowthu, S.; Pyle, A.M. Visualizing the secondary and tertiary architectural domains of lncRNA RepA. *Nat. Chem. Biol.* **2017**, *13*, 282–289. [CrossRef] [PubMed]
199. Smola, M.J.; Christy, T.W.; Inoue, K.; Nicholson, C.O.; Friedersdorf, M.; Keene, J.D.; Lee, D.M.; Calabrese, J.M.; Weeks, K.M. SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 10322–10327. [CrossRef]
200. Owens, M.C.; Clark, S.C.; Yankey, A.; Somarowthu, S. Identifying structural domains and conserved regions in the long non-coding RNA lncTCF7. *Int. J. Mol. Sci.* **2019**, *20*, 4770. [CrossRef]
201. Fatemi, R.P.; Salah-Uddin, S.; Modarresi, F.; Khoury, N.; Wahlestedt, C.; Faghihi, M.A. Screening for small-molecule modulators of long noncoding RNA-protein Interactions using AlphaScreen. *J. Biomol. Screen.* **2015**, *20*, 1132–1141. [CrossRef]
202. Zhang, Y.; Zeng, D.; Cao, J.; Wang, M.; Shu, B.; Kuang, G.; Qu, T.M.; Tan, G.H.; Gu, L.Q.; Huang, Z.S.; et al. Interaction of quindoline derivative withtelomeric repeat-containing RNA induces telomeric DNA-damage response in cancer cells through inhibition of telomeric repeat factor 2. *Biochim. Biophys. Acta* **2017**, *1861*, 3246–3256. [CrossRef] [PubMed]
203. Shi, Y.; Parag, S.; Patel, R.; Lui, A.; Murr, M.; Cai, J.; Patel, N.A. Stabilization of lncRNA GAS5 by a small molecule and its implications in diabetic. *Cell Chem. Biol.* **2019**, *26*, 319–330. [CrossRef] [PubMed]
204. Bowden, R.; Davies, R.W.; Heger, A.; Pugnamenta, A.T.; de Cesare, M.; Oikonen, L.E.; Parkes, D.; Freeman, C.; Dhalla, F.; Patel, S.Y. Sequencing of human genomes with nanopore technology. *Nat. Commun.* **2019**, *10*, 1869. [CrossRef] [PubMed]
205. Makałowski, W.; Shabardina, V.J. Bioinformatics of nanopore sequencing. *Hum. Genet.* **2020**, *65*, 61–67. [CrossRef]
206. Gupta, D.; Bhattacharjee, O.; Mandal, D.; Sen, M.K.; Dey, D.; Dasgupta, A.; Kazi, T.A.; Gupta, R.; Sinharoy, S.; Acharya, K. CRISPR-Cas9 system: A new-fangled dawn in gene editing. *Life Sci.* **2019**, *23*, 116636. [CrossRef]
207. Ali, Z.; Mahas, A.; Mahfouz, M. CRISPR/Cas13 as a tool for RNA interference. *Trends Plant. Sci.* **2018**, *23*, 374–378. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Review

Long Non-Coding RNA Epigenetics

Marek Kazimierczyk and Jan Wrzesinski *

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland; mkazimierczyk@man.poznan.pl
* Correspondence: wrzesinj@ibch.poznan.pl

Abstract: Long noncoding RNAs exceeding a length of 200 nucleotides play an important role in ensuring cell functions and proper organism development by interacting with cellular compounds such as miRNA, mRNA, DNA and proteins. However, there is an additional level of lncRNA regulation, called lncRNA epigenetics, in gene expression control. In this review, we describe the most common modified nucleosides found in lncRNA, 6-methyladenosine, 5-methylcytidine, pseudouridine and inosine. The biosynthetic pathways of these nucleosides modified by the writer, eraser and reader enzymes are important to understanding these processes. The characteristics of the individual methylases, pseudouridine synthases and adenine–inosine editing enzymes and the methods of lncRNA epigenetics for the detection of modified nucleosides, as well as the advantages and disadvantages of these methods, are discussed in detail. The final sections are devoted to the role of modifications in the most abundant lncRNAs and their functions in pathogenic processes.

Keywords: lncRNA epigenetics; RNA modyfiing enzymes; detection modified nucleotides

Citation: Kazimierczyk, M.; Wrzesinski, J. Long Non-Coding RNA Epigenetics. *Int. J. Mol. Sci.* **2021**, *22*, 6166. https://doi.org/10.3390/ijms22116166

Academic Editor: Erik A. C. Wiemer

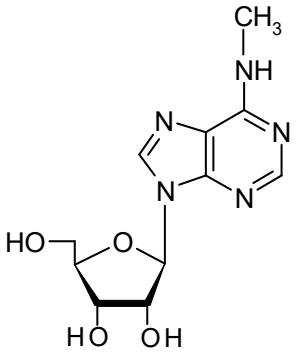
Received: 18 May 2021
Accepted: 1 June 2021
Published: 7 June 2021

1. General Remarks

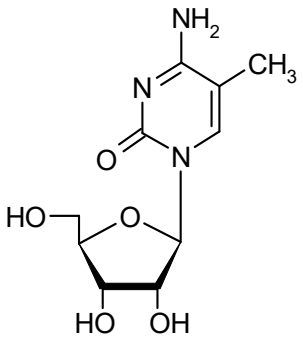
Although almost 80% of the human genome has been transcribed, only 2% of it (mRNA) codes proteins [1,2]. All of the remaining RNAs belong to a vast group of noncoding RNAs (ncRNAs). Based on their functions, ncRNAs can be classified as housekeeping RNAs or regulatory RNAs [3,4]. Housekeeping ncRNAs, including transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and ribosomal RNAs (rRNAs), are commonly expressed constitutively. Regulatory RNAs are a type of ncRNA with a strong regulatory impact on the expression of protein-coding genes. Based on their size, regulatory RNAs can be divided into two groups: small noncoding RNAs (sncRNAs), which are miRNAs and piRNAs, and long noncoding RNAs (lncRNAs). LncRNAs, whose sizes range from several hundred to several thousand nucleotides, are structurally similar to mRNAs [5]. Just like mRNAs, lncRNAs are transcribed by RNA polymerase II, and are likewise capped, often spliced and polyadenylated [6,7]. Despite their large size, lncRNAs do not code protein. A number of lncRNAs contain short ORFs, with fewer than 100 amino acids, which code small proteins or micropeptides, [8–10]. Many lncRNAs are implicated in gene-regulatory roles such as chromosome dosage compensation, imprinting, transcription, translation, splicing, cell cycle control, epigenetic regulation, nuclear and cytoplasmic trafficking, and cell differentiation [11–13]. Recent studies have shown that lncRNAs, like mRNAs, contain modified nucleotides, which regulate cellular activity [14,15].

It has been known for many years that modified deoxynucleotides (m^5C, m^6A) are present in DNA [16,17]. DNA modifications, which do not alter the DNA's sequence, but have an impact on gene activity, are known as epigenetic modifications [18,19]. Similarly, the analysis of RNA post-transcriptional modifications revealed that they occur in every living organism. Such changes are called RNA epigenetics or epitranscriptomics. Over 170 RNA chemical modifications have been found in different organisms to date and over 60 RNA modifications have been identified in eukaryotes Modomics database: http://modomics.genesilico.pl/, accessed on 2 November 2017 [20]. There are

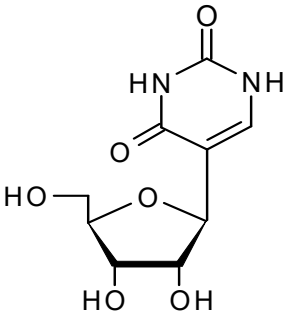
different RNA molecules containing ribonucleotide modifications, such as tRNA, rRNA, mRNA and noncoding RNA. Most of the modifications are found in tRNAs, ranging from adenosine and cytidine methylation to hypermodified nucleosides such as 3-[(2R,3R,4S,5R)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-4,6-dimethylimidazo[1,2-a]purin-9-one (wyosine) [21,22], 2-thio-5-carboxymethyluracyl (s^2mcm^5U) [23], N^6-isopentenyl adenine (i^6A) [24] or N6-threonylcarbamoyladenosine (t^6A) [25], and are not present in other RNA molecules. However, the most abundant modified nucleotides involving lncRNA are methyl nucleotide derivatives, including 5-methyl–cytidine (m^5C) and 6-methyl–adenosine (m^6A) (Figure 1) [26,27]. Modified nucleotides, such as pseudouridine and inosine, constitute a separate group [28,29].



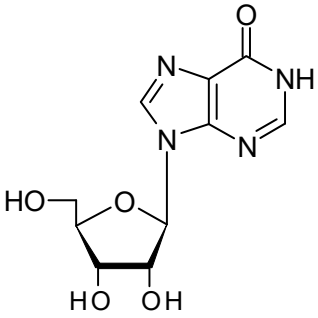
N^6-methyladenosine (m^6A)



5-methylcytidine (m^5C)



Pseudouridine (Ψ)



Inosine (I)

Figure 1. The structure of modified nucleotides found in eukaryotic lncRNA molecules [26–29].

Abundant noncoding RNAs, such as transfer RNAs, ribosomal RNAs and spliceosomal RNAs are modified and depend on the modifications for their biogenesis and function. tRNA is one of the most extensively modified RNAs with, on average, 13 modifications per molecule [30]. It is well documented that most tRNA modifications play one of two major roles: (a) stabilizing tRNA's tertiary structure; and/or (b) aiding in codon–anticodon recognition [31]. Such remarks mainly concern modifications in the D- and T-arms, which stabilize the tRNA structure, while modifications in the anticodon arm affect codon recognition, particularly in the highly modified wobble position [32]. tRNA modification are required for specific aminoacylations by aminoacyl–tRNA synthetases [33]. Yeast $tRNA^{Glu}$, deprived of the hypermodified nucleotide $mcms^2U34$, is a poor substrate for GluRS, exhibiting a 100-fold reduction in its specificity constant (kcat/KM). The presence of the m^6A modification in mRNA regulates several molecular processes, such as transcription, pre-mRNA splicing, mRNA export and stability, as well as translation [34]. While pseudouridine (Ψ) and m^5C modifications affect mRNA stability [35], mRNA containing Ψ synthetized in vitro was more stable than unmodified RNA with the same nucleotide sequences in mammalian cells. A similar effect was observed when m5C was introduced into mRNA [36]. The knockout of methyltransferase NSUN2 causes a decrease in the expression of p16 mRNA. One of the most prevalent forms of post-transcriptional RNA

modifications is the conversion of adenosine nucleosides to inosine (A-to-I), mediated by the adenosine deaminase acting on RNA (ADAR) family of enzymes. A–I editing changes how the codon is read by the ribosome, because I is read as G [37]. Moreover, the presence of inosine in the double-stranded region decreases mRNA stability and alters splice sites. In ribosomal RNAs, the most commonly found modified nucleotides are 2′O methylated nucleotides and Ψ [38]. In *H. sapiens*, 28S RNAs, 10 methylated nucleosides and 95 pseudouridines were found. They are distributed over important regions, including the decoding and tRNA binding sites (the A-, P- and E-sites), the peptidyl-transferase center and the intersubunit interface, facilitating efficient and accurate protein synthesis. Disturbances in tRNA and rRNA methylation processes alter cellular functions and the deregulation of these pathways can lead to complex diseases [39,40].

In this review, we present the characteristics, detection methods and the role of the modified nucleotides present in long noncoding RNAs. We also discuss the influence of lncRNA modifications on the development of mammals and neoplastic processes.

2. RNA Modification Mechanisms

There are three groups of protein involved in modifying RNA metabolism [27,28]. The first group (writers, Figure 2) consists of enzymes introducing modified nucleotides into RNA during posttranscriptional RNA modifications; the second group of proteins interacts with modified nucleotides (readers); and the third group is involved in removing modification labels (erasers).

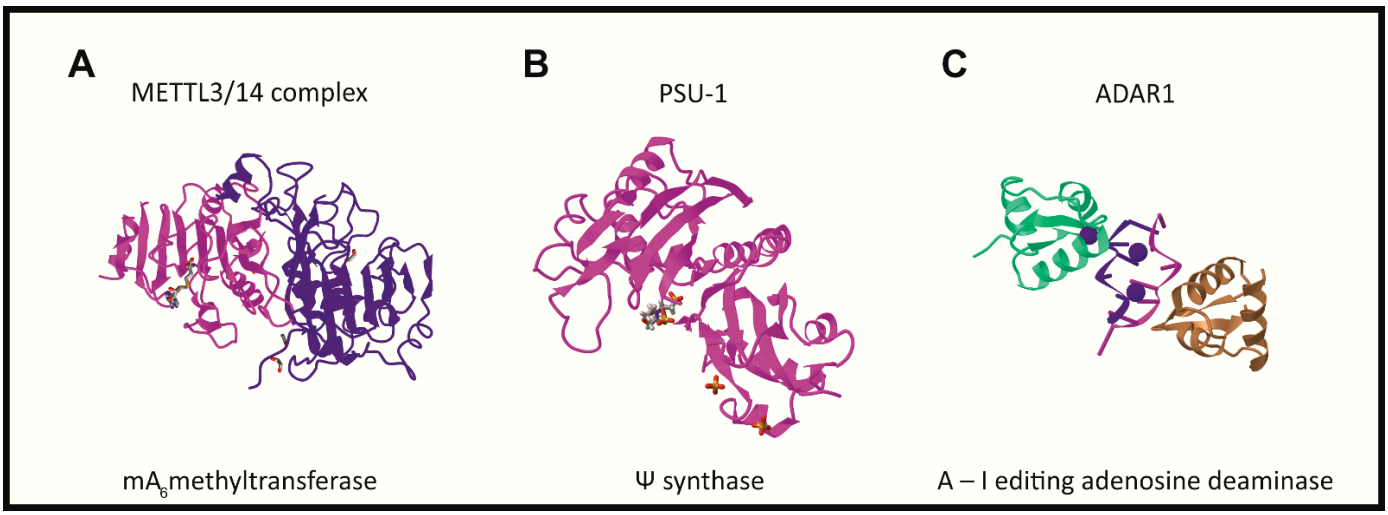


Figure 2. The crystal structure of "writer" enzymes, forming the following lncRNA modifications: (**A**)—METTL3/METTL14 m^6A methyltransferase [41]; (**B**)—PSU1 pseudouridine synthase [42]; and (**C**)—ADAR1 A-I editing adenosine deaminase [43].

