Zbigniew Stanisław Warkocki

Autoreferat

Określenie mechanizmów regulacji podaży i funkcji RNA w komórkach eukariontów przez enzymy działające na koniec 3' RNA

Poznań, 2019-04-25

1. Imię i Nazwisko Zbigniew Warkocki

2. Posiadane dyplomy

2.1. Dyplom doktora nauk o życiu (19 kwietnia 2012r.)

Uniwersytet im. Georga Augusta w Getyndze, Niemcy

Tytuł pracy doktorskiej: "Reconstitution of both steps of yeast pre-mRNA splicing with purified components"

Praca doktorska wykonana w Instytucie im. Maxa Plancka Chemii Biofizycznej w Getyndze

W ramach szkoły doktoranckiej "Gottingen Graduate School for Neurosciences, Biophysics and Molecular Biosciences" zorganizowanej w ramach niemieckiego narodowego centrum doskonałości naukowej

Promotor: Prof. dr hab. Reinhard Lührmann

Zespół promotorski: prof. dr hab. Marina Rodnina, prof. dr hab. Detlef Doenecke, prof. dr Holger Stark, dr Henning Urlaub, dr Caludia Höbartner

Praca i obrona ocenione na *"magna cum laude"*

2.2. Dyplom magistra nauk biologicznych

Uniwersytet im. Adama Mickiewicza w Poznaniu, Wydział Biologii (23 września 2007r.)

Tytuł pracy magisterskiej: "Analiza ekspresji 6 polimeraz RNA zależnych od RNA w Arabidopsis thaliana".

Praca magisterska wykonywana w Instytucie Chemii Bioorganicznej PAN w Poznaniu

Promotor: Prof. dr hab. Marek Figlerowicz

2.3. Dyplom licencjata nauk biologicznych

Uniwersytet im. Adama Mickiewicza w Poznaniu, Wydział Biologii (maj 2005) Tytuł pracy licencjackiej: "Wirusy patogenne roślin – strategie ekspresji genów i metody diagnostyczne".

Promotor: Prof. dr hab. Artur Jarmołowski

3. Informacje o miejscach zatrudnienia/ pracy naukowej

2.2018 – obecnie	Instytut Chemii Bioorganicznej Polskiej Akademii Nauk w Poznaniu Adiunkt, od 8.5.2018r. kierownik Zakładu Metabolizmu RNA
11.2018 – 2.2019	Instytut Genetyki i Medycyny Molekularnej Uniwersytet w Edynburgu Zjednoczone Królestwo (Szkocja) Goszczący pracownik naukowy Laboratorium LINE-1 (dr Jose L. Garcia-Perez) (10 tygodni)
10.2012 – 3.2018	Instytut Biochemii i Biofizyki Polskiej Akademii Nauk w Warszawie Specjalista i asystent Pracownia Biologii RNA i Genomiki Funkcjonalnej prof. dr hab. Andrzej Dziembowski
11.2007 – 10.2012	Instytut im. Maxa Plancka Chemii Biofizycznej w Getyndze, Niemcy Praktykant, doktorant, stażysta podoktorski Zakład Biochemii Komórkowej prof. dr hab. Reinhard Lührmann
8.2007 – 9.2007	Getyńskie Centrum Biologii Molekularnej (GZMB) Uniwersytet im. Georga Augusta w Getyndze, Niemcy Praktykant Zakład Biologii Komórek Macierzystych prof. dr hab. Andreas Wodarz
4.2007 – 7.2007	Instytut im. Albrechta von Hallera Uniwersytet im. Georga Augusta w Getyndze, Niemcy Praktykant Zakład Biochemii Roślin prof. dr hab. Ivo Feussner
5.2005 – 3.2007	Instytut Chemii Bioorganicznej Polskiej Akademii Nauk w Poznaniu Magistrant Pracownia Biologii Molekularnej Roślin prof. dr hab. Marek Figlerowicz

4. Osiągnięcie naukowe¹

4.1. Tytuł

Określenie mechanizmów regulacji podaży i funkcji RNA w komórkach eukariontów przez enzymy działające na koniec 3' RNA

4.2. Publikacje wchodzące w skład osiągnięcia

Warkocki Z \square , Krawczyk PS, Adamska D, Bijata K, Garcia-Perez JL, Dziembowski A \square Uridylation by TUT4/7 restricts retrotransposition of human LINE-1s.Cell 174(6):1537-1548. https://doi.org/10.1016/j.cell.2018.07.022PMID:30122351 (2018)Cytowania=8IF₂₀₁₇² czasopisma=31.8 (Q1)³Pkt MNiSW=50

Łabno A., **Warkocki Z.**, Kuliński T., Krawczyk P., Bijata K, Tomecki R., Dziembowski A⊠ Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs.

Nucleic Acids Research 44, 10437–10453.

https://academic.oup.com/nar/article/44/21/10437/2529690 PMID: 27431325 (2016)⁴ Cytowania=35 IF₂₀₁₆ czasopisma=10.2 (Q1) Pkt MNiSW=45

Razew M, **Warkocki Z**, Taube M, Kolondra A, Czarnocki-Cieciura M, Nowak E, Labedzka-Dmoch K, Kawinska A, Piatkowski J, Golik P, Kozak M, Dziembowski A, Nowotny M^I

Structural analysis of mtEXO mitochondrial RNA degradosome reveals tight coupling of nuclease and helicase components.

Nature Communications 9,97 <u>https://www.nature.com/articles/s41467-017-02570-5</u> PMID: 29311576 (2018)

Cytowania=5 IF₂₀₁₇ czasopisma=12.4 (Q1) Pkt MNiSW=45

Warkocki Z^I, Liudkovska V, Gewartowska O, Mroczek S, Dziembowski A^I
 Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism.
 Philosophical Transactions of the Royal Society of London - B Biological Sciences
 373(1762). https://doi.org/10.1098/rstb.2018.0162
 PMID:30397099 (2018)
 Cytowania=3
 IF₂₀₁₇ czasopisma=5.7 (Q1)
 Pkt MNiSW=45

¹ WSKAZANIE OSIĄGNIĘCIA WYNIKAJĄCEGO Z ART. 16 UST. 2 USTAWY Z DNIA 14 MARCA 2003 R. O STOPNIACH NAUKOWYCH I TYTULE NAUKOWYM ORAZ O STOPNIACH I TYTULE W ZAKRESIE SZTUKI (Dz. U. 2016 r. poz. 882 ze zm. w Dz. U. z 2016 r. poz. 1311.)

² Podawany jest IF za 2017r. ponieważ IF za 2018 nie został jeszcze wyliczony

³ Podawany jest kwartyl jako wskaźnik wypadkowej oceny jakości czasopism w danej dziedzinie, Q1 – czasopismo znajduje się w grupie 25% najlepszych czasopism w danej dziedzinie

⁴ Praca nagrodzona w IBB PAN III nagrodą za najlepszą pracę eksperymentalną z instytutu w 2016r.

Prace wchodzące w skład osiągniecia obejmują 3 prace eksperymentalne wykonane przez mnie w trakcie stażu podoktorskiego w Laboratorium Biologii RNA i Genomiki Funkcjonalnej prowadzonym przez prof. dra hab. Andrzeja Dziembowskiego w Instytucie Biochemii i Biofizyki PAN w Warszawie. Czas spędzony przeze mnie w grupie prof. Andrzeja Dziembowskiego umożliwił mi znaczące rozwinięcie warsztatu naukowego i pozwolił na rozpoczęcie niezależnej pracy badawczej w Instytucie Chemii Bioorganicznej PAN w Poznaniu. W skład dorobku wchodzi 1 praca przeglądowa przygotowana przez mnie w ICHB PAN w Zakładzie Metabolizmu RNA, którym kieruję od maja 2018r. Wszystkie w/w prace zostały opublikowane w ciągu ostatnich 3 lat (2016-2018) w renomowanych czasopismach i do dnia redakcji tego załącznika zostały zacytowane łącznie 51 razy. Praca w Cell jest drugą (pierwsza w 1990r.) w historii pracą opublikowaną w tym prestiżowym czasopiśmie, w której zarówno pierwszy jak i wiodący autorzy są afiliowani w polskiej jednostce naukowej. Zdecydowana większość eksperymentów, których wyniki są zaprezentowane w tej publikacji została przeprowadzona przeze mnie lub z moim znaczącym udziałem w polskim laboratorium. Wyniki opisane w tej publikacji zostały zauważone przez V Narry Kim, szeroko rozpoznawalną w świecie nauki specjalistkę w dziedzinie urydylacji, i skomentowane w krótkiej pracy przeglądowej tzw. News and Views:

Yeo J, VN Kim

U-tail as a guardian against invading RNAs Nature Structural and Molecular Biology 25(10):903-905. doi: 10.1038/s41594-018-0139-0 (2018) https://www.nature.com/articles/s41594-018-0139-0

Ponadto wyniki zostały również opisane w tzw. *Research Highlight* Strzyz P *TUT-TUTting retrotransposons* **Nature Reviews in Molecular and Cell Biology** 19(10):618. doi: 10.1038/s41580-018-0058-2. (2018) https://www.nature.com/articles/s41580-018-0058-2

W dwóch z prac wchodzących w skład osiągnięcia jestem pierwszym i jednocześnie współkorespondencyjnym autorem. W pracach tych byłem wiodącym wykonawcą. Wykonanie wszystkich z wymienionych prac było częściowo możliwe dzięki uzyskanemu przeze mnie finansowaniu (granty FUGA i SONATA z Narodowego Centrum Nauki). Wszystkie z prac wchodzących w skład osiągnięcia oraz towarzyszące zestawy danych są dostępne na stronach internetowych wydawców w formule *open access*.