6-methyl adenosine (m^6A) is one the most abundant RNA modifications, occurring in divergent members of the RNA family such as mRNA, rRNA, tRNA, snRNA and also lncRNA. It is usually present in several sites per transcript [44]. In humans, the formation of the m^6A modification is connected with the methylase complex (writer). Core MTase heterodimer complexes comprise of a methyl transferase such as 3 (METTL3) and a methyltransferase, such as 14 (METTL14) [41,45–47]. Crystallographic and biochemical studies have shown that METTL3 is S-adenosylmethionine methyltransferase with catalytic properties, while METTL14 serves as an RNA binding platform (Figure 2A) [48]. Additionally, interactions of the METTL3/METTL14 complex with other factors like: the Wilms' tumor 1-associating protein (WTAP) [49–51], KIAA1429 (VIRMA) [45,52], the zinc finger CCCH domain containing protein 13 (ZC3H13) [53], the RNA-binding motif protein 15 (RBM15), and its paralog-RBM15B, have been mentioned (Figure 3) [26,54,55]. The WTAP protein is necessary for METTL3/METTL14 complex activity, because the deletion of this protein

decreases the m^{6}A modification levels [27,49,56]. Similarly, KIAA KIAA1429 silencing also causes a substantial reduction in the amount of m^{6}A in RNA. RBM15/15B interacts with METTL3 in a WTAP-dependent manner and mediates the binding of the complex to specific RNA sites [27,57]. Recently, METTL16 was characterized as another "writer" protein, which interacts with several types of RNAs: mRNA, U6 and lncRNA [58–61]. Unlike METTL3/METTL14, which modifies A to m^{6}A in coding RNAs, METTL16 can methylate both coding and noncoding RNAs [62]. While the METTL3/METTL14 complex plays the role of the writer methyltransferase during RNA modification, METTL16, which is involved in MALAT1 epigenetics, acts as both the m^{6}A writer, as well as the reader [63,64].

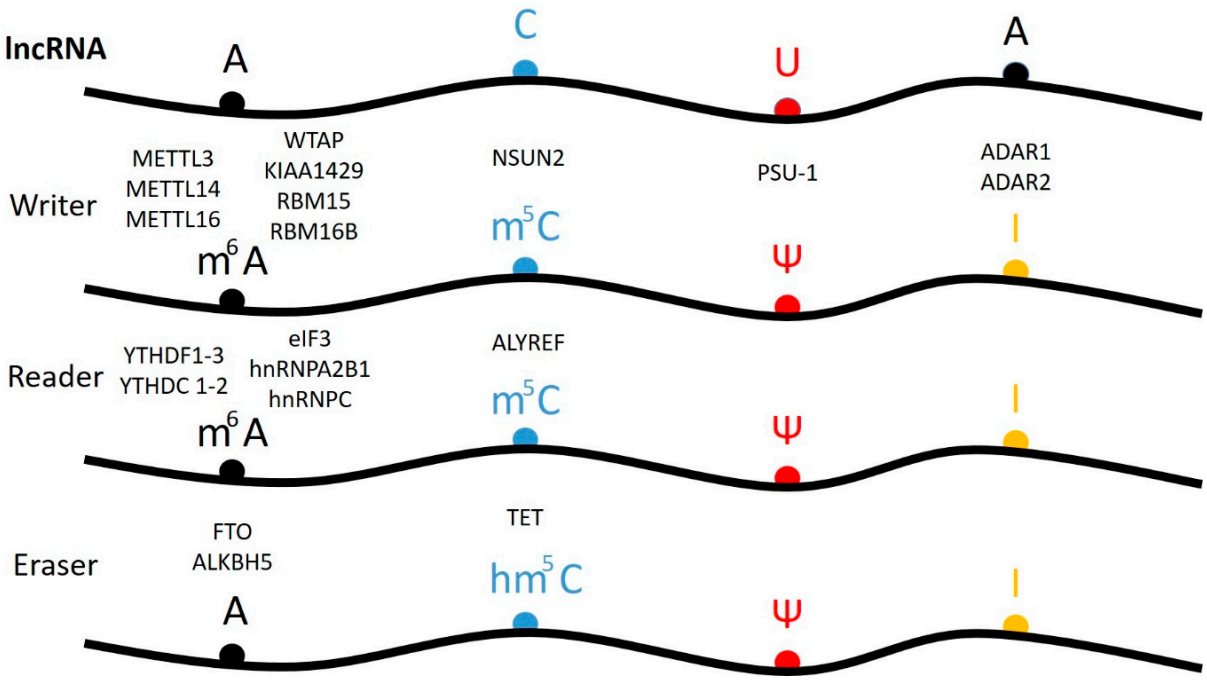


Figure 3. Schematic of lncRNA modification pathways. Writer, reader and eraser proteins are shown. **Nucleotides:** A—adenosine; C—cytidine; U—uridine. **Modifications**: m^{6}A—6-methyladenosine; m^{5}C—5-methylcytosine; Ψ—pseudouridine; I—inosine; hm^{5}C—5-hydroxymethylcytidine. **Writer proteins**: METTL3, 14, 16 methyltransferase-like; WTAP—Wilms' tumor 1 associating protein; KIAA1429—methyltransferase; RBM 15—RNA-binding motif protein 15 and its paralog RBM15B; NSUN—NOL1/NOP2/SUN domain family member; PUS—pseudouridine synthase; ADAR1, ADAR2—adenosine deaminase acting on RNA. **Reader proteins**: YTHDC, YTHDF—subgroups of YTH domain containing proteins; ZCCHC4—zinc-finger CCHC domain-containing protein 4; eIF3—eukaryotic initiation factor 3; hnRNPA2B1 and hnRNPC—heterogeneous nuclear ribonucleoproteins; ALYREF—ALY/REF export factor. **Eraser proteins**: FTO—fat mass and obesity-associated protein; ALKBH5—alpha-ketoglutarate-dependent dioxygenase AlkB homolog 5; TET—ten–eleven translocation protein.

Like DNA and histone modifications pathways, the m^{6}A modification has two specific erasers, FTO and ALKBH5 [64,65]. FTO (fat mass and obesity-associated protein) removes the methylation trace by oxidizing m^{6}A to N^6-hydroxymethyladosine or N^6-formyladenosine, which are chemically unstable and can hydrolyze to the final adenine product [66–68]. Another eraser is the homologous protein ALKBH5, which catalyzes the direct removal of the methyl group from adenine [69]. Both demethylases—namely FTO and ALKBH5—belong to the AlkB family, which contains a conserved iron binding motif and an α-ketoglutarate interaction domain [67,70–72]. FTO is highly expressed in the brain, neurons and in muscle [73], while ALKBH5 is prevalently expressed in the testis and lungs [74]. Whereas the FTO protein has been linked to obesity, the ALKBH5 protein is essential to spermatogenesis.

m^{6}A modification sites are recognized by "reader" proteins to generate functional signaling. The YT521-B homology (YTH) domain-containing protein family are the most

predominant m^6A readers and directly bind to the m^6A-modified RNA bases. The human YTH domain containing protein family consists of five proteins, namely, YTHDF1–3 and YTHDC1–2, which are conserved in mammalian genomes [75]. YTHDC1 is a nuclear protein involved in gene splicing [76–78]. YTHDF1–3 are putative RNA helicases which, apart from the YTH domain, contain a helicase domain, ankyrin repeats, and a DUF1065 domain [79,80]. Moreover, YTHDC2 acts as a scaffold molecule in regulating the spermatogenesis chromosome silencing effect of lncRNA XIST [53,81]. The eukaryotic initiation factor 3 (eIF3) binds to the m^6A located in the 5′UTR region of mRNA and is involved in cap-independent translation [82–84]. However, members of the hnRNP family, including HNRNPA2B1 and HNRNPC, choose their target transcripts by screening the RNA binding motifs (RBMs), which are more accessible to them as a result of the m^6A modification [85,86]. This mechanism is known as the "RNA epigenetic m^6A switch", which means m^6A alters the local structure of mRNA or lncRNA, to facilitate the binding of HNRNPs for biological regulation [87]. Other m^6A readers include the insulin-like growth factor 2 mRNA binding protein (IGF2BP) family, which was reported to regulate the stability of m^6A methylated RNAs [65,88,89].

5-methyl cytidine (m^5C) was found in both DNA and RNA. Two writer m^5C methyltransferases (MTases) have been shown to catalyze the m^5C modification of eukaryotic RNA. However, their substrate specificity and cellular functions are not completely understood [90,91]. Moreover, there are eight known eukaryotic RNA (C5 cytosine) methyltransferases (writers). One of them, RNA DNMT2, resembles DNA methyltransferases in its structure and characteristics [92,93], whereas the second group comprised of seven members, the MTPases (NSUN), contains the conserved NOL1/Nop2/Sun motif [94]. The NSUNs methylate tRNA (NSUN2, NSUN6), rRNA (NSUN1, NSUN5), mRNA (NSUN2), and ncRNA (NSUN2, Figure 2B), as well as mitochondrial rRNA (NSUN4) and mitochondrial tRNA (NSUN3), respectively. NSUN7′s specificity is currently unknown [94]. In 2017, Yang et al. presented evidence through in vitro and in vivo studies that m^5C formation in mRNAs is mainly catalyzed by the NSUN2 type RNA methyltransferase [95]. NSUN2-mediated m^5C methylation promotes the export of mRNA from the nucleus to the cytoplasm in an ALYREF-dependent manner. The RNA binding protein, ALYREF is an m^5C reader and is necessary for the nuclear export of m^5C-modified mRNAs. Unlike m^6A demethylases, it lacks an enzyme that specifically converts m^5C to C. Recently, it has been proven that the ten-eleven translocation (TET 1–3) proteins, a family of 5-methylcytosine dioxygenases, catalyze the successive oxidation of m^5C to the final product, 5-hydroxymethylcytosine (hm^5C), and can therefore be classified as eraser proteins [96]. The TET protein oxidation mechanism can be divided into two steps. In the first step, the activation of the dioxide molecule occurs, requiring the presence of Fe (II) ions and α-ketoglutaric acid (αKG), in order to convert the dioxygen molecule into a highly active Fe (IV)-oxo intermediate. In the second step, the C–H bond is oxidized to form a C–OH hydroxy product [97]. Moreover, hm5C modifications of RNA are involved in stem cell pluripotency and impact translation efficiency [97,98]. The TET2 protein binds to the promoter region of the oncogenic long noncoding RNA (lncRNA-ANRIL) and regulates the expression of ANRIL and its downstream genes. Additionally, the overexpression of the TET2 protein inhibits ANRIL lncRNA abundance, resulting in the decreased risk of gastric cancer [99]. However, the function of hydroxymethyl cytosine in other RNAs, including lncRNA, is not fully understood. The oxidation mechanism may provide an additional layer of epigenetic regulation to the mammalian genome.

Pseudouridine (Ψ) is a ubiquitous modified nucleotide, mainly found in rRNA, tRNA and ncRNA [28]. The estimated cell content is high, exceeding 5%. Ψ has an unusual nucleoside, containing a C–C glycosidic bond, instead of the N-C bond found in the rest of the nucleosides [100]. As a result of U being replaced by Ψ, an additional hydrogen bond donor is present at the non-Watson–Crick edge. The distinct structure of Ψ increases both the rigidity of the phosphodiester backbone, as well as the thermodynamic stability of Ψ–A, compared with U–A [101,102].

Pseudouridine writers, called pseudouridine synthases (PUSs), recognize substrates and catalyze the isomerization of U to Ψ, without the need for cofactors (Figure 2B) [42,100,103]. However, PUS enzymes are unable to isomerize free nucleotides. Pseudourydilation is known to follow two different mechanisms. In the first, the RNA-dependent pathway involves the formation of an RNP complex containing a H/ACA RNA box, cofactors, and four core proteins [104]. Box H/ACA RNAs are among the most evolutionarily conserved families of small ncRNAs and are present in all eukaryotes. In rRNA pseudouridylation, small nucleolar RNAs act as guides that recognize targets with sequence complementarity, thus directing pseudouridylation in a site-specific manner [104–106]. In humans, the four core proteins associated with box H/ACA RNAs are CBF5, NHP2, NOP10 and GAR1 [105,107]. However, only CBF5 catalyzes the U-to-Ψ isomerization reaction. In the solved crystal structure, three of the proteins interact with the H/ACA guide RNA or substrate RNA, while GAR1 may regulate substrate loading and release [104,108].

Alternatively, the RNA-independent pathway relies upon the direct recognition of targets by PUS complexes, often at conserved structural or sequence motifs [109,110]. In contrast to the m^6A modification, specific Ψ eraser or reader proteins have not yet been identified. This is due to the high stability of the C–C bond and the inability of the potential eraser proteins to cleave it, what is necessary for them to exchange Ψ for U. Furthermore, the lack of reader proteins that specifically bind to Ψ means it is difficult for proteins to identify the more subtle modification, which results in the C–N bond being replaced by a C–C bond, between uracil and ribose [109]. In humans, 10 proteins (PUS1–10) involved in RNA modification, with an annotated Ψ synthase domain, have been found (writers). These are classified into five families (TruA, TruB, TruD, RluA, and PUS10), based on their bacterial counterparts [108–111]. Although the primary sequences have diverged, all PUS synthases, including CBF5, share a conserved catalytic domain and likely a conserved catalytic mechanism, based on the solved crystal structure. In the case of m^6A and m^5C, the reader and the eraser proteins have been identified [26,27]. It is also possible that there could be readers and erasers for the Ψ modification; however, their existence has not yet been proven.

Inosine (I) differs from adenine in that it possesses a carbonyl group instead of an amino group at position 6 of the purine ring. This modification only occurs in the double-stranded regions of mRNA, tRNA, rRNA, and ncRNAs and is catalyzed by the writer protein, adenosine deaminase acting on RNA (ADAR). In vertebrates, a family of three ADAR proteins, ADAR1 (Figure 2C), ADAR2, and ADAR3, has been identified [43,112]. Structural analysis has shown that ADAR enzymes contain a C-terminal conserved catalytic deaminase domain, with two (ADAR2 and ADAR3) or three (ADAR1) dsRBDs in the N-terminal portion. The full length ADAR1 protein also contains a N-terminal Zα domain with a nuclear export signal and a Zβ domain, while ADAR3 has an R-domain. ADAR1 and ADAR2 catalyze all the currently known A-to-I editing sites. In contrast, ADAR3 has no documented deaminase activity. It has been postulated that the heterodimerization of ADAR3 with either ADAR1 or ADAR2 might render ADAR1 and ADAR2 inactive. Inosine essentially mimics the chemical properties of guanosine, therefore ADAR proteins introduce an A-to-G substitution in transcripts. These changes can lead to specific amino acid substitutions, altering protein composition. The presence of inosine in RNA influences mRNA alternative splicing, ncRNA-mediated gene silencing, or changes in the transcript's localization and stability [113]. The lncRNA-mRNA duplex is recognized by the ADAR double stranded specific enzyme, which converts adenine to inosine [114]. In the case of lncRNA A-to-I editing, most of the information comes from bioinformatics analysis [12]. About 200,000 editing sites exist in human lncRNAs. Most of them (65%) are located within those sites that influence their secondary structure. Accordingly, both edited and unedited lncRNAs can have different functions [115].

3. Detection of Modified Nucleotides

The information presented above suggests that lncRNA epigenetics are important to cell differentiation and may be involved in controlling organism development. This has motivated many laboratories to screen ncRNAs for the presence of modified nucleotides, and to study the changes of the modification pattern during cell development [26,27].

Application NGS methods to epigenetic mapping is so far difficult because they typically do not detect modified nucleosides [116,117]. Developing single-base resolution sequencing, which could quantify the relatively low abundance of modified nucleotides in lncRNA, is a significant challenge. The identification of transcriptome-wide RNA modifications has been approached using different strategies.

The study of RNA modification started in 1957 when the first modified nucleoside, pseudouridine, was discovered in bulk yeast RNA, using paper chromatography [118]. The first of these, the direct approach, uses antibody immunoprecipitation. The antibody specifically recognizes the modified ribonucleotides, making it possible to determine modifications at a global level [119,120]. The second direct technique uses two-dimensional thin layer chromatography (TLC) to analyze nucleotide composition [120,121]. During the first step, the isolated RNA molecules are hydrolyzed by nuclease, which leaves 3′ phosphate, then after ^{32}P labeling at 5′ side of the nucleotide, 5′3′ diphosphate nucleotide is formed. In the last step, 3′ phosphate is removed and 5′ ^{32}P labeled nucleotides are separated by TLC. There are several TLC methods used to detect a specific modification. Liu et al. developed the SCARLET method. The method relies on site-specific cleavage, radiolabeling, followed by ligation-assisted extraction and thin-layer chromatography [122,123]. Using this technique, the exact location of the m6A residue was determined, which are key parameters in studying the cellular dynamics of m6A modification. The SCARLET method starts with total RNA or with a total polyA$^+$ RNA sample. In the second step, a candidate site in a candidate RNA of interest has to be chosen. In the third step, RNase H cleavage is guided by a complimentary 2′–OMe/2′-H chimeric oligonucleotide to achieve site-specific cleavage 5′ to the candidate site. The cut site is labeled with ^{32}P and the ^{32}P labeled RNA fragment is splint ligated to 116-nucleotide single stranded DNA oligonucleotide, using DNA ligase. The sample is then digested with RNase T_1/A to completely digest all RNA, whereas the ^{32}Plabelled candidate site remains with the DNA nucleotide as DNA-^{32}P(A/m^{6}A)p and DNA-^{32}P(A/m^{6}A)Cp, which migrate as 117/116 mers on denaturing gel. The labeled band is excised from the gel, digested with nuclease P_1 into mononucleotides containing 5′ phosphate, and then the m^{6}A modification status is determined by TLC [124]. This method was successfully used to determine the modified nucleotides like m^{6}A, m^{5}C, Ψ, and possibly other unknown modified nucleotides, in several coding and noncoding RNAs. The method is laborious and with low throughput and it may be possible to substitute it with ultra-performance LC–MS methods (UPLC–MS) [124,125].

The high-throughput m^{6}A mapping strategies were based on the immunoprecipitation of modified RNA molecules, using m^{6}A-specific antibodies coupled to the subsequent NGS sequencing. In the following procedures, m^{6}A Seq and MeRIP-Seq, ncRNA is fragmented to a size of about 200 nucleotides and immunoprecipitated by a m^{6}A specific antibody, attached to the magnetic beads. Then, the RNA separated with a magnet is subjected to a second round of m^{6}A immunoprecipitation. The resulting RNA pool, which is highly enriched with m^{6}A-containing RNAs, is used for library construction and NGS sequencing [117,126]. Both methods provide a rather low resolution. In another direct m^{6}A detection technique, miCLIP [127–129], an additional step was added to the MeRIP-Seq method. The specific antibodies are bound to the m^{6}A mark in the RNA chain and cross-linked using UV light, with a wavelength of 254 nm. miCLIP allows for a high-resolution detection of m^{6}A in RNAs.

The detection of the m^{6}A nucleotide in RNA, using an indirect approach, is difficult because few chemical reagents modify the methyl group. However, NOseq, a method for the detection of m^{6}A in RNA after chemical deamination by nitrous acid, has recently been introduced [130]. Nitrous acid deaminates adenosines to inosine, while the m6A residue is

resistant to such modifications. The application of NGS after modification to detect m6A sites in MALAT1 lncRNA [130].

The indirect approach, in which chemical compounds selectively react with the modified ribonucleotides, was also used during the detection of m^5C or Ψ in RNAs. Reverse transcription, which is utilized in the next step, enables the detection of the modification sites [119,120]. The most common indirect method used to determine m^5C modification sites in RNA, i.e., the bisulfite conversion of cytidine to uridine has been applied successfully to determine the presence of this modified nucleotide in DNA [131]. The method is based on the fact that m^5C modified cytosine is resistant to bisulfite cytosine deamination. Thus, simply comparing the NGS sequences of RNA molecules subjected to bisulfite treatment with those that remain untreated, should pinpoint the modification sites [132–135]. There are several biochemical kits on the market, making it possible to prepare m^5C libraries, which are ready for sequencing. A disadvantage in the detection of m^5C present in RNA using bisulfite is that a large amount of RNA is initially required to compensate for the high losses caused by this reagent. Moreover, other modifications or double-stranded regions may be resistant to bisulfite treatment, especially under the milder reaction conditions required to maintain RNA integrity. Therefore, this method often requires additional confirmation with the aforementioned direct methods used for the detection of m^6A modifications—MeRIP and miCLIP. They differ in the application of m^5C-specific antibodies [136].

An alternative approach, which has been termed the "suicide enzyme trap", has been employed to characterize the substrates of the following m^5C-methyltransferases (m^5C-MTases), NSUN2 and NSUN4 [137,138]. By mutating m^5C-MTases to form irreversible covalent bonds with target residues, the resulting stable enzyme–RNA complexes are suitable for immunoprecipitation and mapping. This is also the case with the AZA-seq methodology, in which the "suicide inhibitor" nucleotide analog of 5-azacytidine is incorporated into cellular RNA and "traps" m^5C-MTases for pulldown and sequencing [139].

Determining the Ψ sites in the RNA chain requires indirectly analyzing this modification using carbodiimide chemistry. The chemical typically used for this is 1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate (CMCT), which reacts with guanosine, uridine and pseudouridine nucleosides. However, in alkaline conditions (pH 10.4), only Ψ–CMCT adducts are stable [140,141]. The bulky CMCT group attached to N3 on Ψ hinders reverse transcription and results in cDNA being truncated. This facilitates the detection of Ψ at a single nucleotide resolution level [140]. A new version of this method, called CeU-seq or Pseudo seq (N3-CMCT-enriched pseudouridine sequencing), was recently developed [142]. The RNA molecules were incubated with a CMCT derivative, and subsequently using the click chemistry approach, the N3-CMCT-Ψ adduct is labeled with a DBCO–(PEG) 4-biotin complex. The immunoprecipitation of Ψ RNA using streptavidin beads results in the enrichment of Ψ-containing RNA. In the next step, the Ψ enriched RNA is used for cDNA library preparation. In the final step, the library is deep-sequenced using the NGS protocol. Sites of pseudouridylation with single nucleotide resolution can be identified by subjecting the data obtained through NGS sequencing of Pseudo-seq libraries, to computational analysis.

Recently, Pan et al. developed a method that uses a CMC-Ψ-induced RT stop with an additional step of site-specific ligation, followed by PCR, to generate two unique PCR products, that correspond to the modified and unmodified uridine. The modification is visualized in the PCR products using gel electrophoresis [143].

The editing events are typically identified using the direct detection method, by comparing the cDNA sequences with the corresponding genomic DNA sequence [144]. The edited inosine base pairs with cytidine in cDNA, hence editing is visible as an A-to-G sequence change. Recently, sophisticated bioinformatics tools have been developed to maximize detection accuracy, while minimizing the detection of false positives [145,146]. Furthermore, considerable effort has been made to comprehensively detect and remove known SNPs from editing datasets and databases [147,148]. As identifying true editing

sites from transcriptome sequencing data is difficult, alternative methods aimed at marking inosine have been developed. Another direct method used to detect editing sites utilizes the co-immunoprecipitation of ADAR enzymes with the bound substrate RNA and the subsequent microarray analysis of these associated RNA sequences [149]. The disadvantage of this method is the fact that the association of ADAR with RNA is not necessarily indicative of editing. An alternative method, based on the finding that glyoxal reacts with guanosine to form a stable adduct, whereas inosine glyoxal adducts are unstable, has been developed. Moreover, guanosine glyoxal/borate adducts are resistant to RNase T_1 digestion [150,151]. RNase T_1 specifically cleaves RNA after guanosine or inosine but is inhibited by guanosine glyoxal/borate adducts. The cleavage of glyoxal-modified RNA creates RNA fragments that carry inosine at their termini, as an input for sequencing. A method similar to the direct approach, used to determine inosine modifications, called inosine chemical erasing (ICE), was developed. [152,153]. ICE involves the treatment of RNA with acrylonitrile, which converts the inosine to N1-cyanoethylinosine in the process of cyanoethylation, and results in the formation of an inosine/acrylonitrile adduct that inhibits base pairing with cytidine and stalls reverse transcription. The total RNA is either treated with acrylonitrile or left untreated and then reverse transcribed into cDNA. In untreated RNA, the A or I at a given position is converted into T or G, respectively. In the treated sample, A is converted to T, while the presence of inosine/acrylonitrile adducts blocks reverse transcription, leading to shorter cDNAs. The ICE method was combined with NGS (ICE-seq), requiring the fragmentation of poly(A)-enriched RNA before cyanoethylation and reverse transcription. ICE-seq makes it possible to identify the editing sites throughout the transcriptome. The gel purification of longer cDNA fragments effectively erases these shorter inosine/acrylonitrile adduct-containing cDNAs from the library. The subsequent sequencing and comparison of the libraries identifies inosines, by detecting erased reads upon cyanoethylation [154].

NGS methods based on short-read sequencing have difficulty in determining the modification patterns of the entire transcript sequence. Recently, a direct modification detection method, called nanopore sequencing, has been developed [155–157]. The Oxford Nanopore Technologies (ONT) sequencer can directly sequence individual native RNA or DNA molecules. The nanopore sequencer can measure disruptions in the current, compared to the raw current intensity, as the RNA or DNA passes through the nanopore in the dielectric membrane. This technology is able, in principle, to identify the nucleotide passing through the nanopore [158]. Most importantly, the method does not require that the RNA be processed, i.e., converted into cDNA by a reverse transcriptase like other NGS methods, and RNA modifications are therefore preserved. ONT data analysis requires that the use of specialized bioinformatic software and several brands of software, such as Tombo, Epinano or ELIGOS, are available [117,155,159]. Tombo is a software used to detect modifications in DNA and RNA, such as the m^5C modification in DNA and RNA and the m6A modification in DNA [160]. The EpiNano software is used to detect the m6A modification in RNA [117,155]. The ELIGOS software compares the error profile between native RNA sequences obtained with direct RNA-seq and a reference sequence, which can be in vitro synthesized RNA, cDNA or the RNA background error model [161]. The greatest limitation of the nanopore sequencer is its comparatively low read accuracy, compared with short read sequencers. It needs to be highlighted that direct RNA modification analysis using nanopore sequencing is rapidly developing and becoming more reliable, so its routine application in the field of RNA epigenetics is expected.

4. The Impact of lncRNA Epigenetics on lncRNA Function

It is estimated that human cells have over 50,000 lncRNA molecules coded in genes. Some of them contain introns and are spliced by the same machinery as pre-mRNAs [12]. Many mature lncRNAs are modified after transcription (Section 2). The application of NGS methods in combination with bioinformatic analysis revealed the occurrence of several modifications in different types of lncRNA molecules. Integrated data analysis of m^6A,

m^5C, Ψ and I sequencing studies was performed, highlighted the total amount of modifications in lncRNA, as well as the number of individual lncRNAs, that contain the respective modified nucleotide. The results were as follows: m^6A—(13357 modifications/12348 lncRNAs); m^5C—(9965 modifications/1072 lncRNAs); Ψ—(162 modifications/150 lncRNAS); I—(11726 modifications/3374 lncRNAs) [26,144,162]. The role modified nucleosides play in lncRNA is not yet completely understood. Most of the information available concerns the most abundant lncRNAs present in the cell [163]. lncRNAs can be classified as being *cis* acting or *trans* acting [164]. The *cis* acting lncRNAs recruit factors, either to the site of lncRNA transcription or to adjacent loci and involve lncRNA XIST and lncRNA H19 [165,166]. Trans acting lncRNAs act independently of their transcription sites, either by regulating expression from other loci in the nucleus or having transcription-unrelated functions anywhere in the cell involving MALAT1, HOTAIR, and lincRNA1238 [166]. Antisense noncoding RNA in the INK4 locus *ANRIL* is included in *cis* and in *trans* gene regulation [167].