Wyniki i zagadnienia przedyskutowane w ww. publikacjach zostały przez mnie w częściach zaprezentowane podczas kilku konferencji i sympozjów naukowych w tym m.in.:

- 1. *"mRNA processing"*, Cold Spring Harbor w Stanach Zjednoczonych Ameryki 2015, plakat
- 2. "Complex life of RNA", EMBL Heidelberg, Niemcy, 2016, plakat
- 3. "The mobile genome", EMBL Heidelberg, Niemcy, 2017, referat ustny
- 4. "RNAtion", Wydział Biologii UAM w Poznaniu, 2017, referat ustny i 3 plakaty
- 5. *"RNA:WIN"*, konferencja w ramach Krajowego Naukowego Ośrodka Wiodącego, Obrzycko, 2018, referat ustny
- 6. "*Complex life of RNA*", EMBL Heidelberg, Niemcy, 2018, 2 plakaty

Dodatkowo zostałem zaproszony do zaprezentowania wyników ujętych w całości lub częściowo w ww. pracach, w szerszym kontekście procesów metabolizmu RNA. Tym sposobem miałem okazję do wygłoszenia kilku ok. godzinnych referatów w kilku instytucjach naukowych w Polsce i za granicą, w tym m.in. w następujących:

- 1. Międzynarodowym Instytucie Biologii Molekularnej i Komórkowej w Warszawie, 2.2.2017 (seminarium instytutowe) i 7.4.2018 (RNA Club Warsaw)
- 2. Instytucie Chemii Bioorganicznej PAN w Poznaniu, 12.5.2017 (wykład inaugurujący sesję sprawozdawczą instytutu za rok 2016)
- 3. Instytucie Genetyki i Medycyny Molekularnej w Edynburgu, 12.12.2018 (seminarium instytutowe)
- 4. Wydziale Biologii UAM w Poznaniu, 1.3.2019 (inicjatywa Do Science! Poznan)
- 5. Wydziale Fizyki UAM w Poznaniu, 12.3.2019 (seminarium wydziałowe)

4.3. Opis osiągnięcia naukowego

Definicja metabolizmu RNA

Ścieżki regulacji ekspresji genów w komórkach eukariotycznych są skomplikowane. Obejmują one szereg różnych procesów w tym tworzenie RNA (transkrypcję RNA), katalizę specyficznych modyfikacji RNA, kontrolę jakości, funkcjonowanie i ostatecznie degradację RNA. Zbiorczo procesy te określane są mianem metabolizmu RNA. Wielość procesów metabolizmu RNA umożliwia właściwe funkcjonowanie komórki oraz jej odpowiedź na wyzwania prezentowane przez procesy rozwojowe, zmiany środowiskowe i patogeny. Zaburzenia w metabolizmie RNA prowadzą natomiast nierzadko do stanów chorobowych, z których wiele objawia się plejotropowymi efektami, tj. takimi, które mają wpływ na wiele procesów komórkowych.

Zróżnicowanie procesów kształtujących końce 3' różnych RNA

Bardzo ważną rolę w metabolizmie RNA odgrywa koniec 3' cząsteczki. Niemal wszystkie rodzaje RNA w komórkach eukariotycznych podlegają po-transkrypcyjnym procesom modyfikującym ich końce 3'. Zakres obróbki końców 3' jest różnoraki, jednak wspólnym elementem dla niemal wszystkich RNA jest endonukleolityczne przecięcie w rejonie końca 3' lub egzonukleolityczne przycięcie od końca 3'. Jest tak w przypadku transkrybowanych przez zależną od DNA polimerazę RNA I (pol I) cząsteczek rRNA, niekodujących, lecz wchodzących w skład rybosomu i tym samym biorącym udział w przepisywaniu informacji genetycznej na białka (w procesie translacji). Końce 3' trzech z czterech rRNA: 28S, 18S i 5.8S rRNA powstają w wyniku endoi egzorybonukleolitycznej obróbki długiego prekursora pre-rRNA [1]. Koniec 3' transkrybowanego przez zależną od DNA polimerazę RNA III (pol III) 5S rRNA podlega enzymatycznemu przycięciu [2]. Podobnie transkrybowane przez pol III tRNA uzyskują swoje końce 3' w wyniku szeregu następujących po sobie aktywności enzymatycznych, których pierwszym efektem jest skrócenie końców 3' [3,4,5]. Inne znaczące pod względem częstości występowania w komórce RNA to snRNA, które odgrywają zasadniczą rolę w katalizie składania egzonów (splicingu pre-mRNA) [6,7]. Spośród nich U6 jest transkrybowane przez pol III i podlega obróbce na końcu 3' [8]. Pozostałe 4 zaangażowane w splicing snRNA (U1, U2, U4 i U5) są transkrybowane przez zależną od DNA polimerazę RNA II (pol II) w formie prekursorów posiadających dłuższe końce 3', które dopiero w wyniku po-transkrypcyjnego przycięcia uzyskują dojrzałą sekwencję [9,10]. Pol II transkrybuje również kodujące białka mRNA. Wszystkie mRNA przechodzą etap cięcia w obrębie ich sekwencji 3' UTR (ang. 3' untranslated region) [11,12].

W przypadku wielu rodzajów RNA kolejnym etapem kształtowania końca 3' jest dodatek niezakodowanych w genomie reszt nukleotydowych. Cząsteczki tRNA uzyskują niezakodowany w genomie koniec 3' o sekwencji CCA, który jest niezbędny w procesie aminoacylacji tj. związania specyficznej reszty aminokwasowej [4]. Większość mRNA uzyskuje na swoich końcach 3' kilkadziesiąt niekodowanych w genomie reszt adenylowych, czasami określanych mianem ogona poli(A) [13,14]. Proces dołączenia reszt adenylowych nosi miano poliadenylacji, która w opisanym tu znaczeniu zachodzi w jądrze komórkowym i jest katalizowana przez jądrowe enzymy PAP α i PAP γ oraz czynniki pomocnicze [11]. Do powstałego ogona poli(A) przyłączają się białka specyficznie go wiążące (ang. poly(A) binding proteins, PABP). Sumarycznie układ ten zapewnia mRNA stabilność oraz wzmacnia ich zdolność do translacji [15,16]. Wyjątkiem wśród mRNA są mRNA kodujące histony, białka niezbędne dla wiązania DNA w postaci chromatyny. Końce 3' mRNA histonów podlegają wyłącznie przycięciu przy udziale innych niż w przypadku pozostałych mRNA enzymów. W wyniku przycięcia na końcach 3' tych mRNA pozostaje struktura typu spinki do włosów, która specyficznie oddziałuje z białkiem (SLBP) [17].

Terminalne transferazy NTP

Opisane w powyższym akapicie poliadenylacja mRNA i dodatek sekwencji CCA do końców 3' tRNA rozpoznano stosunkowo dawno, w latach 70 XX wieku [18,19,20,21]. Natomiast na przestrzeni ostatnich 10-15 lat odkryto i opisano grupę 11 enzymów wspólnie zaliczanych do tzw. terminalnych transferaz NTP (ang. terminal nucleotidyltransferases, TENT). Każde TENT posiada domenę transferazy NTP, zbliżoną do tych obecnych m.in. w polimerazie DNA β oraz enzymach dodających sekwencję CCA do tRNA [22,23]. TENT dodaja niekodowane w genomie reszty nukleotydowe do końców 3' różnych RNA. Na podstawie specyficznej preferencji w kierunku dodawania reszt adenylowych lub urydylowych enzymy te podzielone zostały na (a) tzw. niekanoniczne polimerazy ATP (ang. non-canonical poly(A) polymerases, ncPAPs), odmienne od jądrowych polimeraz ATP odpowiedzialnych za poliadenylację nowych mRNA, oraz (b) terminalne transferazy UTP znane również jako TUTazy (ang. terminal uridyltransferases, TUTs lub TUTases), które katalizują dodatek niezakodowanych w genomie urydyn do końców 3'. Proces przyłączenie urydyn nazywany jest urydylacją. Źródłem energii do wydłużania łańcucha urydylowego jest rozpad trifosforanu w UTP. Podobnie rozpad wiązania trifosforanowego w UTP jest również źródłem energii w przypadku niektórych helikaz [24,25] oraz w fosforylacji przez niektóre kinazy [26]. W komórkach eukariontów potwierdzono występowanie 3 białek, o aktywności terminalnej transferazy UTP nazwanych TENT1, TUT4 i TUT7. Białko TENT1, znane również jako U6 TUTase lub TUT1, jest zlokalizowane w jądrze komórkowym i bierze udział w biogenezie U6 snRNA [8,27,28]. W kontraście do TENT1, TUT4 i TUT7 są białkami zlokalizowanymi w cytoplazmie [29]. W bazie danych gnomAD ilość zidentyfikowanych mutacji typu utraty funkcji (ang. loss of function, LoF) w genach kodujących TUT4 (ZCCHC11) i TUT7 (ZCCHC6) u człowieka znacząco odbiega od przewidywanej na podstawie statystycznej. Sugeruje to, że enzymy te pełnią istotne role w metabolizmie RNA.

Urydylacja RNA i jej wpływu na podaż i funkcje RNA

Dotychczas poznana rola urydylacji w metabolizmie RNA polega na regulacji podaży i funkcji RNA. Wykazano, że TUT4 i TUT7 biorą udział w regulacji i biogenezie mikro RNA, głównie z rodziny let7. W komórkach niezróżnicowanych (pluripotentnych) pre-miRNA let7 wiązane jest przez białko LIN28A, które następnie oddziałuje z TUT4 lub TUT7. W tym trzyskładnikowym kompleksie dochodzi do oligourydylacji pre-miRNA let7, w efekcie prowadzącej do degradacji pre-miRNA [30,31,32,33]. W komórkach zróżnicowanych LIN28A nie jest produkowane. W efekcie odziaływanie TUT4/7 z premiRNA jest krótkotrwałe i w większości wypadków prowadzi do dodania pojedynczej reszty urydylowej do końca 3' pre-miRNA, co jest określane mianem monourydylacji [34,35]. Monourydylacja jest niezbędna w biogenezie specyficznej klasy mikro RNA, umożliwiając wydajną obróbkę przez DICER tych pre-miRNA do dojrzałych cząsteczek mikro RNA [34]. Tym samym różna jakość urydylacji końców 3' pre-miRNA objawia się dwoma przeciwstawnymi efektami na mikro RNA regulując ich podaż. W przypadku wielu innych różnych klas RNA, urydylacja prowadzi do zmniejszenia ich stabilności. Zmniejszanie podaży RNA przez urydylację ma znaczenie w procesach apoptotycznych [36]. Jednak szczególną rolę urydylacja pełni na wczesnych etapach rozwoju osobniczego, w rozwijających się zarodkach, w których urydylowane są mRNA pochodzenia matczynego. Degradacji urydylowanych matczynych mRNA towarzyszy transkrypcja mRNA embrionu, co definiuje dalsze etapy rozwoju i określane jest mianem MZT (ang. *maternal-to-zygotic transition*) [37]. Podobnie indukowana przez urydylację zmiana transkryptomu pełni zasadniczą rolę w spermatogenezie tj. w rozwoju męskich gamet [38]. Podsumowując rola urydylacji polega głównie na indukowaniu degradacji urydylowanych RNA, z prominentnymi wyjątkami dotyczącymi biogenezy pewnych mikro RNA oraz co opisałem poniżej powstrzymywania retrotranspozycji LINE-1.

Degradacja RNA od końca 3'

Istnieją trzy możliwości enzymatycznej degradacji RNA polegające na rozrywaniu wiązań fosfodiestrowych. Pierwsza polega na rozerwaniu łańcucha RNA wewnątrz cząsteczki RNA i określana jest mianem reakcji endorybonukleolitycznej. Druga i trzecia możliwość polega na odłączaniu pojedynczego rybonukleotydu od łańcucha RNA. Może to mieć miejsce albo od końca 5' albo od końca 3' RNA i nosi miano reakcji ezgorybonukleolitycznych, odpowiednio 5'-3' i 3'-5'. Na podstawie obecnie dostępnych wyników badań, endo- i egzonukleolityczne ścieżki degradacji współistnieją w komórkach i uzupełniają się wzajemnie, jednak zablokowanie jednej z nich ma najczęściej poważne konsekwencje dla homeostazy komórki [39]. Degradacja RNA od końca 3' ma szczególne znaczenie w powiązaniu z urydylacją oraz w mitochondriach, komórkowych fabrykach energii. Urydylowane RNA są preferencyjnie degradowane przez egzorybonukleazę 3'-5' DIS3L2. Funkcjonalna współpraca pomiędzy TUTazami a DIS3L2 w regulacji podaży wielu różnych klas RNA została eksperymentalnie potwierdzona w pracy wchodzacej w skład osiagniecia [40] oraz niezależnie potwierdzona przez 2 inne grupy badaczy [41,42]. Wcześniej zaproponowano i wykazano rolę DIS3L2 w degradacji urydylowanych pre-miRNA let7 [43,44,45], co znalazło potwierdzenie w strukturze DIS3L2, w której ścieżka prowadząca RNA do centrum egzorybonukleolitycznego zawiera fragmenty specyficznie oddziałujące z urydynami [46,47]. Poza DIS3L2 w komórkach eukariontów działa również kompleks egzosomu [48]. Ze względu na użycie określenia "egzosom" również w przypadku pęcherzyków wydzielanych przez komórkę na zewnątrz, dla uściślenia, stosuję się również określenie egzosom RNA (ang. RNA exosome). Egzosom RNA przyjmuje kształt beczułki składającej się z 9 strukturalnych białek, które dodatkowo wiążą aktywność egzorybonukleazy 3'-5' [49,50]. W zależności od lokalizacji subkomórkowej może być to DIS3 (w jądrze) [51,52], DIS3L (w cytoplazmie) lub RRP6 (w jąderku) [53]. Kompleks ten odgrywa główną rolę w degradacji 3'-5' u eukariontów i ma swoje odpowiedniki w bakteriach i archeonach.

W mitochondriach szczególna rola degradacji RNA od końca 3' wynika z zestawu mitochondrialnych rybonukleaz, wśród których dotychczas nie zidentyfikowano egzorybonukleazy 5'-3', zaś endorybonukleazy pełnią rolę głównie w obróbce

prekursorowych RNA [54]. Rolę egzorybonukleazy 3'-5' w mitochondriach odgrywa kompleks degradosomu zwany również mtEXO [55]. Skład mtEXO wykazuje różnice pomiędzy eukariontów. O ile helikaza SUV3 została zidentyfikowana jako składowa mtEXO zarówno u człowieka jak i u drożdży, komponent egzorybonukleazowy jest u tych organizmów różny (Dss1p u drożdży i PNPaza u człowieka) [55,56,57,58]. Struktura ludzkiego mtEXO nie została dotychczas poznana.

Omówienie głównych odkryć wchodzących w skład osiągnięcia

W pracach wchodzących w skład osiągniecia przedstawiliśmy sposób w jaki aktywności enzymatyczne zmieniające końce 3' RNA wpływają na metabolizm RNA u eukariontów. Opisane odkrycia dotyczą urydylacji RNA u człowieka i identyfikacji funkcjonalnej współpracy pomiędzy enzymami urydylującymi RNA (TUT4 i TUT7) i egzorybonukleazą 3'-5' DIS3L2 w regulacji podaży kilku różnych klas RNA w tym snRNA oraz transkryptów pol III. Odkryliśmy również molekularny mechanizm przeciwdziałania retrotranspozycji LINE-1 przez urydylację końców 3' LINE-1 mRNA. Zaproponowaliśmy, że przeciwdziałanie retrotranspozycji zachodzi na etapie inicjacji odwrotnej transkrypcji retrotranspozonu podczas tworzenia nowej insercji w jądrowym DNA. Mechanizm ten ze względu na silnie ewolucyjnie konserwowany sposób inicjacji odwrotnej transkrypcji przez LINE-1 może mieć zasadnicze znaczenie dla powstrzymywania retrotranspozycji na etapie po-transkrypcyjnym we wszystkich typach komórek i na każdym etapie rozwojowym. Dodatkowe implikacje wynikające z obecności ogonów poli(A) zapewniają możliwości dalszych badań np. w kontekście zależnych od LINE-1 schorzeń autoimmunologicznych oraz dla rozróżnienia funkcji pełnionych w komórce przez TUT4 i TUT7. W kolejnej pracy pokazaliśmy strukturę kompleksu egzorybonukleolitycznego 3'-5' – mtEXO, będącego głównym regulatorem podaży i funkcji RNA w komórkach drożdży. Nasze odkrycia zostały podsumowane i ukazane w szerszym kontekście w pracy przeglądowej również ujętej w ramach osiągnięcia [59].

W pracy Łabno i wsp., 2016 [40] wykazaliśmy funkcjonalną współpracę pomiędzy TUT4/7 a egzorybonukleazą 3'-5' DIS3L2 w specyficznej regulacji podaży różnych RNA w cytoplazmie, w tym niektórych mRNA, głównie mRNA histonów oraz krótkiego transkryptu z rejonu 5' UTR mRNA ferrytyny, który odkryliśmy. Niemniej głównym odkryciem naukowym było zidentyfikowanie dotychczas nieznanych substratów RNA dla cytoplazmatycznych TUTaz i wykazanie, że te urydylowane RNA są specyficznie degradowane przez DIS3L2.

Zaobserwowaliśmy, że w komórkach pozbawionych aktywności DIS3L2 i wyrażających katalitycznie nieaktywną wersję tego białka (D391N) [60] gromadzą się snRNA o niewłaściwych końcach 3', wydłużonych o sekwencje kodowane w genomie poza rejon spotykany w dojrzałych snRNA. Gromadzące się wydłużone snRNA były dodatkowo urydylowane. Potencjalnie, te wydłużone snRNA mogły reprezentować transkrypty prekursorowe, z których powstaną dojrzałe snRNA. Jednak w naszych badaniach wykluczyliśmy udział TUT4/7-DIS3L2 w dojrzewaniu snRNA poprzez obróbkę ich końców 3', co znalazło wkrótce potem potwierdzenie w postaci zidentyfikowania jądrowego enzymu odpowiedzialnego za tą obróbkę – TOE1 [61]. Wykazaliśmy natomiast, że ścieżka regulacji snRNA przez TUT4/7-DIS3L2 służy usuwaniu niewłaściwie przyciętych w jądrze snRNA, które przedostają się do cytoplazmy gdzie są urydylowane przez TUT4/7 a następnie degradowane przez DIS3L2. Ponadto dodatkowe analizy pozwoliły wykazać brak wpływu DIS3L2 (i TUT4/7) na podaż dojrzałych jądrowych snRNA. Zaproponowany przez nas mechanizm najprawdopodobniej pełni istotną rolę w zapewnieniu fizjologicznej podaży i jakości snRNA eliminując snRNA o wadliwych końcach 3'.

Wykazaliśmy również, że zależna od TUT4/7-DIS3L2 regulacja podaży RNA dotyczy transkryptów produkowanych przez pol III. Polimeraza ta naturalnie zaopatrza produkowane przez siebie transkrypty w 3-6 reszt urydylowych na końcach 3', ponieważ ciąg kilku urydyn w RNA stanowi sygnał końca (terminacji) transkrypcji przez pol III [62,63]. Niemniej, w naszych badaniach pokazaliśmy, że m.in. Y RNA i vault RNA będące transkryptami pol III podlegają dodatkowej urydylacji w cytoplazmie i gromadzą się w komórkach pozbawionych aktywności DIS3L2. W testach in cellulo i w biochemicznych eksperymentach in vitro wykazaliśmy, że nieurydylowane Y i vault RNA cechują się znaczną stabilnością i są odporne na egzorybonukleolityczną aktywność DIS3L2. Z dużym prawdopodobieństwem odporność ta związana jest z wysokim stopniem ustrukturalizowania tych krótkich RNA oraz, w warunkach in vivo, z ich specyficznym oddziaływaniem z białkami [64,65]. Jednak dodanie 2 lub więcej reszt urydylowych na końcach 3' Y i vault RNA znacząco wzmaga wydajność i kinetykę ich degradacji przez DIS3L2. Innym transkryptem pol III podlegającym zależnej od TUT4/7-DIS3L2 degradacji jest BC200, transkrypt typu Alu (ang. *ALU-like element RNA*) [66,67,68].

Sumarycznie w pracy Łabno i wsp., 2016 opisaliśmy zmiany zachodzące w komórkach, w których ścieżka degradacji RNA zależna od TUT4/7 i DIS3L2 nie jest funkcjonalna ze względu na brak aktywności DIS3L2. Zaobserwowane przez nas zmiany mają charakter plejotropowy i dotyczą wielu różnych RNA wpływając na utratę homeostazy komórkowej. Istotnym jest, że mutacje w DIS3L2, które znoszą aktywność tego białka są jednymi z najczęstszych mutacji obserwowanych w rzadkich schorzeniach genetycznych: syndromie Perlmana i nerczaku zarodkowym (guzie Wilmsa) [69]. Opisane przez nas efekty mutacji w DIS3L2 wskazują na zmiany w metabolizmie RNA mogące występować w syndromie Perlmana, nerczaku zarodkowym i innych schorzeniach powodowanym przez mutacje w DIS3L2 oraz potencjalnie TUT4/7. Czy funkcjonalna współpraca TUT4/7 z DIS3L2 gra rolę w trakcie rozwoju zarodowego przy zastępowaniu ekspresji trankryptów matczynych przez transkrypty zygoty/embrionu (MZT) [38,70] nie zostało dotychczas zbadane.