XIST is a 17.5 kb capped, spliced and polyadenylated nucleotide transcript, transcribed from the *XIST* gene and involved in X chromosome inactivation [168,169]. Due to its size, XIST has many modifications like 78 sites m6A. 5 sites m5C and single site of Ψ [26]. In humans, multiple m^6A sites in XIST repeat regions have been identified. Patil and colleagues have shown that m^6A formation in XIST, as well as in cellular mRNAs, is mediated by RBM15 and RBM15B, which bind the m^6A-methylation complex and recruit it to specific sites in RNA [83]. Additionally, the knockdown of RBM15 and RBM15B, or the knockdown of METTL3 methyltransferase, impairs XIST-mediated gene silencing. The depletion of YTHDC1 was shown to result in defective X-chromosome inactivation, whereas tethering YTHDC1 to XIST rescues the phenotype in the absence of a functional m^6A methylation complex [83]. These data suggest that the biogenesis of m6A and its recognition is required for XIST-mediated transcriptional repression.

Many pathways contribute to the control of gene expression during development. Polycomb repressive complex (PRC2) and XIST are associated with gene repression in various developmental processes, such as X chromosome inactivation and genomic imprinting.

PRC2 binds with high affinity to the 5′-end region of XIST called the repeat A-region [137]. This functionally important domain is composed of 8.5 repeats of a 26-nucleotide sequence. In human XIST, five m^5C marks were also detected. The presence of m^5C methylation in the XIST transcript prevents the binding of the PRC2 in vitro. In Xist lncRNA, the presence of pseudouridylation and A–I editing sites has been confirmed, however, their roles are unknown at this stage so far [170,171].

The metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) gene is located at chromosome 11q13 and encodes a 7.9 kb transcript. MALAT1 is a highly modified lncRNA, carrying multiple m^6A modification sites. Moreover, a modification of the m^6A site in MALAT1 mediates the m^6A switch, which allows for the HNRNPC (reader) protein to be bound to a U-rich tract in the strand opposite to the m^6A modification site [172]. Thus, the m^6A modification regulates the binding of proteins to MALAT1 lncRNA. The recent analysis of MeRIP seq data revealed thousands of m^6A switches, which are involved in alternative RNA splicing and abundance [173]. The MALAT1 transcript has a triple-helical RNA stability element at the 3′ end, which can be recognized by METTL16 [61,63]. This suggests the existence of a possible m6 A modification site or a function for guiding METTL16 onto its targets. Recently Wang et al. postulated that m^6A modifications in MALAT1 are important to the metastatic capacity of esophageal cancer cells, both in vitro and in vivo [174]. They also show that the recognition of the m^6A mark in MALAT1 by the YTHDC1 reader protein plays a critical role in maintaining the composition of nuclear speckles and their genomic binding sites. Interestingly, the phenotypes induced by MALAT1-m^6A deficiency could be largely rescued both in vitro and in vivo, by artificially tethering YTHDC1 onto MALAT1. In addition, *MALAT1* lncRNA is subject to post-transcriptional m5C modification; five m^5C sites been found to regulate chromatin-related roles in other lncRNAs, such as HOTAIR and XIST [135]. Although the exact role played by pseudouridine and inosine

modifications remains to be explained, and the presence of each of the three inosines increases the stability of MALAT1 by 2–3 kcal/mol [28,170].

HOX transcript antisense RNA (HOTAIR) lncRNA is a transcript of the antisense strand of the *hoxC* gene, which is spliced and polyadenylated [175] Within HOTAIR, 14 individual m^6A sites were identified, with a single site (A783) being consistently methylated. HOTAIR interacts with the nuclear m^6A reader YTHDC1 at the methylated A783 and at additional sites [176]. Localization in chromatin strongly depends on the m6A modification at site A783 of HOTAIR, while the modification of other m^6A sites mediates high HOTAIR levels. Additionally, m^6A-dependent YTHDC1–HOTAIR interactions are required for gene repression, independent of the expression level and chromatin recruitment. The previous results demonstrate that site-specific cytosine methylation occurs in lncRNA HOTAIR [135]. The methylation of C1683 is widespread in different cell types and it is not limited by the abundance of HOTAIR RNA levels. Furthermore, the degree of methylation of C1683 appears to be remarkably high, suggesting that it might be important to HOTAIR's functioning. In this respect, it is interesting to note that C1683 is located in the vicinity of the region that has previously been shown to interact with the LSD1 complex [13]. It is therefore tempting to speculate that the methylation of this cytosine may affect the ability of HOTAIR to interact with LSD1.

Yang et al. identified the cytoplasmic long intergenic noncoding RNA 1281 (lincRNA 1281), whose function is regulated by m^6A modifications [177]. This lincRNA is necessary for the differentiation of mouse embryonic stem cells (mESCs) and acts as a ceRNA by sequestering let-7 miRNAs [177]. Notably, lincRNA 1281 contains m^6A marks in its 3′-end region, which are required for the binding of let-7 miRNA. It has been proposed that the presence of m^6A in lincRNA1281 can act as a m^6A-switch for specific RNA binding proteins, which will eventually regulate their interaction with let-7 miRNA. However, the identity of such proteins has not yet been discovered. A similar mechanism has already been proposed for the binding of the HuR (ELAVL1) protein and miRNAs, to the mRNAs encoding developmental regulators in mESCs.

H19, an imprinted lncRNA with a size of 2.3 kb, plays an important role in embryonic development [178]. The knockdown of METTL3 or METTL14 notably reversed the hypoxic preconditioning-induced (HPC-induced) enhancement of cell viability, anti-apoptosis ability, and lncRN AH19 expression [179]. Methylated RNA immunoprecipitation (IP) indicated that the knockdown of METTL3 or METTL14 decreased the m^6A levels in lncRNA H19. The RNA binding protein immunoprecipitation (RIP) assay showed that METTL3 and METTL14 can directly bind with lncRNA H19. The m^5C modification in lncRNAH19 can increase its stability [180]. Furthermore, m^5C-modified lncRNA H19 can be specifically bound by G3BP1, a well-known oncoprotein, which results in MYC accumulation. NSUN2 was shown to methylate lncRNA H19 and affect its stability. The Ras-GTPase-activating protein-binding protein 1 (G3BP1) was confirmed to bind methylated lncRNAH19, based on the presence of NSUN2. lncRNA H19 has two editing sites, and their presence increases the lncRNA stabilization energy by 3 kcal/mol [171].

The human steroid receptor RNA activator (SRA) is a transcript of the *sra1* gene, whose size ranges from 0.7 to 0.9 kb. It is a dual-function RNA, which acts as both an lncRNA and an mRNA [181]. lncRNA SRA regulates several processes, such as cell cycle proliferation, as well as insulin, Notch, and TNFa signaling [182]. Currently, depending on the technique used, the data show the presence of 1–4 m^6A modifications in lncRNA SRA, of an unknown function [26]. In a subsequent study, the same authors identified a specific uridine residue in SRA1 (U206), whose modification by PUS1 (or PUS3) might induce a functional switch, which regulates nuclear receptor signaling [26,181]. As of now, inosine has not been detected in this lncRNA.

The human plasmacytoma variant translocation 1 (PVT1 lncRNA) is a large locus, which is longer than 30 kb and is located at 8q24.21. It serves an oncogenic role in a variety of malignant tumors, such as colorectal cancer [183]. PVT1 lncRNA is highly modified and contains m^6A, m^5C and Ψ nucleosides [36]. Recently, it has been proven that the m^6A

PVT1 lncRNA modification regulates epidermal stemness via its interaction with the MYC protein [184]. The suggested mechanism involves the association of Pvt1 with the MYC protein and prevents MYC degradation. m^6A methylation is an important regulator of this process. Moreover, the RNA m^6A modification mediated by the METTL3/METTL14 complex (Figure 2A) regulates epidermal stemness by controlling Pvt1 and MYC interactions through Pvt1 methylation, uncovering a key and novel molecular mechanism underlying skin tissue homeostasis, regeneration and wound repair. Additionally, PVT1 lncRNA has been identified as a pseudourydilation target [26,27]. Some of the Ψ sites were located within functional lncRNA motifs, indicating the potential regulatory impact of Ψ on lncRNAs.

5. Conclusions and Future Directions

Information available in the literature suggests that lncRNA epigenetics play a role in gene regulation through various mechanisms. First, the modification of lncRNAs may change their structure and affect their interactions with proteins, like in the case of lncRNA. Second, the modification of lncRNAs could mediate transcription repression [54]. Third, lncRNA modifications might alter its subcellular distribution [185,186]. Finally, lncRNA modifications regulate the stability of lncRNAs like H19, MALAT1, XIST, etc. However, overall, there are still relatively few studies concerning lncRNA epigenetics. Due to the fact that lncRNA has many distinct functions in the cell, there is a need to further elucidate the components required for lncRNAs modification and recognition. It is likely that the development of new sequencing techniques, such as nanopore sequencing, will facilitate the search for new modified lncRNAs. Modified lncRNAs are involved in oncogenesis [176,187] and they can therefore be excellent biomarkers of neoplastic processes [188]. The underlying mechanisms by which lncRNA modifications contribute to gene regulation and whether and how the RNA epigenetics of mRNAs differs from that of lncRNAs, as well as which proteins are involved in both RNAs remain to be elucidated. Understanding these mechanisms makes it possible to develop lncRNA-targeted therapies.

Author Contributions: M.K. and J.W. wrote manuscript and prepared figures. J.W. finally accepted the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Polish National Centre of Science 2016/21/NZ2/03468.

Acknowledgments: We thank B. Imiołczyk for critically reading the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Xue, C. Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108. [CrossRef] [PubMed]
2. Kapranov, P.; Willingham, A.T.; Gingeras, T.R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **2007**, *8*, 413–423. [CrossRef] [PubMed]
3. Layton, E.; Fairhurst, A.M.; Griffiths-Jones, S.; Grencis, R.K.; Roberts, I.S. Regulatory RNAs: A universal language for inter-domain communication. *Int. J. Mol. Sci.* **2020**, *21*, 8919. [CrossRef] [PubMed]
4. Uchida, S.V.; Adams, J.C. Physiological roles of non-coding RNAs. *Am. J. Physiol. Cell* **2019**, *317*, C1–C2. [CrossRef] [PubMed]
5. Quinn, J.J.; Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **2017**, *17*, 47–62. [CrossRef]
6. St Laurent, G.; Wahlestedt, C.; Kapranov, P. The landscape of long non-coding RNA classification. *Trends Genet.* **2015**, *31*, 239–251. [CrossRef] [PubMed]
7. Schulz, D.; Schwalb, B.; Kiesel, A.; Baejen, C.; Torkler, P.; Gagneur, J.; Soeding, J.; Cramer, P. Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* **2013**, *155*, 1075–1087. [CrossRef] [PubMed]
8. Zhu, S.; Wang, J.; He, Y.; Meng, N.; Yan, G.R. Peptides/Proteins encoded by non-coding RNA: A novel resource bank for drug targets and biomarkers. *Front. Pharmacol.* **2018**, *9*, 1295. [CrossRef]
9. Chugunova, A.; Navalayeu, T.; Dontsova, O.; Sergiev, P. Mining for small translated ORFs. *J. Proteome Res.* **2018**, *17*, 1–11. [CrossRef]
10. Matsumoto, A.; Nakayama, K.I. Hidden peptides encoded by putative noncoding RNAs. *Cell Struct. Funct.* **2018**, *43*, 75–83. [CrossRef]