Wkrótce po opublikowaniu naszej pracy, wnioski w niej zawarte (m.in. funkcjonalna współpraca TUT4/7 i DIS3L2 oraz urydylacja niekodujących RNA i transkryptów pol III) znalazły niezależne potwierdzenie w pracach z laboratoriów Stepanki Vanacovej [42] i Richarda Gregorego [41].

W pracy eksperymentalnej Warkocki i wsp., 2018 [29] odkryliśmy molekularny mechanizm, który zmniejsza zdolność retrotranspozonów LINE-1 do wbudowywania się w nowe miejsca w ludzkim genomie.

Retrotranspozony są grupą mobilnych elementów genetycznych, których aktywność kształtowała genomy eukariontów w trakcie ewolucji [71,72]. U współczesnego człowieka pozostałości aktywności mobilnych elementów genetycznych stanowią blisko połowę genomu, zaś retrotranspozonów ok. 30% [71,73], przy czym retrotranspozony są jedyną grupą mobilnych elementów genetycznych, które nadal posiadają zdolność mobilizacji tj. wbudowywania się w nowe miejsca w genomie [74,75]. Retrotranspozony obejmują 3 grupy: LINE-1, ALU i SVA, z których tylko LINE-1 koduje wszystkie elementy niezbędne do insercji w nowe genomowe loci. Cykl życiowy LINE-1 obejmuje transkrypcję genu LINE-1 przez pol II, co prowadzi do powstania kapowanego (posiadającego specyficzną strukturę stabilizującą koniec 5') i poliadenylowanego mRNA. LINE-1 mRNA jest odmienne od innych ludzkich mRNA, ponieważ koduje 2 białka: L1-ORF1p i L1-ORF2p, w dwóch następujących po sobie ramkach odczytu. Pierwsze z białek posiada domenę typu coiled-coil, tworzy homotrimery i wiąże in cis LINE-1 mRNA, z którego powstało [76,77,78]. Białko L1-ORF2p posiada natomiast dwie aktywności enzymatyczne: nikazy DNA (tj. zdolność do nacinania jednej nici DNA w dwuniciowej cząsteczce DNA) [79] oraz odwrotnej transkryptazy (tj. przepisywania RNA na komplementarne DNA – cDNA) [80]. L1-ORF2p powstaje w wyniku nietypowego mechanizmu reinicjacji translacji, co sprawia, że ilość kopii tego białka jest ok 1000 razy niższa niż L1-ORF1p [81]. L1-ORF2p podobnie jak L1-ORF1p wiąże się z LINE-1 mRNA w obrebie jego poliadenylowanego końca 3' [82] tworząc LINE-1 RNP [83], które transportowane jest do jadra komórkowego. W nim może zajść wbudowanie LINE-1 w genomowe DNA, w procesie określanym mianem TPRT (ang. target-primed reverse transcription). Pierwszym etapem TPRT jest nacięcie genomowego DNA przez L1-ORF2p w obrębie konserwowanej ewolucyjnie sekwencji A|TTTTT, gdzie | oznacza miejsce cięcia [84]. W wyniku nacięcia możliwe jest uwolnienie krótkiego odcinka oligo(dT), który paruje się z resztami A w ogonie poli(A) LINE-1 mRNA stanowiąc starter dla odwrotnej transkrypcji przez L1-ORF2p [85]. Właśnie na tym etapie zasadniczego znaczenia nabiera sekwencja końca 3' LINE-1 mRNA. W naszej pracy wykazaliśmy, że LINE-1 mRNA podlegają wydajnej urydylacji przez TUT4/7 w różnych ludzkich liniach komórkowych w tym w ludzkich komórkach embrionalnych i komórkach progenitorowych neuronów, w których LINE-1 podlegają wysokiej ekspresji (podczas gdy w większości komórek somatycznych ich ekspresja jest blokowana epigenetycznie uniemożliwiając transkrypcję). Stosując nowoczesne podejście eksperymentalne oparte o wysokoprzepustowe sekwencjonowanie [86] przeprowadziliśmy analizę końców 3' LINE-1 mRNA. Blisko 40% LINE-1 mRNA było urydylowanych, a tylko ok 50% posiadało 3' adeniny, przy czym większość z tych transkryptów posiadała krótkie ogony oligo(A) liczące do ok. 10 adenin. Wśród zsekwencjonowanych końców 3' praktycznie brak było ogonów poli(A) liczących więcej niż 50 adenin. Ok 5% zsekwencjonowanych końców 3' nie posiadało ani ogona oligo/poli(A) ani reszt urydylowych, będąc potencjalnie fragmentami pochodzącymi z degradacji 3'-5' LINE-1 mRNA. Powyższe obserwacje poczyniliśmy dla różnych klas endogennych LINE-1 wyrażanych z genomowych loci oraz LINE-1 mRNA powstających na matrycy plazmidów zawierających tzw. reportery LINE-1. Co istotne, w naszej pracy wykazaliśmy, że obecność nawet pojedynczej reszty urydylowej na końcu 3' RNA zmniejsza poziom retrotranspozycji LINE-1 (zdolność do wbudowania się w genomowe DNA) o 20-30%, a 3-4 urydyny o ponad 50%. W zgodzie z tymi obserwacjami w komórkach wyrażających zwiększone ilości TUT4 lub TUT7 poziom retrotranspozycji reporterów LINE-1 był znacząco zredukowany. Przeciwny efekt obserwowaliśmy w komórkach, w których aktywność TUT4 i/lub TUT7 została wyciszona przez interferencję RNA. Wykazaliśmy, że podaż TUT4 i/lub TUT7 w komórkach ma ograniczony wpływ na ilość LINE-1 mRNA oraz retrotranspozonowych białek L1-ORF1p i L1-ORF2p. Podobnie obserwowany przez nas wpływ TUT4/7 na stabilność i dynamikę degradacji LINE-1 mRNA w komórce był ograniczony. Tym samym potwierdzając, że obserwowany efekt urydylacji na retrotranspozycję LINE-1 wynika w przeważającej mierze z jakościowych zmian końców 3' LINE-1 mRNA, a nie ze zmian ilościowych.

W dalszej kolejności wykazaliśmy, że urydylacja końców 3' LINE-1 mRNA jest możliwa dzięki funkcjonalnej współpracy TUT4/7 z białkiem MOV10, które zidentyfikowaliśmy jako białko specyficznie wzbogacone w eksperymentach immunoprecypitacji z TUT4 oraz, w mniejszym stopniu, z TUT7. MOV10 jest helikazą 5'-3' oddziałującą z końcami 3' RNA w tym LINE-1 mRNA [87]. Wykazaliśmy, że MOV10 zwiększa poziom urydylacji LINE-1 mRNA *in vivo* i *in vitro*. Dodatkowo pokazaliśmy, że TUT4 ale nie TUT7 akumuluje się w cytoplazmatycznych *foci* ("ciałkach"), co może mieć wpływ zarówno na zakres substratów jak i odmienny efekt jaki urydylacja przez TUT4 lub TUT7 ma na RNA.

Podsumowując opisany w pracy Warkocki i wsp., 2018 molekularny mechanizm zmniejszający wydajność retrotranspozycji LINE-1 ma zasadnicze znaczenia dla utrzymania stabilności ludzkiego genomu. Jest to szczególnie istotne w komórkach, w których LINE-1 podlegają wydajnej ekspresji. Dzieje się tak w najistotniejszych dla rozwoju i proliferacji człowieka okresach, w tym na wczesnym rozwoju zarodkowym i w trakcie tworzenia komórek rozrodczych. Wtedy znaczniki epigenetyczne są czasowo znoszone aby umożliwić rozwój nowego organizmu i nabycie przezeń indywidualnych cech [88]. Wysoka konserwacja ewolucyjna sekwencji rozpoznawanej przez L1-ORF2p i uwalnianie krótkiego oligo(dT) jako startera dla odwrotnej transkrypcji [84,89] sprawia, że odkryty przez nas mechanizm jest wyjątkowo skuteczny w zapobieganiu nowym potencjalnie chorobotwórczym insercjom LINE-1. Innymi słowy udało nam się odkryć piętę Achillesową LINE-1.

Urydylacja końców 3' LINE-1 mRNA może mieć jednak inne daleko idące efekty na homeostazę komórki. Badaniu tych procesów zamierzam poświęcić dużo uwagi jako lider grupy (Zakład Metabolizmu RNA) w IBCH PAN w Poznaniu. W pracy Razew i wsp., 2018 [90] przedstawiliśmy strukturę drożdżowego kompleksu mtEXO, zwanego również mitochondrialnym degradosomem. MtEXO degraduje RNA od końca 3' w komórkowych fabrykach energii – mitochondriach. W pracy prezentujemy w jaki sposób dwa białka tworzące ten unikalny "kompaktowy" kompleks współpracują w degradacji RNA.

Kompleks białkowy drożdżowego mtEXO zawiera białka Dss1p i Suv3p, oba kodowane przez genom jądrowy [91,92,93,94]. Dss1p jest 3'-5' exorybonukleazą należącą do klasy enzymów podobnych do RNAzy II (ang. RNAse II-like), tej samej klasy enzymów co omawiane powyżej DIS3L2 [46,95]. Suv3p jest helikazą (białkiem rozplatającym dwuniciowe RNA) wykazującą szereg podobieństw do rodziny helikaz SF2, jednak posiadającą pewne wyróżniające cechy [96]. Razem te dwa białka tworzą kompleks mtEXO odpowiedzialny za regulację podaży i funkcji RNA w mitochondriach drożdzy. W porównaniu do metabolizmu RNA w cytoplazmie obejmującego m.in. urydylację, sposób regulacji podaży RNA oraz obróbki końców 3' w mitochondriach reprezentuje uproszczony system, w którym główna rola przypada właśnie mtEXO [97,98]. Dzieje się tak dlatego iż system regulacji transkrypcji mitochondrialnego opiera się na dwóch filarach: transkrypcji i degradacji RNA [97]. Ponadto dotychczas nie opisano żadnej aktywności egzorybonukleolitycznej 5'-3' w mitochondriach [54,55]. Kodowane przez mitochondrialny genom białka wchodzą w skład łańcucha oddechowego pełniąc istotne role w tworzeniu energii w postaci ATP dla procesów komórkowych [99]. Białka te podlegają translacji w mitochondriach z udziałem mitochondrialnych rybosomów i tRNA, niemniej większość mitochondrialnych białek kodowanych jest w genomie jądrowym i powstaje w wyniku translacji w cytoplazmie [99,100].

W naszej pracy uzyskaliśmy kryształy i rozwiązaliśmy struktury Dss1p i mtEXO z drożdży Candida glabrata. Uzyskaliśmy kryształy Dss1p, skróconego na końcu N o 69 aminokwasów (Cq-Dss1⁷⁰⁻⁹⁰⁰), oraz białka skróconego i posiadającego mutację w miejscu katalitycznym (Cq-Dss1⁷⁰⁻⁹⁰⁰ D477N). Cq-Dss1⁷⁰⁻⁹⁰⁰ posiada typową dla RNAz typu II domenę katalityczną RNB (ang. RNA binding domain). Białko wyróżnia się jednak spośród innych enzymów tej grupy kompozycją domen dodatkowych, wśród których możliwe jest wyróżnienie domeny typu KOW (w obrębie β -barrel), domeny WH (ang. winged helix), podwójnej domeny HTH (ang. helix-turn-helix) i C-końcowej domeny S1. Dss1p posiada również N-końcową helisę, natomiast sam koniec Ν najprawdopodobniej nie posiada zdefiniowanej struktury. Wymienione domeny zastępują domeny CSD1 i CSD2 (ang. cold-shock domain) typowo obserwowane w białkach typu RNAzy II [101], w tym w DIS3L2 [46] i RRP44 [102]. Ww.domeny podobnie jak CSD1 i CSD2 dekorują domenę RNB tworząc lejek, który może oddziaływać z RNA i prowadzi do wnętrza domeny RNB, na której spodzie znajduje się miejsce egzorybonukleolityczne 3'-5'.

Uzyskana przez nas struktura *Cg*-mtEXO powstała w oparciu o Dss1⁷⁰⁻⁹⁰⁰ D477N i Suv3^{43–685}. Mutacja w miejscu katalitycznym Dss1p umożliwiła krystalizację w obecności RNA i ATP/Mg²⁺ i wizualizację RNA w strukturze białka. Struktura Dss1p

w *Cg*-mtEXO jest bardzo zbliżona do struktury białka w izolacji, a struktura Suv3p w *Cg*-mtEXO jest zbliżona do wcześniej określonej struktury ludzkiego białka [96]. W *Cg*-mtEXO Suv3p oddziałuje z domenami dekorującymi Dss1p. Kilka miejsc kontaktu znajduje się w obrębie HTH w Dss1p i w domenie RecA1 w Suv3p, w obrębie unikalnej dla Suv3 helisy. Dodatkowe miejsca oddziaływania białek w kompleksie znajdują się odpowiednio w C-końcowej domenie *Cg*-Suv3p i domenie WH w *Cg*-Dss1p. Ponadto N-końcowa domena *Cg*-Suv3p jest obecna w szczelinie pomiędzy WH, HTH i RNB. Wykazaliśmy, że główne miejsce oddziaływania oparte o unikalne dla Suv3p i Dss1p domeny/reszty aminokwasowe ma zasadnicze znaczenie dla struktury mtEXO. W eksperymentach *in vivo* wykazaliśmy, że zmutowanie aminokwasów istotnych dla tego oddziaływania objawia się uszkodzoną zdolnością do oddychania tlenowego. Za pomocą metody biofizycznej SAXS (ang. *small angle X-ray scattering*) potwierdziliśmy, że obserwowane przez nas ułożenie Suv3p i Dss1p w krysztale, ma miejsce również w roztworze wodnym, a zatem jest to funkcjonalna i biologicznie ważna konformacja mtEXO.

Uzyskana przez nas struktura umożliwiła zobrazowanie kanału w mtEXO, którym do miejsca katalitycznego prowadzone jest RNA. Struktura i dodatkowe badania z wykorzystaniem różnych innowacyjnych i tradycyjnych podejść biochemicznych pozwoliły uściślić długość łańcucha RNA w mtEXO do 16-18 nukleotydów. Wykazaliśmy, że Suv3p w mtEXO umożliwia wydajną i szybką degradację RNA posiadających struktury drugorzędowe (w tym fragmenty LINE-1 3' UTR i vaultRNA1-2) co nie jest możliwe w przypadku samego Dss1p. Nasze badania pozwoliły potwierdzić model działania mtEXO, w którym RNA jest aktywnie wplatane przez Suv3p do kanału mtEXO w kierunku miejsca katalitycznego, gdzie zachodzi egzorybonukleolityczna 3'-5' degradacja RNA.

Podsumowując w pracy Razew i wsp., 2018 zaprezentowaliśmy całościową analizę struktury i funkcji drożdżowego mtEXO degradującego RNA od końca 3'. Uzyskane dane pozwalają zrozumieć degradację RNA w komórkowych fabrykach energii, mitochondriach, których prawidłowe funkcjonowanie jest nieodzowne dla eukariontów.

W pracy przeglądowej Warkocki i wsp., 2018 [59] podsumowaliśmy i usystematyzowaliśmy odkrycia dotyczące TENT, stosując ujednolicone nazewnictwo tych enzymów wprowadzone przez HUGO Gene Nomenclature Committee (https://www.genenames.org/) w 2018r. Omawiane białka obejmują TENT2, TENT4A, TENT4B, TENT5A, TENT5B, TENT5C, TENT5D, MTPAP (TENT6), TENT1, TUT4 i TUT7. Enzymy te omówiliśmy w podziale na adenyltransferazy i urydyltransferazy. Szczególną uwagę poświęciliśmy wyborowi i omówieniu najważniejszych odkryć. W mojej ocenie praca znacząco wykracza poza podsumowanie i zebranie w jednym opracowaniu odkryć innych badaczy. Jest tak dlatego, że w oparciu o nagromadzone dane odnieśliśmy się do niektórych z wcześniejszych odkryć i propozycji mechanizmów działania TENT, a także ich specyficzności substratowej krytycznie oceniając zaproponowane mechanizmy. Praca w mojej ocenie stanowi dobry punkt wyjścia dla

osób chcących zaznajomić się z procesami metabolizmu RNA regulowanymi przez TENT.

4.4. Opis dorobku naukowego spoza zbioru prac wliczanych do osiągnięcia

W skład pozostałego dorobku naukowego wchodzą 4 prace eksperymentalne i 1 praca przeglądowa.

Praca eksperymentalna

Szczesny R, Kowalska K, Klosowska-Kosicka K, Chlebowski A, Owczarek E, Warkocki Z, Kulinski T, Adamska D, Affek K, Jedroszkowiak A, Kotrys A, Tomecki R, Krawczyk P, Borowski L, Dziembowski A

Versatile approach for functional analysis of human proteins and efficient stable cell line generation using FLP-mediated recombination system.

PLoS ONE 13, e0194887

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194887 PMID: 29590189 (2018)

Cytowania=2 IF5₂₀₁₇ czasopisma=3.4 Pkt MNiSW=35

Praca została wykonana w Instytucie Biochemii i Biofizyki PAN w Warszawie. Zaprezentowaliśmy w niej zestaw 52 plazmidów do tworzenia stabilnych linii komórkowych korzystając z komórek HEK293 Flp-in T-rex. Potwierdziliśmy użyteczność zestawu plazmidów w tworzeniu linii komórkowych oraz eksperymentów biologii molekularnej m.in. do różnych procedur wykorzystujących (w tym immunoprecypitację oraz do badania funkcji białek in cellulo za pomocą zestawów do eksperymentów typu "depletion and rescue"), biologii komórkowej (w tym badań lokalizujących białka w komórce za pomocą szeregu różnych białek fluorescencyjnych dołączanych jako metka do białek badanych) oraz biochemii (oczyszczanie białek z komórek ludzkich). Na poczet tej pracy zaprojektowałem kilka wektorów serii pKK, część z nich również przygotowałem i potwierdziłem ich użyteczność (wektory kodujące sekwencję MBP, 3x FLAG, HA i wektory do znakowania RNA (pKK-RNAtag vector series). Opisane w pracy podejścia badawcze oraz przygotowane plazmidy zostały z powodzeniem wykorzystane w pracach eksperymentalnych wchodzących w skład opisanego powyżej osiągnięcia.

Pozostałe 3 prace eksperymentalne wykonane zostały w Instytucie im. Maxa Plancka w Getyndze (niem. Max-Planck-Institut für biophysikalische Chemie) w Zakładzie Biochemii Komórkowej kierowanym przez prof. Reinharda Lührmanna:

Warkocki Z, Odenwälder P, Schmitzová J, Platzmann F, Stark H, Urlaub H, Ficner R, Fabrizio P, Lührmann R.

Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components.

Nature Structural and Molecular Biology 16(12):1237-43.

https://www.nature.com/articles/nsmb.1729 PMID:19935684 (2009) Cytowania=111

IF₂₀₀₉ czasopisma=12 (Q1) Pkt MNiSW=45

Warkocki Z, Schneider C, Mozaffari-Jovin S, Schmitzová J, Höbartner C, Fabrizio P, Lührmann R

The G-patch protein Spp2 couples the spliceosome-stimulated ATPase activity of the DEAH-box protein Prp2 to catalytic activation of the spliceosome.

Genes & Development 29(1):94-107. <u>http://genesdev.cshlp.org/content/29/1/94.long</u> PMID:25561498 (2015)

Cytowania=28

IF₂₀₁₅ czasopisma=10 (Q1) Pkt MNiSW=45

Ohrt T, Odenwälder P, Dannenberg J, Prior M, **Warkocki Z**, Schmitzová J, Karaduman R, Gregor I, Enderlein J, Fabrizio P, Lührmann R.

Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system.