11. Statello, L.; Guo, C.J.; Chen, L.L.; Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 96–118. [CrossRef]
12. Kazimierczyk, M.; Kasprowicz, M.K.; Kasprzyk, M.E.; Wrzesinski, J. Human long noncoding RNA interactome: Detection, characterization and function. *Int. J. Mol. Sci.* **2020**, *21*, 1027. [CrossRef] [PubMed]
13. Rinn, J.L.; Chang, H.Y. Long noncoding RNAs: Molecular modalities to organismal functions. *Annu. Rev. Biochem.* **2020**, *89*, 283–308. [CrossRef] [PubMed]
14. Zhang, X.; Wang, W.; Zhu, W.; Dong, J.; Cheng, Y.; Yin, Z.; Shen, F. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* **2019**, *20*, 5573. [CrossRef]
15. Roundtree, I.A.; Molly, E.; Evans, M.E.; Pan, T.; He, C. Dynamic RNA modifications in gene expression regulation. *Cell* **2017**, *169*, 1187–1200. [CrossRef] [PubMed]
16. Suzuki, M.M.; Bird, A. DNA methylation landscapes: Provocative insights from epigenomics. *Nat. Rev Genet.* **2008**, *9*, 465–476. [CrossRef] [PubMed]
17. Kim, M.; Costello, J. DNA methylation: An epigenetic mark of cellular memory. *Exp. Mol. Med.* **2017**, *49*, e322. [CrossRef] [PubMed]
18. Kumar, S.; Chinnusamy, V.; Mohapatra, T. Epigenetics of modified DNA bases: 5-methylcytosine and beyond. *Front. Genet.* **2018**, *9*, 640. [CrossRef] [PubMed]
19. Carter, C.; Zhao, K. The epigenetic basis of cellular heterogenity. *Nat. Rev. Genet.* **2020**, *169*, 1187–1200.
20. Boccaletto, P.; Machnicka, M.A.; Purta, E.; Piatkowski, P.; Baginski, B.; Wirecki, T.K.; de Crecy-Lagard, V.; Ross, R.; Limbach, P.A.; Kotter, A.; et al. MODOMICS: A database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* **2018**, *46*, D303–D307. [CrossRef]
21. Itaya, T.; Kanai, T.; Sawada, T. Structure of wyosine, the condensed tricyclic nucleoside of torula yeast phenylalanine transfer ribonucleic acid. *Chem. Pharm. Bull.* **2002**, *50*, 547–548. [CrossRef]
22. Grosjean, H. *DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution*; Landes Bioscience: Austin, TX, USA; CRC Press: Boca Raton, FL, USA, 2009.
23. Ajitkumar, P.; Cherayil, J.D. Thionucleosides in transfer ribonucleic acid: Diversity, structure, biosynthesis, and function. *Microbiol. Rev.* **1988**, *52*, 103–113. [CrossRef]
24. Schweizer, U.; Bohleber, S.; Fradejas-Villar, N. The modified base isopentenyladenosine and its derivatives in tRNA. *RNA Biol.* **2017**, *14*, 1197–1208. [CrossRef]
25. Thiaville, P.C.; Iwata-Reuyl, D.; de Crecy-Lagard, V. Diversity of the biosynthesis pathway for threonylcarbamoyladenosine (t^6A), a universal modification of tRNA. *RNA Biol.* **2014**, *11*, 1529–1539. [CrossRef]
26. Jacob, R.; Zander, S.; Gutschner, T. The dark side of the epitranscriptome: Chemical modifications in long non-coding RNAs. *Int. J. Mol. Sci.* **2017**, *18*, 2387. [CrossRef]
27. Dinescu, S.; Ignat, S.; Lazar, A.D.; Constantin, C.; Neagu, M.; Costache, M. Epitranscriptomic signatures in lncRNAs and their possible roles in cancer. *Genes* **2019**, *10*, 52. [CrossRef]
28. De Zoysa, M.D.; Yu, Y.T. Posttranscriptional RNA pseudouridylation. *Enzymes* **2017**, *41*, 151–167. [PubMed]
29. Torres, A.G.; Pineyro, D.; Filonava, L.; Stracker, T.H.; Batlle, E.; Ribas de Pouplana, L. A-to-I editing on tRNAs: Biochemical, biological and evolutionary implications. *FEBS Lett.* **2014**, *588*, 4279–4286. [CrossRef] [PubMed]
30. Pan, T. Modifications and functional genomics of human transfer RNA. *Cell Res.* **2018**, *28*, 1–10. [CrossRef] [PubMed]
31. Lyons, S.M.; Fay, M.M.; Ivanov, P. The role of RNA modifications in the regulation of tRNA cleavage. *FEBS J.* **2018**, *592*, 2828–2844. [CrossRef] [PubMed]
32. Avcilar-Kucukgoze, I.; Kashina, A. Hijacking tRNAs from translation: Regulatory functions of tRNAs in mammalian cell physiology. *Front. Mol. Biosci.* **2020**, *7*, 610–617. [CrossRef]
33. Sylvers, L.A.; Rogers, K.C.; Shimizu, M.; Othsuka, E.; Soll, D. A 2-thiouridine derivative in tRNAGlu is a positive determinant for aminoacylation by Escherichia coli glutamyl-tRNA synthetase. *Biochemistry* **1993**, *32*, 3836–3841. [CrossRef]
34. Boo, S.H.; Kim, Y.K. The emerging role of RNA modifications in the regulation of mRNA stability. *Exp. Mol. Med.* **2020**, *52*, 400–408. [CrossRef]
35. Karikó, K.; Muramatsu, H.; Welsh, F.A.; Ludwig, J.; Kato, H.; Akira, S.; Weissman, D. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Mol. Ther.* **2008**, *16*, 1833–1840. [CrossRef] [PubMed]
36. Li, Q.; Li, X.; Tang, H.; Jiang, B.; Dou, Y.; Gorospe, M.; Wang, W. NSUN2—Mediated m5C methylation and METTL3/METTL14—mediated m6A methylation cooperatively enhance p21 translation. *J. Cell Biochem.* **2017**, *118*, 2587–2598. [CrossRef] [PubMed]
37. Hoernes, T.P.; Faserl, K.; Juen, M.A.; Kremser, J.; Gasser, C.; Fuchs, E.; Shi, X.; Siewert, A.; Lindner, H.; Kreutz, C.; et al. Translation of non-standard codon nucleotides reveals minimal requirements for codon-anticodon interactions. *Nat. Commun.* **2018**, *9*, 4865. [CrossRef] [PubMed]
38. Sloan, K.E.; Warda, A.S.; Sharma, K.E.; Entian, K.D.; Lafontaine, D.L.J.; Bohnsack, M.T. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.* **2017**, *14*, 1138–1152. [CrossRef] [PubMed]
39. Decatur, W.A.; Fournier, M.J. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **2002**, *27*, 344–351. [CrossRef]
40. Popis, M.C.; Blanco, S.; Frye, M. Posttranscriptional methylation of transfer and ribosomal RNA in stress response pathways, cell differentiation, and cancer. *Curr. Opin. Oncol.* **2016**, *28*, 65–71. [CrossRef] [PubMed]

41. Wang, X.; Feng, J.; Xue, Y.; Guan, Z.; Zhang, D.; Liu, Z.; Gong, Z.; Wang, Q.; Huang, J.; Tang, C.; et al. METTL3/METTL14 m6A methyltransferase. *Nature* **2016**, *534*, 575–578. [CrossRef]
42. Czudnochowski, N.; Wang, A.L.; Finer-Moore, J.; Stroud, R.M.J. In human pseudouridine synthase 1 (hPus1), a C-terminal helical insert blocks tRNA from binding in the same orientation as in the Pus1 bacterial homologue TruA, consistent with their different target selectivities. *Mol. Biol.* **2013**, *425*, 3875–3887. [CrossRef] [PubMed]
43. Placido, D.; Brown, B.A.; Lowenhaupt, K.; Rich, A.; Athanasiadis, A. A left-handed RNA double helix bound by the Z alpha domain of the RNA-editing enzyme ADAR1. *Structure* **2007**, *15*, 395–404. [CrossRef] [PubMed]
44. McCown, P.J.; Ruszkowska, A.; Kunkler, C.N.; Breger, K.; Hulewicz, J.P.; Wang, M.C.; Springer, N.A.; Brown, J.A. Naturally occurring modified ribonucleosides. *Wiley Interdiscip. Rev. RNA* **2020**, *11*, e1595. [CrossRef]
45. Yang, Y.; Hsu, P.J.; Chen, Y.S.; Yang, Y.G. Dynamic transcriptomic m6 A decoration: Writers, erasers, readers and functions in RNA metabolism. *Cell Res.* **2018**, *28*, 616–624. [CrossRef] [PubMed]
46. Bedi, R.K.; Huang, D.; Sledz, P.; Caflisch, A. Crystal structure of the human METTL3-METTL14 complex bound to compound 1/adenosine (DHU_M3M_023) . *Chem. Med. Chem.* **2020**, *15*, 744–748.
47. Wang, P.; Katelyn, A.; Doxtader, K.A.; Nam, Y. Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* **2016**, *63*, 306–317. [CrossRef]
48. Liu, P.; Li, F.; Lin, J.; Fukumoto, T.; Nacarelli, T.; Hao, X.; Kossenkov, A.V.; Simon, M.C.; Zhang, R. m6A-independent genome-wide METTL3 and METTL14 redistribution drives the senescence-associated secretory phenotype. *Nat. Cell Biol.* **2021**, *23*, 355–365. [CrossRef]
49. Ping, X.L.; Sun, B.F.; Wang, L.; Xiao, W.; Yang, X.; Wang, W.J.; Adhikari, S.; Shi, Y.; Lv, Y.; Chen, Y.-S.; et al. Mammalian WTAP is a regulatory subunit of the RNA N6—methyladenosine methyltransferase. *Cell Res.* **2014**, *24*, 177–189. [CrossRef]
50. Kobayashi, M.; Ohsugi, M.; Sasako, T.; Awazawa, M.; Umehara, T.; Iwane, A.; Kobayashi, N.; Okazaki, Y.; Kubota, N.; Suzuki, R.; et al. The RNA methyltransferase complex of WTAP, METTL3, and METTL14 regulates mitotic clonal expansion in adipogenesis. *Mol. Cell Biol.* **2018**, *38*, e00116–e00118. [CrossRef]
51. Huang, J.; Chen, Z.; Chen, X.; Chen, J.; Cheng, Z.; Wang, Z. The role of RNA N6-methyladenosine methyltransferase in cancers. *Mol. Ther. Nucleic Acids* **2021**, *23*, 887–896. [CrossRef] [PubMed]
52. Schwartz, S.; Mumbach, M.R.; Jovanovic, M.; Wang, T.; Maciag, K.; Bushkin, G.G.; Mertins, P.; Ter-Ovanesyan, D. Perturbation of m6A writers reveals two distinct classes of mRNA methylation at Internal and 5′ Sites. *Cell Rep.* **2014**, *8*, 284–296. [CrossRef] [PubMed]
53. Wen, J.; Lv, R.; Ma, H.; Shen, H.; He, C.; Wang, J.; Jiao, F.; Liu, H.; Yang, P.; Tan, L.; et al. Zc3h13 regulates nuclear RNA m6A methylation and mouse embryonic stem cell self-renewal. *Mol. Cell* **2018**, *69*, 1028–1038. [CrossRef] [PubMed]
54. Patil, D.P.; Chen, C.K.; Pickering, B.F.; Chow, A.; Jackson, C.; Guttman, M.; Jaffrey, S.R. m^{6}A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **2016**, *537*, 369–373. [CrossRef] [PubMed]
55. Chen, Y.; Lin, Y.; Shu, Y.; He, J.; Gao, Y. Interaction between N6 -methyladenosine (m^6 A) modification and noncoding RNAs in cancer. *Mol. Cancer* **2020**, *19*, 94. [CrossRef]
56. Sorci, M.; Ianniello, Z.; Cruciani, S.; Larivera, S.; Ginistrelli, L.C.; Capuano, E.; Marchioni, M.; Fazi, F.; Fatica, A. METTL3 regulates WTAP protein homeostasis. *Cell Death Dis.* **2018**, *9*, 796. [CrossRef]
57. Meyer, K.D.; Jaffrey, S.R. Rethinking m (6) a readers, writers, and erasers. *Annu. Rev. Cell Dev. Biol.* **2017**, *33*, 319–342. [CrossRef]
58. Warda, A.S.; Kretschmer, J.; Hackert, P.; Lenz, C.; Urlaub, H.; Höbartner, C.; Sloan, K.E.; Markus, T.; Bohnsack, M.T. Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* **2017**, *18*, 2004–2014. [CrossRef]
59. Ruszkowska, A.; Ruszkowski, M.; Dauter, Z.; Brown, J.A. Structural insights into the RNA methyltransferase domain of METTL16. *Sci. Rep.* **2018**, *8*, 53. [CrossRef]
60. Mendel, M.; Chen, K.M.; Homolka, D.; Gos, P.; Pandey, R.R.; McCarthy, A.A.; Pillai, R.S. Methylation of structured RNA by the m^{6}A writer METTL16 Is essential for mouse embryonic development. *Mol. Cell* **2018**, *71*, 986–1000.e11. [CrossRef]
61. Ruszkowska, A. METTL16, methyltransferase-like protein 16: Current insights into structure and function. *Int. J. Mol. Sci.* **2021**, *22*, 2176. [CrossRef]
62. Pendleton, K.E.; Chen, B.; Liu, K.; Hunter, O.V.; Xie, Y.; Tu, B.P.; Conrad, N.K. The U6 snRNA m^{6}A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* **2017**, *168*, 824–835. [CrossRef] [PubMed]
63. Brown, J.A.; Charles, G.; Kinzig, C.G.; DeGregorio, S.J.; Steitz, J. Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 14013–14018. [CrossRef]
64. Nance, D.J.; Satterwhite, E.R.; Bhaskar, B.; Misra, S.; Carraway, K.R.; Mansfield, K.D. Characterization of METTL16 as a cytoplasmic RNA binding protein. *PLoS ONE* **2020**, *15*, e0227647. [CrossRef]
65. Zaccara, S.; Ries, R.J.; Jaffrey, S.R. Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 608–624. [CrossRef]
66. Fu, Y.; Dominissini, D.; Rechavi, G.; He, C. Gene expression regulation mediated through reversible m^{6}A RNA methylation. *Nat. Rev. Genet.* **2014**, *15*, 203–306. [CrossRef]
67. Gerken, T.; Girard, C.A.; Tung, Y.C.; Webby, C.J.; Saudek, V.; Hewitson, K.S.; Yeo, G.S.; McDonough, M.A. The obesity-associated FTO gene encodes a 2-oxoglutaratedependent nucleic acid demethylase. *Science* **2007**, *318*, 1469–1472. [CrossRef] [PubMed]