RNA 19(7):902-15 <u>https://rnajournal.cshlp.org/content/19/7/902.long PMID:23685439</u> (2013)

Cytowania=35

IF₂₀₁₃ czasopisma=4.6 (Q1) Pkt MNiSW=35

Prace te poświęcone są procesowi składania egzonów nazywanemu również splicingiem (ang. pre-mRNA splicing). Proces polega na wycinaniu z cząsteczki prekursorów mRNA zwanych pre-mRNA sekwencji niekodujących (intronów) i składaniu sekwencji kodujących (egzonów) w dojrzałą cząsteczkę mRNA, która może zostać przetłumaczona na białko w procesie translacji. Składanie egzonów zachodzi w dwóch następujących po sobie reakcjach transesteryfikacji (etapy 1 i 2) i jest katalizowane przez kompleks białkowo-rybonukleinowy nazywany spliceosomem [6]. Spliceosom jest niewątpliwie jednym z najbardziej skomplikowanych kompleksów w biologii molekularnej, ponieważ w jego skład poza pre-mRNA wchodzi 5 różnych snRNA oraz ponad 80 białek u drożdży i ponad 100 u człowieka. Spliceosom składa się de novo na każdym pre-mRNA i przechodzi szereg dynamicznych zmian w swoim składzie białkowo-RNA i w swojej strukturze [6]. Zasadniczą rolę w reorganizacji spliceosomu odgrywa 8 helikaz RNA/RNPaz (białek rozrywających kontakty białek z RNA), które wykorzystując energię z hydrolizy NTP modelują nie tylko strukturę RNA ale wpływają również na skład białkowy spliceosomu. Gdy przystępowałem do realizacji mojego projektu doktorskiego większość wiedzy o splicingu oparta była o eksperymenty genetyczne in vivo lub przeprowadzane in vitro w obecności ekstraktów komórkowych lub z częściowo oczyszczonymi kompleksami spliceosomu. W tych badaniach niemożliwe było określenie pełnego składu białkowego aktywnego spliceosomu. Podobnie utrudnione było przeprowadzenie szczegółowych badań biochemicznych nad regulacją i efektami działania helikaz/RNPaz na zmianę składu białkowego i RNA kompleksu oraz na jego strukturę na poziomie kontaktów pomiędzy poszczególnymi białkami lub RNA i na poziomie makromolekularnym.

Moim głównym osiągnięciem naukowym było opracowanie systemu do biochemicznego i biofizycznego badania splicingu w warunkach *in vitro* z wykorzystaniem oczyszczonych kompleksów drożdżowych spliceosomów i rekombinowanych białek. W pierwszej kolejności prowadziłem reakcje splicingu in vitro z zastosowaniem transkrybowanego in vitro RNA i ekstraktów przygotowanych z komórek ze zmutowanym genem helikazy/RNPazy Prp2 (prp2-1). Wykorzystanie termicznie inaktywowanych ekstraktów z komórek prp2-1 umożliwiło złożenie kompleksu spliceosomu in vitro, lecz brak katalitycznie aktywnego białka Prp2p uniemożliwiał katalizę etapu 1 splicingu. Złożone kompleksy mogły zostać wydajnie oczyszczone w wieloetapowym procesie obejmującym rozdział w gradiencie glicerolowym, chromatografię powinowactwa kompleksu do złoża oraz ponowny rozdział w gradiencie glicerolowym [103,104]. W podobny sposób w laboratorium oczyszczono m.in. spliceosomy zatrzymane w procesie splicingu pomiędzy etapami 1 i 2 [103]. Na podstawie analiz składu RNA i białek oczyszczonych spliceosomów prp2-1 i porównaniu składu tychże ze spliceosomami zatrzymanymi pomiędzy etapami 1 i 2 [103] możliwe stało się zidentyfikowanie białek niezbędnych do katalizy splicingu, wśród nich białka Cwc25p o w owym czasie nieznanej funkcji [104]. Innymi słowy, w tych badaniach oczyściliśmy oraz zdefiniowaliśmy skład drożdżowego spliceosomu na etapie, który zdefiniowaliśmy jako B^{act∆prp2}. Następnie poprzez dodanie rekombinowanego białka Prp2p i inkubację z ATP możliwe było uzyskanie kompleksu, który wraz ze współpracownikami określiliśmy mianem B*. Skład i struktura B* różniły się od kompleksu B^{actAprp2} definiując tym samym rolę Prp2p w procesie, który określiliśmy mianem aktywacji katalitycznej spliceosomu. Wykazaliśmy, że B* wydajnie katalizuje etap 1 splicingu tylko w obecności białka Cwc25p, które najprawdopodobniej stabilizuje strukturę RNA kompleksu sprzyjającą katalizie. W wyniku katalizy etapu 1 powstaje kompleks C. W pracy zademonstrowaliśmy, że dodanie do C kolejnych 3 białek (Prp16p, Slu7p i Prp22p) prowadzi do etapu 2 tj. łączenia egzonów. Tym samym z wykorzystaniem systemu udało się nam określić skład białkowy aktywnego spliceosomu i uzyskać narzędzie do szczegółowych badań ról jakie w katalizie etapów 1 i 2 pełnią różne białka. Opracowany system rekonstytucji splicingu in vitro pozwolił m.in. na szczegółowe zbadanie roli Prp2p. Najważniejszym efektem działalności Prp2p jest zależne od hydrolizy NTP obniżenie powinowactwa wiązania do spliceosomu 14 z ponad 45 białek obecnych w kompleksie B^{act} oraz zmiany jego makromolekularnej struktury [104]. Ponadto aktywność Prp2p prowadzi do stworzenia miejsca wiązania dla białka Cwc25p [104,105,106]. Białko to wiąże się specyficznie ze spliceosomem B* lecz nie z B^{act} i jest niezbędne do wydajnej katalizy etapu 1.

System rekonstytucji splicingu *in vitro* oraz przedstawione tu pokrótce wyniki zostały opublikowane w renomowanym czasopiśmie *Nature Structural & Molecular Biology* w artykule, którego jestem pierwszym autorem. W tej pracy oraz w mojej pracy doktorskiej (opublikowanej przez wydawnictwo Sierke, ISBN: 978-3-86844-439-1) zamieszczone zostały również struktury uzyskane za pomocą mikroskopii elektronowej z kontrastowaniem związkami uranu (ang. *negative stain EM*) oraz przy pomocy techniki cryo-EM polegającej na witryfikacji kompleksu białkowo-RNP. Obrazy te uzyskane

zostały we współpracy z grupą mikroskopii elektronowej kierowaną przez prof. dra Holgera Starka. W badaniach zajmowałem się oczyszczaniem spliceosomu z reakcji in vitro z zastosowaniem różnych buforów, dodatków niskocząsteczkowych oraz różnych sekwencji etapów w procesie oczyszczania. Badania te wpisywały się w szeroki nurt badań strukturalnych nad spliceosomem, których ukoronowaniem były prace z lat 2016-2018 przedstawiające struktury spliceosomów [107,108,109], w tym praca przedstawiająca drożdżowy kompleks B^{act} [110]. Opracowany przeze mnie system nadal może potencjalnie posłużyć do stworzenia rodzaju "filmu" tj. dokumentacji dynamicznych przemian spliceosomu ze stadium B^{act_prp2} do stadium postkatalitycznego wykorzystując do tego oczyszczone spliceosomy i dodatek rekombinowanych białek i ATP. Faktycznie w bardzo niedawno opublikowanej pracy z laboratorium Yigong Shi, autorzy wykorzystując podejście eksperymentalne zbliżone do opracowanego przeze mnie pokazują struktury umożliwiające wgląd w proces, w którym Prp2p tworzy miejsce katalityczne dla Cwc25p i Yju2p [111]. Część moich pozostałych wyników ukazała się w publikacji w RNA (z 2013r.). Opracowany przeze mnie system z powodzeniem został wykorzystany do przeprowadzenia w laboratorium prof. Lührmanna szeregu szczegółowych badań [105,112,113,114].

W trakcie realizacji ww.projektów zdobyłem bogate doświadczenie w oczyszczaniu i analizie biochemicznej kompleksów białkowo rybonukleinowych, a także w biochemii enzymów operujących w takich kompleksach – helikaz/RNPaz. Uczestniczyłem w projektowaniu eksperymentów kinetycznych i biofizycznych.

Moja dalsza praca w zespole, w tym po obronie pracy doktorskiej, dotyczyła badania roli jaką w katalizie splicingu spełnia białko Spp2p. W owym czasie wiedziano, że Spp2p i Prp2p są zdolne do oddziaływania [115]. Spekulowano również rolę Spp2p w tworzeniu miejsca wiązania dla Prp2p w spliceosomie [116]. Jednocześnie dotychczasowe badania dotyczące zapotrzebowania w procesie splicingu na energie z hydrolizy NTP przeprowadzano wyłącznie z wykorzystaniem rekombinowanych białek przy braku spliceosomu, lub też z nieoczyszczonymi spliceosomami w ekstraktach komórkowych. W moich badaniach udało mi się opracować warunki oczyszczania drożdżowego spliceosomu, które pozwalają na specyficzne usunięcie Spp2p z kompleksu (i otrzymanie kompleksu, który nazwaliśmy B^{actAprp2Aspp2}). Wykazałem, że wbrew wcześniejszym sugestiom Spp2p nie jest niezbędne do wiązania Prp2p ze spliceosomem. Natomiast Spp2p ma znaczący wpływ na aktywność NTPazową Prp2p stymulując hydrolize NTP przez to białko. Przeprowadziłem eksperymenty biochemiczne w celu określenia konsumpcji UTP przez Prp2p podczas konwersji oczyszczonego B^{act} do B*. Monitorowałem konsumpcję UTP ponieważ w skład kompleksów B^{act} i B* wchodzi białko Brr2p będące ATPazą i białko Snu114p będące GTPazą. Aktywności ATP- i GTPazowe Brr2p i Snu114p są niezbędne w innych etapach splicingu niż konwersja B^{act} do B*, gdyż ta zachodzi równie wydajnie w obecności każdego NTP, co wykazałem. Na podstawie

18

wyników moich badań zasugerowałem, że w obecności oczyszczonych spliceosomów Spp2p umożliwia wykorzystanie energii generowanej przez Prp2p podczas hydrolizy NTP do przemodelowania kompleksu, podczas gdy przy braku Spp2p wytworzona energia jest zużywana w sposób bezproduktywny. W moich eksperymentach jako pierwszy badacz miałem okazję oszacować konsumpcję UTP przez spliceosomową helikazę/RNPazę w obecności oczyszczonego kompleksu. Wyniki mojej pracy zostały wpierw zaprezentowane podczas konferencji "RNA Society Meeting 2013" w Davos w Szwajcarii, a następnie w 2015r. opublikowane w czasopiśmie Genes and Development, w publikacji, której jestem pierwszym autorem.

Grupa prowadzona przez prof. Lührmanna należy do najlepszych w świecie w badaniu splicingu, wiodąc prym wpierw w badaniach biochemicznych, a następnie korzystając z opracowanych w niej sposobów oczyszczania spliceosomów na badaniach strukturalnych spliceosomu. W Instytucie Chemii Biofizycznej im. Maxa-Plancka pracowało 4 laureatów nagrody Nobla w tym laureat w dziedzinie fizjologii i medycyny z roku 2014 prof. Stefan Hell. Pracę w tym znakomitym naukowo miejscu uważam za bardzo cenne doświadczenie i atut w planowanej dalszej pracy naukowej.

W trakcie doktoratu wziąłem czynny udział w kilkunastu konferencjach i sympozjach naukowych w tym w konferencji "mRNA processing" w Cold Spring Harbor w USA (2009), konferencji "Molecular Life Sciences" we Frankfurcie (2011), konferencji "Multipole Approach to Structural Biology" w Warszawie (2011), konferencji "Mossbacher Kolloqium" w Mossbach (2010), w "Complex Life of mRNA" w Heidelbergu (2012) i corocznych "Horizons in Molecular Biology" w Getyndze. W ramach szkoły doktoranckiej GGNB (*Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences*) uczestniczyłem jako słuchacz w wykładach i w kilkunastu kilkudniowych kursach metod w biologii molekularnej, biochemii i biofizyki. Byłem również osobą prowadzącą 3 kursy poświęcone oczyszczaniu kompleksów makromolekularnych metodami chromatografii powinowactwa.

Ostatnią w tym spisie, choć *de facto*, moją pierwszą pracą była publikacja przeglądowa poświęcona krótkim interferencyjnym RNA działającym *in trans* (ang. *trans-acting siRNA*) w komórkach roślin.

Warkocki Z, Figlerowicz MKrótkie interferencyjne RNA działające in transPostępy Biochemii 52 (3), 253-259, 2006http://www.postepybiochemii.pl/#File?./html/3 2006/253.htmlCytowania: 3 (Google Scholar)Pkt MNiSW=5

Publikacja powstała z mojej inicjatywy w trakcie wykonywania przeze mnie projektu magisterskiego w Pracowni Biologii Molekularnej Roślin kierowanej przez prof. dra hab. Marka Figlerowicza w ICHB PAN w Poznaniu. W owym czasie tasiRNA były nowo odkrytą grupą krótkich RNA [117,118,119]. Wykazano, że biorą one udział w potranskrypcyjnej regulacji ekspresji genów (w tym m.in. genów z rodziny ARF, ang. *auxin response factors*), podobnie jak siRNA wchodząc w skład kompleksu RISC prowadząc do przecięcia i w efekcie degradacji mRNA. Charakterystyczny jest sposób powstawania tasiRNA, w który zaangażowane są m.in. zależne od RNA polimerazy RNA, enzymy obecne u roślin ale nie u zwierząt [120]. W 2006r. moja praca przeglądowa podsumowująca odkrycia dotyczące tasiRNA była pierwszą w języku polskim i jedną z pierwszych na świecie podsumowująca to nowe odkrycie.

4.5. Wybrane nagrody i przyznane granty na badania⁵

2018 – 2021	SONATA grant Narodowego Centrum Nauki (Polska)
2018	Stypendium wyjazdowe EMBO do Uniwersytetu w Edynburgu (UE)
2016	III nagroda za najlepszą pracę eksperymentalną IBB PAN w 2016r. (Polska)
2012 – 2015	FUGA grant Narodowego Centrum Nauki (Polska)
2014	Stypendium START Fundacji na Rzecz Nauki Polskiej (Polska)
2008 – 2010	Stypendium im. Augusta Kekulé z Fonds der chemischen Industire (Niemcy)
2007	Stypendium Erasmus na pobyt na Uniwersytecie Georga Augusta (UE)
2006	Stypendium Ministra Nauki i Szkolnictwa Wyższego (Polska)
2002	Laureat XXXI Olimpiady Biologicznej dla uczniów szkół średnich (Polska)

- 4.6. Spis literatury
- 1 Tomecki, R., Sikorski, P. J. and Zakrzewska-Placzek, M. (2017) *FEBS letters*, 591(13), pp. 1801– 1850. PMID:28524231.
- 2 Gerstberger, S. *et al.* (2017) *Cell reports*, 21(3), pp. 758–772. PMID:29045842.
- *3* Maraia, R. J. and Lamichhane, T. N. (2011) *Wiley Interdisciplinary Reviews: RNA*, pp. 362–375. PMID:21572561.
- 4 Wellner, K., Betat, H. and Mörl, M. (2018) *Biochimica et Biophysica Acta (BBA) Gene Regulatory Mechanisms*, 1861(4), pp. 433–441. PMID:29374586.
- 5 Wichtowska, D., Turowski, T. W. and Boguta, M. (2013) *Wiley Interdisciplinary Reviews: RNA*, pp. 709–722. PMID:24039171.
- *6* Wahl, M. C., Will, C. L. and Lührmann, R. (2009) *Cell*, 136(4), pp. 701–718. PMID:19239890.
- Will, C. L. and Lührmann, R. (2011) Cold Spring Harbor perspectives in biology, 3(7).
 PMID:21441581.
- 8 Mroczek, S. and Dziembowski, A. (2013) *Wiley interdisciplinary reviews. RNA*, 4(5), pp. 581–592. PMID:23776162.
- *9* Chen, J. and Wagner, E. J. (2010) *Biochemical Society Transactions*, 38(4), pp. 1082–1087. PMID:20659008.

⁵ W nawiasach kraj pochodzenia organizacji przyznającej, UE – Unia Europejska

- 10 Patel, S. B. and Bellini, M. (2008) *Nucleic Acids Research*, 36(20), pp. 6482–6493. PMID:18854356.
- 11 Hill, C. H. et al. (2019) Molecular Cell, 73(6), p. 1217–1231.e11. PMID:30737185.
- 12 Dominski, Z. and Marzluff, W. F. (2007) *Gene*, 396(2), pp. 373–390. PMID:17531405.
- Colgan, D. F. and Manley, J. L. (1997) Genes & development, 11(21), pp. 2755–66.
 PMID:9353246.
- 14 Proudfoot, N. J. (2011) Genes & development, 25(17), pp. 1770–82. PMID:21896654.
- *15* Munroe, D. and Jacobson, A. (1990) *Molecular and Cellular Biology*, 10(7), pp. 3441–3455. PMID:1972543.
- 16 Vicens, Q., Kieft, J. S. and Rissland, O. S. (2018) *Molecular Cell*, 72(5), pp. 805–812. PMID:30526871.
- 17 Marzluff, W. F. and Koreski, K. P. (2017) *Trends in Genetics*, 33(10), pp. 745–759
- *18* Lim, L. and Canellakis, E. S. (1970) *Nature*, 227(5259), pp. 710–2. PMID:5432070.
- 19 Edmonds, M., Vaughan, M. H. and Nakazato, H. (1971) *Proceedings of the National Academy* of Sciences of the United States of America, 68(6), pp. 1336–40. PMID:5288383.
- 20 Adesnik, M. *et al.* (1972) *Journal of molecular biology*, 71(1), pp. 21–30. PMID:4539023.
- 21 Fukamachi, S., Bartoov, B. and Freeman, K. B. (1972) *The Biochemical journal*, 128(2), pp. 299– 309. PMID:4404394.
- 22 Martin, G. and Keller, W. (2007) *RNA (New York, N.Y.)*, 13(11), pp. 1834–49. PMID:17872511.
- 23 Kuchta, K. et al. (2009) Nucleic Acids Research, 37(22), pp. 7701–7714. PMID:19833706.
- 24 Kim, S. H. *et al.* (1992) *The EMBO journal*, 11(6), pp. 2319–26. PMID:1534753.
- 25 Tanaka, N. and Schwer, B. (2005) *Biochemistry*, 44(28), pp. 9795–9803. PMID:16008364.
- 26 Shugar, D. (1996) Acta biochimica Polonica, 43(1), pp. 9–23. PMID:8790708.
- 27 Hirai, H. et al. (1988) Journal of biochemistry, 104(6), pp. 991–4. PMID:3243771.
- 28 Trippe, R. et al. (2006) RNA (New York, N.Y.), 12(8), pp. 1494–1504. PMID:16790842.
- 29 Warkocki, Z. et al. (2018) Cell, 174(6), p. 1537–1548.e29. PMID:30122351.
- 30 Heo, I. et al. (2008) Molecular Cell, 32(2), pp. 276–284. PMID:18951094.
- 31 Heo, I. et al. (2009) Cell, 138(4), pp. 696–708. PMID:19703396.
- *Hagan, J. P., Piskounova, E. and Gregory, R. I. (2009) Nature Structural & Molecular Biology,* 16(10), pp. 1021–1025. PMID:19713958.
- *33* Piskounova, E. *et al.* (2011) *Cell*, 147(5), pp. 1066–1079. PMID:22118463.
- 34 Heo, I. et al. (2012) Cell, 151(3), pp. 521–532. PMID:23063654.
- 35 Kim, B. *et al.* (2015) *The EMBO journal*, 34(13), pp. 1801–1815. PMID:25979828.
- 36 Thomas, M. P. et al. (2015) Cell Reports, 11(7), pp. 1079–1089. PMID:25959823.
- 37 Morgan, M. et al. (2017) Nature, 548(7667), pp. 347–351. PMID:28792939.
- 38 Morgan, M. et al. (2019) Cell Research. Nature Publishing Group, 29(3), pp. 221–232
- *Houseley, J. and Tollervey, D. (2009) Cell, 136(4), pp. 763–776. PMID:19239894.*
- 40 Łabno, A. et al. (2016) Nucleic Acids Research, 44(21), pp. 10437–10453. PMID:27431325.
- 41 Pirouz, M. et al. (2016) Cell Reports, 16(7), pp. 1861–1873. PMID:27498873.
- 42 Ustianenko, D. et al. (2016) The EMBO journal, 35(20), pp. 2179–2191. PMID:27647875.
- 43 Malecki, M. et al. (2013) The EMBO Journal, 32(13), pp. 1842–1854. PMID:23503588.
- 44 Lubas, M. et al. (2013) The EMBO Journal, 32(13), pp. 1855–1868. PMID:23756462.
- 45 Ustianenko, D. *et al.* (2013) *RNA (New York, N.Y.)*, 19(12), pp. 1632–1638. PMID:24141620.
- 46 Faehnle, C. R., Walleshauser, J. and Joshua-Tor, L. (2014) *Nature*, 514(7521), pp. 252–256. PMID:25119025.
- 47 Lv, H. *et al.* (2015) *Acta crystallographica. Section D, Biological crystallography*, 71(Pt 6), pp. 1284–94. PMID:26057668.
- 48 Chlebowski, A. et al. (2013) Biochimica et Biophysica Acta (BBA) Gene Regulatory Mechanisms, 1829(6–7), pp. 552–560. PMID:23352926.
- 49 Makino, D. L., Baumgärtner, M. and Conti, E. (2013) *Nature*, 495(7439), pp. 70–75. PMID:23376952.
- 50 Makino, D. L. et al. (2015) Nature, 524(7563), pp. 54–8. PMID:26222026.