68. Jia, G.; Fu, Y.; Zhao, X.; Dai, Q.; Zheng, G.; Yang, Y.; Yi, C.; Lindahl, T.; Pan, T.; Yang, Y.G.; et al. 6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **2011**, *7*, 885–887. [CrossRef] [PubMed]
69. Zheng, G.; Dahl, J.A.; Niu, Y.; Fedorcsak, P.; Huang, C.M.; Li, C.J.; Vagbo, C.B.; Shi, Y.; Wang, W.-L.; Song, S.-H.; et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell.* **2013**, *49*, 18–29. [CrossRef]
70. Feng, C.; Liu, Y.; Wang, G.; Deng, Z.; Zhang, Q.; Wu, W.; Tong, Y.; Cheng, C.; Chen, Z. Crystal structures of the human RNA demethylase Alkbh5 reveal basis for substrate recognition. *J. Biol. Chem.* **2014**, *289*, 11571–11583. [CrossRef] [PubMed]
71. Zhang, X.; Wei, L.H.; Wang, Y.; Xiao, Y.; Liu, J.; Zhang, W.; Yan, N.; Amu, G.; Tang, X.; Zhang, L.; et al. Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 2919–2924. [CrossRef] [PubMed]
72. Bayoumi, M.; Munir, M. Structural insights into m^6A-erasers: A step toward understanding molecule specificity and potential antiviral targeting. *Front. Cell Dev. Biol.* **2021**, *8*, 58710. [CrossRef]
73. McTaggart, J.S.; Lee, S.; Iberl, M.; Church, C.; Cox, R.D.; Ashcroft, F.M. FTO is expressed in neurones throughout the brain and Its expression Is unaltered by fasting. *PLoS ONE* **2011**, *6*, e27968.
74. Zou, S.; Toh, J.D.W.; Wong, K.H.Q.; Gao, Y.G.; Hong, W.; Woon, E.C.Y. N6-Methyladenosine: A conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. *Sci. Rep.* **2016**, *6*, 25677. [CrossRef] [PubMed]
75. Liao, S.; Sun, H.; Xu, C. YTH Domain: A Family of N^6 -methyladenosine (m^6A) Readers. *Genom. Proteom. Bioinform.* **2018**, *16*, 99–107. [CrossRef] [PubMed]
76. Xiao, W.; Adhikari, S.; Dahal, U.; Chen, Y.S.; Hao, Y.J.; Sun, B.F.; Sun, H.Y.; Li, A.; Ping, X.L.; Lai, W.-Y.; et al. Nuclear m^6A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* **2016**, *61*, 925. [CrossRef]
77. Li, A.; Chen, Y.-S.; Ping, X.-L.; Yang, X.; Xiao, W.; Yang, Y.; Sun, H.-Y.; Zhu, Q.; Baidya, P.; Wang, X.; et al. Cytoplasmic m^6A reader YTHDF3 promotes mRNA translation. *Cell Res.* **2017**, *27*, 444–447. [CrossRef] [PubMed]
78. Roundtree, I.A.; Luo, G.Z.; Zhang, Z.; Wang, X.; Zhou, T.; Cui, Y.; Sha, J.; Huang, X.; Guerrero, L.; Xie, P.; et al. YTHDC1 mediates nuclear export of N^6-methyladenosine methylated mRNAs. *eLife* **2017**, *6*, e3131. [CrossRef] [PubMed]
79. Dai, X.; Wang, T.; Gonzalez, G.; Wang, Y. Identification of YTH domain-containing proteins as the readers for *N*1-methyladenosine in RNA. *Anal. Chem.* **2018**, *90*, 6380–6384. [CrossRef]
80. Mao, Y.; Dong, L.; Liu, X.M.; Guo, J.; Ma, H.; Shen, B.; Qian, S.B. m^6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat. Commun.* **2019**, *10*, 5332. [CrossRef]
81. Nesterova, T.B.; Wei, G.; Coker, H.; Pintacuda, G.; Bowness, J.S.; Zhang, T.; Almeida, M.; Bloechl, B.; Moindrot, B.; Carter, E.J.; et al. Xist RNA, the master regulator of X chromosome inactivation, acts in *cis* to induce chromosome-wide silencing. *Nat. Commun.* **2019**, *10*, 3129. [CrossRef]
82. Meyer, K.D.; Patil, D.P.; Zhou, J.; Zinoviev, A.; Skabkin, M.A.; Elemento, O.; Pestova, T.V.; Qian, S.-B.; Jaffrey, S.R. 5′ UTR m^6A promotes cap-independent translation. *Cell* **2015**, *164*, 999–1010. [CrossRef]
83. Patil, D.P.; Pickering, B.F.; Jaffrey, S.R. Reading m ^{6} A in the transcriptome: M ^{6} A-binding proteins. *Trends Cell Biol.* **2018**, *28*, 113–127. [CrossRef] [PubMed]
84. Liu, T.; Wei, Q.; Jin, J.; Luo, Q.; Liu, Y.; Yang, Y.; Cheng, C.; Li, L.; Pi, J. The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. *Nucleic Acids Res.* **2020**, *48*, 3816–3831. [CrossRef] [PubMed]
85. Alarcón, C.R.; Goodarzi, H.; Lee, H.; Liu, X.; Tavazoie, S.; Tavazoie, S.F. HNRNPA2B1 is a mediator of m^6A-dependent nuclear RNA processing events. *Cell* **2015**, *162*, 1299–1308. [CrossRef]
86. Tong, J.; Flavell, R.A.; Li, H.B. RNA m ^{6} A modification and its function in diseases. *Front. Med.* **2018**, *12*, 481–489. [CrossRef] [PubMed]
87. Liu, N.; Dai, Q.; Zheng, G.; He, C.; Parisien, M.; Pan, T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **2015**, *518*, 560–564. [CrossRef] [PubMed]
88. Huang, H.; Weng, H.; Sun, W.; Qin, X.; Shi, H.; Wu, H.; Zhao, B.S.; Mesquita, A.; Liu, C.; Yuan, C.L.; et al. Recognition of RNA N6 methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* **2018**, *20*, 285–295. [CrossRef]
89. Bell, J.L.; Wächter, K.; Mühleck, B.; Pazaitis, N.; Köhn, M.; Lederer, M.; Hüttelmaier, S. Insulin like growth factor 2 mRNA-binding proteins (IGF2BPs): Post-transcriptional drivers of cancer progression? *Cell Mol. Life Sci.* **2013**, *70*, 2657–2675. [CrossRef] [PubMed]
90. Motorin, Y.; Luko, F.; Helm, M. 5-methylcytosine in RNA: Detection, enzymatic formation and biological functions. *Nucleic Acids Res.* **2010**, *38*, 1415–1430. [CrossRef]
91. Bohnsack, K.E.; Höbartner, C.; Markus, T.; Bohnsack, M.T. Eukaryotic 5-methylcytosine (m5C) RNA methyltransferases: Mechanisms, cellular functions, and links to disease. *Genes* **2019**, *10*, 102. [CrossRef]
92. Jeltsch, A.; Nellen, W.; Lyko, F. Two substrates are better than one: Dual specificities for Dnmt2 methyltransferases. *Trends Biol. Sci.* **2006**, *31*, 306–308. [CrossRef]
93. Jeltsch, A.; Ehrenhofer-Murray, A.; Jurkowski, T.P.; Lyko, F.; Reuter, G.; Ankri, S.; Nellen, W.; Schaefer, M.; Mark Helm, M. Mechanism and biological role of Dnmt2 in nucleic acid methylation. *RNA Biol.* **2017**, *14*, 1108–1123. [CrossRef]
94. Chi, L.; Delgado-Olguín, P. Expression of NOL1/NOP2/sun domain (Nsun) RNA methyl- transferase family genes in early mouse embryogenesis. *Gene Expres. Patterns* **2013**, *13*, 319–327. [CrossRef] [PubMed]
95. Yang, X.; Yang, Y.; Sun, B.F.; Chen, Y.-S.; Xu, J.-W.; Lai, W.-Y.; Li, A.; Wang, X.; Bhattarai, D.P.; Xiao, W.; et al. 5-methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m^5C reader. *Cell Res.* **2017**, *27*, 606–625. [CrossRef]

96. Wu, H.; Zhang, Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.* **2011**, *25*, 2436–2452. [CrossRef]
97. Yang, J.; Bashkenova, N.; Zang, R.; Huang, X.; Wang, J. The roles of TET family proteins in development and stem cells. *Development* **2020**, *147*, dev183129. [CrossRef] [PubMed]
98. Shen, H.; Ontiveros, R.J.; Owen, M.C.; Liu, M.Y.; Ghanty, U.; Kohli, R.M.; Liu, K.F. TET-mediated 5-methylcytosine oxidation in tRNA promotes translation. *J. Biol. Chem.* **2021**, *296*, 100087. [CrossRef] [PubMed]
99. Deng, W.; Zhang, Y.; Cai, J.; Zhang, J.; Liu, X.; Yin, J.; Bai, Z.; Yao, H.; Zhang, Z. LncRNA-ANRIL promotes gastric cancer progression by enhancing NF-kB signaling. *Exp. Biol. Med.* **2019**, *244*, 953–959. [CrossRef] [PubMed]
100. Gray, M.C.; Charette, M.W. Pseudouridine in RNA: What, where, how, and why. *IUBMB Life* **2000**, *49*, 341–351. [CrossRef] [PubMed]
101. Kierzek, E.; Malgowska, M.; Lisowiec, J.; Turner, D.H.; Gdaniec, Z.; Kierzek, R. The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res.* **2014**, *42*, 3492–3501. [CrossRef] [PubMed]
102. Deb, I.; Popenda, Ł.; Sarzyńska, J.; Małgowska, M.; Lahiri, A.; Gdaniec, Z.; Kierzek, R. Computational and NMR studies of RNA duplexes with an internal pseudouridine-adenosine base pair. *Sci. Rep.* **2019**, *9*, 16278. [CrossRef]
103. Borchardt, E.K.; Martinez, N.M.; Gilbert, W.V. Regulation and Function of RNA Pseudouridylation in Human Cells. *Annu. Rev. Genet.* **2020**, *54*, 309–336. [CrossRef]
104. Zhao, Y.; Dunker, V.; Yu, Y.-T.; Karijolich, J. The Role of Noncoding RNA Pseudouridylation in Nuclear Gene Expression Events. *Front. Bioeng. Biotechnol.* **2018**, *6*, 8. [CrossRef]
105. Ni, J.; Tien, A.L.; Fournier, M.J. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* **1997**, *89*, 565–573. [CrossRef]
106. Caton, E.A.; Kelly, E.K.; Kamalampeta, R.; Kothe, U. Efficient RNA pseudouridylation by eukaryotic H/ACA ribonucleoproteins requires high affinity binding and correct positioning of guide RNA. *Nucleic Acids Res.* **2018**, *46*, 905–916. [CrossRef]
107. Yu, Y.T.; Meier, U.T. RNA-guided isomerization of uridine to pseudouridine—Pseudouridylation. *RNA Biol.* **2014**, *11*, 1483–1494. [CrossRef]
108. Li, X.; Ye, K. Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* **2006**, *443*, 302–307. [CrossRef] [PubMed]
109. Li, X.; Ma, S.; Yi, C. Pseudouridine: The fifth RNA nucleotide with renewed interests. *Curr. Opin. Chem. Biol.* **2016**, *33*, 108–116. [CrossRef] [PubMed]
110. Zhao, B.S.; He, C. Pseudouridine in a new era of RNA modifications. *Cell Res.* **2015**, *25*, 153–154. [CrossRef] [PubMed]
111. Hamma, T.; Feree D'Amare, A. Pseudouridine synthases. *Chem. Biol.* **2006**, *13*, 1125–1135. [CrossRef]
112. Tomaselli, S.; Bonamassa, B.; Alisi, A.; Nobili, V.; Locatelli, F.; Gallo, A. ADAR enzyme and miRNA story: A nucleotide that can make the difference. *Int. J. Mol. Sci.* **2013**, *14*, 22796–22816. [CrossRef]
113. Farajollahi, S.; Maas, S. Molecular diversity through RNA editing: A balancing act. *Trends Genet.* **2010**, *26*, 221–230. [CrossRef] [PubMed]
114. Torsin, L.L.; Petrescu, G.E.D.; Sabo, A.A.; Chen, B.; Brehar, F.M.; Dragomir, M.P.; Calin, G.A. Editing and chemical modifications on non-coding RNAs in cancer: A New Tale with Clinical Significance. *Int. J. Mol. Sci.* **2021**, *22*, 581. [CrossRef] [PubMed]
115. Silvestris, D.A.; Scopa, C.; Hanchi, S.; Locatelli, F.; Gallo, A. De Novo A-to-I RNA Editing Discovery in lncRNA. *Cancers* **2020**, *12*, 2959. [CrossRef] [PubMed]
116. Dominissini, D.; Moshitch-Moshkovitz, S.; Schwartz, S.; Salmon-Divon, M.; Ungar, L.; Osenberg, S.; Cesarkas, K.; Jacob-Hirsch, J.; Amariglio, N.; Kupiec, M.; et al. Topology of the human and mouse m6A RNA methylomes revealed by m^{6}A-seq. *Nature* **2012**, *485*, 201–206. [CrossRef] [PubMed]
117. Liu, H.; Begik, O.; Lucas, M.C.; Ramirez, J.M.; Mason, C.E.; Wiener, D.; Schwartz, S.; Mattick, J.S.; Smith, M.A.; Novoa, E.M. Accurate detection of m^6 RNA modifications in native RNA sequence. *Nat. Commun.* **2019**, *10*, 4079. [CrossRef]
118. Davis, F.F.; Allen, F.W. Ribonucleic acids from yeast which contain a fifth nucleotide. *J. Biol. Chem.* **1957**, *227*, 907–915. [CrossRef]
119. Matsuzawa, S.; Wakata, Y.; Ebi, F.; Isobe, M.; Kurosawa, N. Development and validation of monoclonal antibodies against N6-methyladenosine for the detection of RNA modifications. *PLoS ONE* **2019**, *14*, e0223197. [CrossRef]
120. Muthmann, N.; Katja Hartstock, K.; Rentmeister, A. Chemo-enzymatic treatment of RNA to facilitate analyses. *WIREs RNA* **2020**, *11*, e1561. [CrossRef]
121. Bochner, B.R.; Bruce, N.; Ames, B.N. Complete analysis of cellular nucleotides by two-dimensional thin layer chromatography. *J. Biol. Chem.* **1982**, *257*, 9759–9769. [CrossRef]
122. Bodi, Z.; Fray, R.G. Detection and Quantification of N ^{6}-Methyladenosine in Messenger RNA by TLC. *Methods Mol. Biol.* **2017**, *1562*, 79–87. [PubMed]
123. Liu, N.; Parisien, M.; Dai, Q.; Zheng, G.; He, C.; Pan, T. Probing N^6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* **2013**, *19*, 1848–1856. [CrossRef] [PubMed]
124. Liu, N.; Pan, T. Probing RNA modification status at single-nucleotide resolution in total RNA. *Methods Enzymol.* **2015**, *560*, 149–159. [PubMed]
125. Ivleva, V.B.; Yu, Y.Q.; Gilar, M. Ultra-performance liquid chromatography/tandem mass spectrometry (UPLC/MS/MS) and UPLC/MS(E) analysis of RNA oligonucleotides. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 2631–2640. [CrossRef]
126. Zhang, C.; Fu, J.; Zhou, Y. A review in research progress concerning m^{6}A methylation and immunoregulation. *Front. Immunol.* **2019**, *10*, 922. [CrossRef]

127. Sugimoto, Y.; König, J.; Hussain, S.; Zupan, B.; Curk, T.; Frye, M.; Ule, J. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol.* **2012**, *13*, R67. [CrossRef] [PubMed]
128. Linder, B.; Grozhik, A.V.; Olarerin-George, A.O.; Meydan, C.; Mason, C.E.; Jaffrey, S.R. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* **2015**, *12*, 767–772. [CrossRef] [PubMed]
129. Hawley, B.R.; Jaffrey, S.R. Transcriptome-Wide Mapping of m^6 A and m^6 Am at Single-Nucleotide Resolution Using miCLIP. *Curr. Protoc. Mol. Biol.* **2019**, *126*, e88. [CrossRef]
130. Werner, S.; Galliot, A.; Pichot, F.; Kemmer, T.; Marchand, V.; Sednev, M.V.; Lence, T.; Roignant, J.Y.; König, J.; Höbartner, C.; et al. NOseq: Amplicon sequencing evaluation method for RNA m^6A sites after chemical deamination. *Nucleic Acids Res.* **2021**, *49*, e23. [CrossRef]
131. Clark, S.J.; Frommer, M. Bisulfite genomic sequencing of methylated cytosines. In *Laboratory Methods for the Detection of Mutations and Polymorphisms in DNA*; Taylor, G., Ed.; CRC Press: New York, NY, USA, 1997; pp. 151–162.
132. Schaefer, M.; Pollex, T.; Hanna, K.; Lyko, F. RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res.* **2009**, *37*, e12. [CrossRef]
133. Xue, C.; Yalei Zha, Y.; Li, L. Advances in RNA cytosine-5 methylation: Detection, regulatory mechanisms, biological functions and links to cancer. *Biomark. Res.* **2020**, *8*, 43. [CrossRef]
134. Chen, Y.S.; Ma, H.L.; Yang, Y.; Lai, W.Y.; Sun, B.F.; Yang, Y.G. 5-Methylcytosine analysis by RNA-BisSeq. *Methods Mol. Biol.* **2019**, *1870*, 237–248. [PubMed]
135. Amort, T.; Soulière, M.F.; Wille, A.; Jia, X.Y.; Fiegl, H.; Wörle, H.; Micura, R.; Lusser, A. Long non-coding RNAs as targets for cytosine methylation. *RNA Biol.* **2013**, *10*, 1003–1008. [CrossRef] [PubMed]
136. Chen, Y.S.; Yang, W.L.; Zhao, Y.L.; Yang, Y.G. Dynamic transcriptomic m ^{5} C and its regulatory role in RNA processing. *Wiley Interdiscip. Rev. RNA* **2021**, e1639.
137. Metodiev, M.D.; Spåhr, H.; Loguercio Polosa, P.; Meharg, C.; Becker, C.; Altmueller, J.; Habermann, B.; Larsson, N.G.; Ruzzenente, B. NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. *PLoS Genet.* **2014**, *10*, e1004110. [CrossRef]
138. Khoddami, V.; Cairns, B. Transcriptome—Wide target profiling of RNA cytosine methyltransferases using the mechanism-based enrichment procedure Aza-IP. *Nat. Protoc.* **2014**, *9*, 337–361. [CrossRef]
139. Uddin, M.B.; Zhishan Wang, Z.; Yang, C. Dysregulations of functional RNA modifications in cancer, cancer stemness and cancer therapeutics. *Theranostic* **2020**, *10*, 3164–3189. [CrossRef] [PubMed]
140. Ofengand, J.; Bakin, A.; Wrzesinski, J.; Nurse, K.; Lane, B.G. The pseudouridine residues of ribosomal RNA. *Biochem. Cell Biol.* **1995**, *73*, 915–924. [CrossRef] [PubMed]
141. Ofengand, J.; Del Campo, M.; Kaya, Y. Mapping pseudouridines in RNA molecules. *Methods* **2001**, *25*, 365–373. [CrossRef]
142. Li, X.; Zhu, P.; Ma, S.; Song, J.; Bai, J.; Sun, F.; Yi, C. Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* **2015**, *11*, 592–597. [CrossRef] [PubMed]
143. Zhang, W.; Eckwahl, M.J.; Zhou, K.I.; Pan, T. Sensitive and quantitative probing of pseudouridine modification in mRNA and long noncoding RNA. *RNA* **2019**, *25*, 1218–1225. [CrossRef] [PubMed]
144. Shafik, A.; Schumann, U.; Evers, M.; Sibbritt, T.; Preiss, T. The emerging epitranscriptomics of long noncoding RNAs. *Biochim. Biophys. Acta* **2016**, *1859*, 59–70. [CrossRef] [PubMed]
145. Lo Giudice, C.; Silvestris, D.A.; Roth, S.H.; Eisenberg, E.; Pesole, G.; Gallo, A.; Picardi, E. Quantifying RNA editing in deep transcriptome datasets. *Front. Genet.* **2020**, *11*, 194. [CrossRef]
146. Malik, T.N.; Cartailler, J.P.; Emeson, R.B. Quantitative analysis of adenosine-to-inosine RNA editing. *Methods Mol. Biol.* **2021**, *2181*, 97–111. [PubMed]
147. Zerbino, D.R.; Frankish, A.; Flicek, P. Progress, challenges, and surprises in annotating the human genome. *Annu. Rev. Genom. Hum. Genet.* **2020**, *21*, 55–79. [CrossRef]
148. Wang, Y.; Song, X.; Xu, T. Identification and analysis of RNA editing events in ovarian serous cystadenoma using RNA-seq data. *Curr. Gene Ther.* **2021**, *21*, 5. [CrossRef]
149. Ohlson, J.; Ohman, M. A method for finding sites of selective adenosine deamination. *Methods Enzymol.* **2007**, *424*, 289–300.
150. Morse, D.P.; Bass, B.L. Detection of inosine in messenger RNA by inosine-specific cleavage. *Biochemistry* **1997**, *36*, 8429–8434. [CrossRef]
151. Cattenoz, P.B.; Taft, R.J.; Westhof, D.; Mattick, J.S. Transcriptome-wide identification of A—I RNA editing sites by inosine specific cleavage. *RNA* **2013**, *19*, 257–270. [CrossRef]
152. Sakurai, M.; Yano, T.; Kawabata, H.; Ueda, H.; Suzuki, T. Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat. Chem. Biol.* **2010**, *6*, 733–740. [CrossRef]
153. Suzuki, T.; Ueda, H.; Okada, S.; Sakurai, M. Transcriptome-wide identification of adenosine-to-inosine editing using the ICE-seq method. *Nat. Protoc.* **2015**, *10*, 715–732. [CrossRef]
154. St Laurent, G.; Tackett, M.R.; Nechkin, S.; Shtokalo, D.; Antonets, D.; Savva, Y.A.; Maloney, R.; Kapranov, P.; Lawrence, C.E.; Reenan, R.A. Genome-Wide Analysis of A-to-I RNA Editing by Single-Molecule Sequencing in Drosophila. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1333–1339. [CrossRef] [PubMed]
155. Xu, L.; Seki, M. Recent advances in the detection of base modifications using the Nanopore sequencer. *J. Hum. Genet.* **2020**, *65*, 25–33. [CrossRef] [PubMed]

156. Motorin, Y.; Marchand, V. Analysis of RNA modifications by second- and third-generation deep sequencing: 2020 Update. *Genes* **2021**, *12*, 278. [CrossRef] [PubMed]
157. Furlan, M.; Tanaka, I.; Leonardi, T.; de Pretis, S.; Pelizzola, M. Direct RNA sequencing for the study of synthesis, processing, and degradation of modified transcripts. *Front. Genet.* **2020**, *11*, 394. [CrossRef] [PubMed]
158. Eisenstein, M. Oxford Nanopore announcement sets sequencing sector abuz. *Nat. Biotech.* **2012**, *30*, 295–296. [CrossRef]
159. Anreiter, I.; Mir, Q.; Simpson, J.T.; Janga, S.C.; Soller, M. New twists in detecting mRNA modification dynamics. *Trends Biotechnol.* **2021**, *39*, 72–89. [CrossRef]
160. Jia, L.; Chen, J.; Liu, H.; Fan, W.; Wang, D.; Li, J.; Liu, D. Potential m6A and m5C methylations within the genome of a chinese african swine fever virus strain. *J. Virol. Sin.* **2020**, *36*, 321–324. [CrossRef]
161. Jenjaroenpun, P.; Wongsurawat, T.; Wadley, T.D.; Wassenaar, T.M.; Liu, J.; Dai, Q.; Wanchai, V.; Akel, N.S.; Jamshidi-Parsian, A.; Franco, A.T.; et al. Decoding the epitranscriptional landscape from native RNA sequences. *Nucleic Acids Res.* **2021**, *49*, e7. [CrossRef]
162. Szcześniak, M.W.; Makałowska, I. lncRNA-RNA Interactions across the Human Transcriptome. *PLoS ONE* **2016**, *11*, e0150353. [CrossRef]
163. Cabili, M.N.; Dunagin, M.C.; McClanahan, P.D.; Biaesch, A.; Padovan-Merhar, O.; Regev, A.; Rinn, J.L.; Raj, A. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol.* **2015**, *16*, 20. [CrossRef]
164. Perry, R.B.T.; Ulitsky, I. The functions of long noncoding RNAs in development and stem cells. *Development* **2016**, *143*, 3882–3894. [CrossRef]
165. Kopp, F.; Mendell, J.T. Functional classification and experimental dissection of long noncoding RNAs. *Cell* **2018**, *172*, 393–407. [CrossRef] [PubMed]
166. Zhao, Y.; Teng, H.; Yao, F.; Yap, S.; Sun, Y.; Ma, L. Challenges and Strategies in Ascribing Functions to Long Noncoding RNAs. *Cancers* **2020**, *12*, 1458. [CrossRef]
167. Chi, J.S.; Li, J.Z.; Jia, J.J.; Zhang, T.; Liu, X.M.; Yi, L. Long non-coding RNA ANRIL in gene regulation and its duality in atherosclerosis. *Curr. Med. Sci.* **2017**, *37*, 816–822. [CrossRef]
168. Loda, A.; Heard, E. *Xist* RNA in action: Past, present, and future. *PLoS Genet.* **2019**, *15*, e1008333. [CrossRef] [PubMed]
169. Cerase, A.; Pintacuda, G.; Tattermusch, A.; Avner, P. Xist localization and function: New insights from multiple levels. *Genome Biol.* **2015**, *16*, 166. [CrossRef] [PubMed]
170. Trotman, J.B.; Braceros, K.C.A.; Cherney, R.E.; Murvin, M.M.; Calabrese, J.M. The control of polycomb repressive complexes by long noncoding RNAs. *Wiley Interdiscip. Rev. RNA* **2021**, e1657.
171. LNCipedia/NONCODE. Available online: http://bioinfo.life.hust.edu.cn/LNCediting/ (accessed on 1 June 2021).
172. Zhou, K.I.; Parisien, M.; Dai, Q.; Liu, N.; Diatchenko, L.; Sachleben, J.R.; Pan, T. N(6)-methyladenosine modification in a long loncoding RNA hairpin predisposes its conformation to protein binding. *J. Mol. Biol.* **2016**, *428*, 822–833. [CrossRef]
173. Liu, N.; Zhou, K.I.; Parisien, M.; Dai, Q.; Diatchenko, L.; Pan, T. N^6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* **2017**, *45*, 6051–6063. [CrossRef] [PubMed]
174. Wang, X.; Liu, C.; Zhang, S.; Yan, H.; Zhang, L.; Jiang, A.; Liu, Y.; Feng, Y.; Li, D.; Guo, Y.; et al. N6-methyladenosine modification of MALAT1 promotes metastasis via reshaping nuclear speckles. *Dev. Cell.* **2021**, *56*, 702–715. [CrossRef]
175. Bhan, A.; Mandal, S.S. LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. *Biochim. Biophys. Acta* **2015**, *1856*, 151–164. [CrossRef]
176. Porman, A.M.; Roberts, J.R.; Chrupcala, M.; Kennedy, M.; Williams, M.M.; Richer, J.K.; Johnson, A.M. A single N6-methyladenosine site in lncRNA HOTAIR regulates its function in breast cancer cells. *BioRaiv* **2020**. [CrossRef]
177. Yang, D.; Qiao, J.; Wang, G.; Lan, Y.; Li, G.; Guo, X.; Xi, J.; Ye, D.; Zhu, S.; Chen, W.; et al. N6-Methyladenosine modification of lincRNA 1281 is critically required for mESC differentiation potential. *Nucleic Acids Res.* **2018**, *46*, 3906–3920. [CrossRef]
178. Gabory, A.; Dandolo, L.; Yoshimizu, T.; Ripoche, M.A. The H19 gene: Regulation and function of a non-coding RNA. *Cytogenet. Genome Res.* **2006**, *113*, 188–193. [CrossRef]
179. Su, Y.; Xu, R.; Zhang, R.; Qu, Y.; Zuo, W.; Ji, Z.; Geng, H.; Pan, M.; Ma, G. N^6-methyladenosine methyltransferase plays a role in hypoxic preconditioning partially through the interaction with lncRNA H19. *Acta Biochim. Biophys. Sin.* **2020**, *52*, 1306–1315. [CrossRef] [PubMed]
180. Sun, Z.; Xue, S.; Zhang, M.; Xu, H.; Hu, X.; Chen, S.; Liu, Y.; Guo, M.; Cui, H. Aberrant NSUN2-mediated m5 C modification of H19 lncRNA is associated with poor differentiation of hepatocellular carcinoma. *Oncogene* **2020**, *39*, 6906–6919. [CrossRef] [PubMed]
181. Lanz, R.B.; McKenna, N.J.; Onate, S.A.; Albrecht, U.; Wong, J.; Tsai, S.Y.; Tsai, M.J.; O'Malley, B.W. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell* **1999**, *97*, 17–27. [CrossRef]
182. Sheng, L.; Ye, L.; Zhang, D.; Cawthorn, W.P.; Xu, B. New Insights Into the Long Non-coding RNA SRA: Physiological Functions and Mechanisms of Action. *Front. Med.* **2018**, *5*, 244. [CrossRef] [PubMed]
183. Jin, K.; Wang, S.; Zhang, Y.; Xia, M.; Mo, Y.; Li, X.; Li, G.; Zeng, Z.; Xiong, W.; He, Y. Long non-coding RNA PVT1 interacts with MYC and its downstream molecules to synergistically promote tumorigenesis. *Cell Mol. Life Sci.* **2019**, *76*, 4275–4289. [CrossRef]
184. Lee, J.; Wu, Y.; Harada, B.T.; Li, Y.; Zhao, J.; He, C.; Ma, Y.; Wu, X. N^6-methyladenosine modification of lncRNA Pvt1 governs epidermal stemness. *EMBO J.* **2021**, *40*, e106276. [CrossRef]