- 51 Dziembowski, A. *et al.* (2007) *Nature Structural & Molecular Biology*, 14(1), pp. 15–22. PMID:17173052.
- 52 Liu, Q., Greimann, J. C. and Lima, C. D. (2006) *Cell*, 127(6), pp. 1223–1237. PMID:17174896.
- 53 Tomecki, R. *et al.* (2010) *The EMBO journal*, 29(14), pp. 2342–2357. PMID:20531386.
- 54 Bruni, F., Lightowlers, R. N. and Chrzanowska-Lightowlers, Z. M. (2017) *The FEBS Journal*, 284(12), pp. 1767–1777. PMID:27926991.
- 55 Szczesny, R. J. et al. (2012) Biochimica et Biophysica Acta (BBA) Gene Regulatory Mechanisms, 1819(9–10), pp. 1027–1034
- 56 Wang, D. D.-H. *et al.* (2009) *Journal of Biological Chemistry*, 284(31), pp. 20812–20821. PMID:19509288.
- 57 Borowski, L. S. et al. (2013) Nucleic Acids Research, 41(2), pp. 1223–1240. PMID:23221631.
- 58 Dhir, A. et al. (2018) Nature, 560(7717), pp. 238–242. PMID:30046113.
- 59 Warkocki, Z. et al. (2018) Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1762), p. 20180162. PMID:30397099.
- 60 Malecki, M. et al. (2013) The EMBO journal, 32(13), pp. 1842–1854. PMID:23503588.
- 61 Lardelli, R. M. et al. (2017) Nature Genetics, 49(3), pp. 457–464
- 62 Landrieux, E. *et al.* (2006) *The EMBO journal*, 25(1), pp. 118–28. PMID:16362040.
- 63 Turowski, T. W. and Tollervey, D. (2016) *Biochemical Society Transactions*, 44(5), pp. 1367–1375. PMID:27911719.
- 64 Kowalski, M. P. and Krude, T. (2015) *The International Journal of Biochemistry & Cell Biology*, 66, pp. 20–29. PMID:26159929.
- 65 Gopinath, S. C. B., Wadhwa, R. and Kumar, P. K. R. (2010) *Molecular Cancer Research*, 8(11), pp. 1536–1546. PMID:20881010.
- 66 Martignetti, J. A. and Brosius, J. (1993) *Proceedings of the National Academy of Sciences of the United States of America*, 90(24), pp. 11563–7. PMID:8265590.
- 67 Lin, D. et al. (2008) Molecular and Cellular Biology, 28(9), pp. 3008–3019
- 68 Shin, H. et al. (2018) Molecules and cells, 41(12), pp. 993–999. PMID:30590906.
- 69 Astuti, D. et al. (2012) Nature Genetics, 44(3), pp. 277–284. PMID:22306653.
- 70 Chang, H. et al. (2018) Molecular Cell, 70(1), p. 72--82.e7. PMID:29625039.
- 71 Cordaux, R. and Batzer, M. A. (2009) *Nature Reviews. Genetics*, 10(10), pp. 691–703. PMID:19763152.
- 72 Faulkner, G. J. and Garcia-Perez, J. L. (2017) *Trends in genetics: TIG*. PMID:28797643.
- 73 Lander, E. S. et al. (2001) Nature, 409(6822), pp. 860–921. PMID:11237011.
- 74 Beck, C. R. *et al.* (2010) *Cell*, 141(7), pp. 1159–1170. PMID:20602998.
- 75 Brouha, B. et al. (2003) Proceedings of the National Academy of Sciences of the United States of America, 100(9), pp. 5280–5285. PMID:12682288.
- 76 Martin, S. L. et al. (2003) Proceedings of the National Academy of Sciences of the United States of America, 100(24), pp. 13815–13820. PMID:14615577.
- 77 Khazina, E. *et al.* (2011) *Nature Structural & Molecular Biology*, 18(9), pp. 1006–1014. PMID:21822284.
- 78 Naufer, M. N. et al. (2016) Nucleic Acids Research, 44(1), pp. 281–293. PMID:26673717.
- 79 Feng, Q. *et al.* (1996) *Cell*, 87(5), pp. 905–916. PMID:8945517.
- 80 Mathias, S. L. *et al.* (1991) *Science (New York, N.Y.)*, 254(5039), pp. 1808–1810. PMID:1722352.
- 81 Alisch, R. S. et al. (2006) Genes & Development, 20(2), pp. 210–224. PMID:16418485.
- 82 Doucet, A. J. *et al.* (2015) *Molecular Cell*, 60(5), pp. 728–741. PMID:26585388.
- 83 Doucet, A. J. et al. (2010) PLoS genetics, 6(10). PMID:20949108.
- *S4* Jurka, J. (1997) *Proceedings of the National Academy of Sciences of the United States of America*, 94(5), pp. 1872–1877. PMID:9050872.
- 85 Monot, C. et al. (2013) PLoS genetics, 9(5), p. e1003499. PMID:23675310.
- 86 Chang, H. et al. (2014) Molecular Cell, 53(6), pp. 1044–1052. PMID:24582499.
- 87 Gregersen, L. H. *et al.* (2014) *Molecular Cell*, 54(4), pp. 573–585. PMID:24726324.

- 88 Smallwood, S. A. and Kelsey, G. (2012) Trends in Genetics, 28(1), pp. 33–42. PMID:22019337.
- 89 Figiel, M. et al. (2018) Journal of Biological Chemistry, 293(1), pp. 191–202. PMID:29122886.
- 90 Razew, M. et al. (2018) Nature Communications, 9(1). PMID:29311576.
- 91 Dziembowski, A. et al. (2003) The Journal of biological chemistry, 278(3), pp. 1603–11. PMID:12426313.
- 92 Malecki, M. et al. (2007) Journal of molecular biology, 372(1), pp. 23–36. PMID:17658549.
- *93* Dmochowska, A., Golik, P. and Stepien, P. P. (1995) *Current genetics*, 28(2), pp. 108–12. PMID:8590460.
- 94 Stepien, P. P. et al. (1992) Proceedings of the National Academy of Sciences of the United States of America, 89(15), pp. 6813–7. PMID:1379722.
- 95 Mian, I. S. (1997) Nucleic acids research, 25(16), pp. 3187–95. PMID:9241229.
- Jedrzejczak, R. et al. (2011) Acta crystallographica. Section D, Biological crystallography, 67(Pt
 11), pp. 988–96. PMID:22101826.
- 97 Szczesny, R. J. et al. (2012) Biochimica et Biophysica Acta (BBA) Gene Regulatory Mechanisms, 1819(9–10), pp. 1027–1034. PMID:22178375.
- 98 Borowski, L. S. et al. (2010) Biochimica et Biophysica Acta (BBA) Bioenergetics, 1797(6–7), pp. 1066–1070
- 99 Hahn, A. and Zuryn, S. (2019) Trends in Cell Biology, 29(3), pp. 227–240. PMID:30509558.
- 100 Lipinski, K. A., Kaniak-Golik, A. and Golik, P. (2010) *Biochimica et Biophysica Acta (BBA)* -*Bioenergetics*, 1797(6–7), pp. 1086–1098
- 101 Barbas, A. et al. (2009) Journal of Biological Chemistry, 284(31), pp. 20486–20498
- 102 Lorentzen, E. et al. (2008) Molecular Cell, 29(6), pp. 717-728
- 103 Fabrizio, P. et al. (2009) Molecular Cell, 36(4), pp. 593-608. PMID:19941820.
- 104 Warkocki, Z. et al. (2009) Nature Structural and Molecular Biology, 16(12), pp. 1237–1243. PMID:19935684.
- 105 Ohrt, T. et al. (2012) RNA, 18(6), pp. 1244–1256. PMID:22535589.
- 106 Ohrt, T. et al. (2013) RNA, 19(7)
- 107 Bertram, K. et al. (2017) Cell, 170(4), p. 701-713.e11. PMID:28781166.
- 108 Bertram, K. et al. (2017) Nature, 542(7641), pp. 318–323. PMID:28076346.
- 109 Haselbach, D. et al. (2018) Cell, 172(3), p. 454-464.e11. PMID:29361316.
- 110 Rauhut, R. et al. (2016) Science, 353(6306), pp. 1399–1405. PMID:27562955.
- 111 Wan, R. et al. (2019) Cell. PMID:30879786.
- 112 Bao, P. et al. (2017) RNA, 23(12), pp. 1770–1779. PMID:28864812.
- 113 Fourmann, J.-B. et al. (2013) Genes & Development, 27(4), pp. 413-428. PMID:23431055.
- 114 Rasche, N. et al. (2012) The EMBO Journal, 31(6), pp. 1591–1604. PMID:22246180.
- 115 Roy, J. et al. (1995) RNA (New York, N.Y.), 1(4), pp. 375–90. PMID:7493316.
- 116 Silverman, E. J. *et al.* (2004) *Molecular and cellular biology*, 24(23), pp. 10101–10. PMID:15542821.
- 117 Allen, E. et al. (2005) Cell, 121(2), pp. 207–221. PMID:15851028.
- 118 Axtell, M. J. et al. (2006) Cell, 127(3), pp. 565–577. PMID:17081978.
- 119 Yoshikawa, M. et al. (2005) Genes & Development, 19(18), pp. 2164–2175. PMID:16131612.

23

120 Nonomura, K.-I. (2018) Plant Reproduction, 31(1), pp. 21–29. PMID:29350289.