185. Xu, Z.; Ji, G.; Cui, Y.; Cui, X. The impacts of non-coding RNAs and N6-methyladenosine on cancer: Past, present, and future. *Curr. Cancer Drug Targets* **2021**. [CrossRef] [PubMed]
186. He, Y.; Shi, Q.; Zhang, Y.; Yuan, X.; Yu, Z. Transcriptome-wide 5-methylcytosine functional profiling of long non-coding RNA in hepatocellular carcinoma. *Cancer Manag. Res.* **2020**, *12*, 6877–6885. [CrossRef]
187. Chen, M.; Wong, C.-M. The emerging roles of N6-methyladenosine (m6A) deregulation in liver carcinogenesis. *Mol. Cancer* **2020**, *19*, 44. [CrossRef] [PubMed]
188. Yang, X.; Liu, M.; Li, M.; Zhang, S.; Hiju, H.; Sun, J.; Mao, Z.; Zheng, M.; Feng, B. Epigenetic modulations of noncoding RNA: A novel dimension of Cancer biology. *Mol. Cancer* **2020**, *19*, 64. [CrossRef] [PubMed]

Życiorys naukowy

1. EDUKACJA

2019 - 2023	Studia III stopnia Jednostka naukowa: **Instytut Chemii Bioorganicznej PAN** Specjalizacja: **Profil biologiczny** Tryb: **Środowiskowe Studium Doktoranckie ICHB PAN**
2013 - 2018	Studia I i II stopnia Jednostka naukowa: **Uniwersytet Przyrodniczy w Poznaniu** Kierunek: **Biotechnologia** Specjalizacja: **Genetyka Zwierząt** Poziom wykształcenia: **magister inżynier**

2. PRZEBIEG KARIERY ZAWODOWEJ

06.2018 – 07.2018	Stanowisko: **Stażysta naukowy** Miejsce pracy: **LUKE (Natural Resources Institute Finland) Jokioinen, Finlandia**
03.2017 – 04.2018	Stanowisko: **Magistrant** Miejsce pracy: **Wydział Żywienia Zwierząt i Gospodarki Paszowej, Uniwersytet Przyrodniczy w Poznaniu** **Poznań**
02.2017 – 02.2017	Stanowisko: **Stażysta** Miejsce pracy: **Poznański Park Naukowo-Technologiczny Fundacji UAM Poznań**
10.2016 – 02.2017	Stanowisko: **Inżynierant** Miejsce pracy: **Zakład Hodowli Koni, Uniwersytet Przyrodniczy w Poznaniu Poznań**
06.2015 – 06.2015	Stanowisko: **Praktykant** Miejsce pracy: **Zakład Farmaceutyczny Colfarm S.A. Mielec**

3. PUBLIKACJE

tRNA-derived fragments from the Sus scrofa tissues provide evidence of their conserved role in mammalian development

Marek Kazimierczyk, Agata Jędroszkowiak, Dorota Kowalczykiewicz, Maciej Szymański, Barbara Imiołczyk, Jerzy Ciesiołka, Jan Wrzesiński

Biochem Biophys Res Commun.2019 Dec 10;520(3):514-519. doi: 10.1016/j.bbrc.2019.10.062.

Characteristics of Transfer RNA-Derived Fragments Expressed during Human Renal Cell Development: The Role of Dicer in tRF Biogenesis

Marek Kazimierczyk, Marta Wojnicka, Ewa Biała, Paulina Żydowicz-Machtel, Barbara Imiołczyk, Tomasz Ostrowski, Anna Kurzyńska-Kokorniak, Jan Wrzesinski

Int J Mol Sci. 2022 Mar 26;23(7):3644. doi: 10.3390/ijms23073644.

Human Long Noncoding RNA Interactome: Detection, Characterization and Function

Marek Kazimierczyk, Marta K. Kasprowicz, Marta E. Kasprzyk, and Jan Wrzesinski

Int J Mol Sci. 2020 Feb; 21(3): 1027. Published online 2020 Feb 4. doi: 10.3390/ijms21031027

Long Non-Coding RNA Epigenetics

Marek Kazimierczyk, Jan Wrzesinski

Int J Mol Sci. 2021 Jun 7;22(11):6166. doi: 10.3390/ijms22116166.

4. KONFERENCJE

,,The role of Piwi/piRNA in function of renal cell.''

The Non-Coding Genome conference in Heidelberg 2019 poster no. 139

Oświadczenia autora

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

Marek Kazimierczyk

Zakład Biochemii RNA

Instytut Chemii Bioorganicznej

Polskiej Akademii Nauk

61-704 Poznań

Ul. Noskowskiego 12/14

Oświadczenie doktoranta dotyczące jego udziału w powstawaniu pracy naukowej wchodzącej w skład rozprawy doktorskiej.

1. Kazimierczyk, M., Jędroszkowiak, A., Kowalczykiewicz, D., Szymanski, M., Imiołczyk, B., Ciesiołka, J., Wrzesinski J (2019) tRNA-derived fragments from the Sus scrofa tissues provide evidence of their conserved role in mammalian development. *Biochem. Biophys. Res. Commun.* **520**, 514-519. DOI: 10.1016/j.bbrc.2019.10.062.

IF = 2.985, MNiSW = 100

Oświadczam, że mój udział w tworzeniu niniejszej pracy polegał na:

a) Końcowej analizie downstream wyników sekwencjonowania
b) Przygotowaniu opisu oraz analizie ilościowej oddziaływania białka Piwil4 z wybranymi sncRNA
c) Finalyme opracowaniu danych ilościowych z wykonanych eksperymentów
d) Techniczna oraz merytoryczna praca nad manuskryptem publikacji

Autor

.....
Podpis

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

2. Kazimierczyk, M., Kasprowicz, M.K., Kasprzyk, K.E., Wrzesinski, J. (2020) Human Long Noncoding RNA Interactome: Detection, Characterization and Function. *Int. J. Mol. Sci.* **21**, 1027; DOI:10.3390/ijms21031027.

IF = 5.924, MNiSW = 140

Oświadczam, że mój udział w tworzeniu niniejszej pracy polegał na:

a) Przygotowaniu merytorycznym oraz technicznym wszystkich figur w publikacji
b) Propozycji i realizacji merytorycznej treści rozdziału „The LncRNA Interactome: The Nucleic Acid Story"
c) Technicznej oraz merytorycznej pracy nad manuskryptem publikacji

3. Kazimierczyk, M., Wrzesinski, J. (2021) Long Non-Coding RNA Epigenetics. *Int. J. Mol. Sci.* ***22***(11), 6166; DOI: 10.3390/ijms22116166

IF = 6.208, MNiSW = 140

Oświadczam, że mój udział w tworzeniu niniejszej pracy polegał na:

a) Przygotowaniu merytorycznym oraz technicznym wszystkich figur w publikacji
b) Przygotowaniu pierwotnej wersji manuskryptu oraz po szerokiej konsultacji finalnej wersji manuskryptu

Autor

.....
Podpis

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

4. Kazimierczyk, M., Wojnicka, M., Biała, E., Żydowicz-Machtel, P., Imiołczyk, B., Ostrowski, T., Kurzyńska-Kokorniak, A., Wrzesinski, J. (2022) Characteristics of Transfer RNA-Derived Fragments Expressed during Human Renal Cell Development: The Role of Dicer in tRF Biogenesis. *Int. J. Mol. Sci.* **23**(7), 3644; DOI:10.3390/ijms23073644.

IF = 5.6, MNiSW = 140

Oświadczam, że mój udział w tworzeniu niniejszej pracy polegał na:

a) Opracowaniu globalnych wyników sekwencjonowania RNA
b) Wstępnej analizie downstream wyników sekwencjonowania RNA, wybraniu cząsteczek do dalszej analizy eksperymentalnej
c) Prowadzeniu hodowli komórkowych hESC, HEK293, HK-2, A498 oraz izolacji materiału do analiz
d) Nadzorze nad prowadzonymi eksperymentami Northern blot oraz analizie wyników
e) Analizie baz danych w celu odnalezienia potencjalnych białek regulowanych przez tRFy.
f) Opracowaniu wszystkich grafik w publikacji
g) Udziale merytorycznym oraz technicznym w tworzeniu manuskryptu

Autor

.....
Podpis

Oświadczenia autora korespondencyjnego

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

Jan Wrzesiński

Zakład Biochemii RNA

Instytut Chemii Bioorganicznej

Polskiej Akademii Nauk

61-704 Poznań

Ul. Noskowskiego 12/14

Oświadczenie autora do korespondencji dotyczące udziału doktoranta w powstawaniu pracy naukowej wchodzącej w skład rozprawy doktorskiej.

Udział mgr inż. Marka Kazimierczyka w publikacja będących jego pracą doktorską oceniam bardzo wysoko. Świadczy o tym bycie pierwszym autorem we wszystkich publikacjach.

1. Kazimierczyk, M., Jędroszkowiak, A., Kowalczykiewicz, D., Szymanski, M., Imiołczyk, B., Ciesiołka, J., Wrzesinski J (2019) tRNA-derived fragments from the Sus scrofa tissues provide evidence of their conserved role in mammalian development. *Biochem. Biophys. Res. Commun.* **520**, 514-519. DOI: 10.1016/j.bbrc.2019.10.062.

   IF = 2.985, MNiSW = 100

Doktorant miał udział w następujących zadaniach

a) Analizie danych bioinformatycznych sekwencjonowania
b) Przygotowaniu danych ilościowych wykonanych eksperymentów
c) Opis i ocena ilościowa oddziaływań białka Piwil4 z małymi RNA
d) Udział w przygotowaniu manuskryptu publikacji

Autor korespondencyjny

.....
Podpis

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

2. Kazimierczyk, M., Kasprowicz, M.K., Kasprzyk, K.E., Wrzesinski, J. (2020) Human Long Noncoding RNA Interactome: Detection, Characterization and Function. *Int. J. Mol. Sci.* **21**, 1027; DOI:10.3390/ijms21031027.

   IF = 5.924, MNiSW = 140

   W tej publikacji przeglądowej doktorant

   a) Przygotował pierwotna wersją rozdziału „The LncRNA Interactome: The Nucleic Acid Story"
   b) Zaproponował i opracował szatę graficzną publikacji
   c) Brał udział w przygotowaniu ostatecznej wersji manuskryptu

3. Kazimierczyk, M., Wrzesinski, J. (2021) Long Non-Coding RNA Epigenetics. *Int. J. Mol. Sci.* ***22***(11), 6166; DOI: 10.3390/ijms22116166

   IF = 6.208, MNiSW = 140

   W tej pracy przeglądowej udział mgr Marka Kazimierczyka był podobny jak w poprzedniej.

   a) Doktorant przygotował pierwotną wersję manuskryptu a po dyskusji z autorem korespondencyjnym wersję ostateczną
   b) Zaproponował i opracował szatę graficzną publikacji

Autor korespondencyjny

.....
Podpis

INSTYTUT CHEMII BIOORGANICZNEJ
Polskiej Akademii Nauk

ul. Noskowskiego 12/14, 61-704 Poznań
tel.: centrala 61 852 85 03, sekretariat 61 852 89 19
faks: 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 NIP 777-00-02-062
http://www.ibch.poznan.pl

4. Kazimierczyk, M., Wojnicka, M., Biała, E., Żydowicz-Machtel, P., Imiołczyk, B., Ostrowski, T., Kurzyńska-Kokorniak, A., Wrzesinski, J. (2022) Characteristics of Transfer RNA-Derived Fragments Expressed during Human Renal Cell Development: The Role of Dicer in tRF Biogenesis. *Int. J. Mol. Sci.* **23**(7), 3644; DOI:10.3390/ijms23073644.

   IF = 5.6, MNiSW = 140

   Udział doktoranta w tej publikacji eksperymentalnej był wieloraki.

a) Analiza bioinformatyczna wyników głębokiego sekwencjonowania oraz wyszukanie fragmentów tRNA o najwyższej ekspresji do potwierdzenia eksperymentalnego.
b) Opracowanie protokół hodowli komórkowych (komórki hESC, HEK 293T, HK2, A498)
c) Nadzór na przebiegiem hodowli
d) Izolacja małych RNA do dalszych analiz
e) Wykonanie części eksperymentów Northern blot oraz ich analiza ilościowa.
f) Analiza komputerowa oddziaływań tRF z białkami występującymi w nerkach
g) Udział w przygotowaniu manuskryptu oraz po dyskusji jego wersji finalnej.

Autor korespondencyjny

.....
Podpis