

Dr hab. Michał Ponczek
Uniwersytet Łódzki
Wydział Biologii i Ochrony Środowiska
Katedra Biochemii Ogólnej

Recenzja osiągnięcia naukowego „**Opracowanie metod i narzędzi analizy ilościowej i jakościowej kolistych RNA oraz ich wykorzystanie w różnych układach biologicznych**” oraz pozostałego dorobku naukowego, dydaktycznego, organizacyjnego oraz działalności popularyzującej naukę
dr Anny Philips dotycząca wniosku o nadanie stopnia doktora habilitowanego

Odpowiadając na decyzję Rady Doskonałości Naukowej o wyznaczeniu mnie do komisji habilitacyjnej jako recenzenta dr Anny Philips w postępowaniu w sprawie nadania stopnia doktora habilitowanego w dziedzinie nauk ścisłych i przyrodniczych w dyscyplinie nauki biologiczne w szczególności dnia 17 grudnia 2021 roku przedstawiam recenzję dorobku naukowego, działalności dydaktycznej i organizacyjnej oraz popularyzującej naukę. Recenzja będzie składać się z następujących części: odniesienia się do sylwetki naukowej habilitantki, oceny osiągnięcia naukowego, oceny pozostałego dorobku naukowego, oceny działalności dydaktycznej, organizacyjnej, popularyzującej naukę i wniosków końcowych.

1. Sylwetka naukowa

Dr Anna Philips uzyskała tytuł magistra w dyscyplinie informatyka (specjalność informatyka stosowana) w 2008 roku na Uniwersytecie im. Adama Mickiewicza w Poznaniu (UAM), Wydział Fizyki, Laboratorium Bioinformatyki. Promotorami i opiekunami pracy byli prof. Andrzej Dobek i dr Leszek Rychlewski, a tytuł pracy magisterskiej: „*Analiza strukturalna konformacji liganda w dokowaniu białko-ligand*” wskazuje na wczesne zainteresowania zjawiskami biologicznymi na poziomie molekularnym w powiązaniu z informatyką. Dalsze etapy rozwoju naukowego obejmowały studia doktoranckie na Wydziale Biologii UAM, Instytut Biologii Molekularnej i Biotechnologii pod kierunkiem prof. dr. hab. Janusza Bujnickiego zakończone obroną w 2013 roku wyróżnionej pracy doktorskiej „*Nowe metody bioinformatyczne do przewidywania miejsc wiązania metali i ligandów w strukturach RNA*”. Była dodatkowo zatrudniona w Instytucie Biologii Molekularnej i Komórkowej w Warszawie w latach 2008–2009 jako bioinformatyk, a w latach 2010–2013 w ramach grantu TEAM prof. dr. hab. J. Bujnickiego, gdzie zajmowała się modelowaniem oddziaływań ligandów z RNA. Pracuje jako adiunkt w Instytucie Chemii Bioorganicznej Polskiej Akademii Nauk w Poznaniu (ICHB PAN), od 2013 roku po uzyskaniu stopnia doktora, a od 2020 roku jest kierownikiem Pracowni Bioinformatyki w tejże jednostce naukowej.

2. Omówienie i ocena osiągnięcia naukowego

Osiągnięcie naukowe Dr Anna Philips stanowi cykl powiązanych tematycznie czterech artykułów naukowych opatrzonych wspólnym tytułem „**Opracowanie metod i narzędzi analizy ilościowej i jakościowej kolistych RNA oraz ich wykorzystanie w różnych układach biologicznych**” opublikowanych w języku angielskim

w czasopiśmie międzynarodowym o wskaźniku cytowań *Clarivate Analytics* od 3,26 do 6,6 w latach 2019-2021 (dwie prace w 2020 roku, po jednej pracy odpowiednio w 2019 i 2021 roku). Habilitantka jest pierwszą autorką dwóch z tych prac, w jednej autorką korespondencyjną. Prace te były cytowane bez autocytowań odpowiednio 2, 2, 0 i 14 razy wg bazy *Scopus* (dane z dnia 27.04.2022). Największe zainteresowanie innych naukowców wzbudziła praca z 2019 roku, dotycząca dystrofii miotonicznej u człowieka, a w pozostałych publikacjach badania skupiały się na roślinie rzodkiewnik pospolity (*Arabidopsis thaliana* (L.) Heynh) i jako najnowsze publikacje, dwie z 2020 i jedna z 2021 roku, nie zdążyły zaskarbić sobie jeszcze dużego międzynarodowego zainteresowania badaczy w postaci dużej liczby cytowań. Osiągnięcie naukowe było w większości finansowane i realizowane w ramach projektu „Identyfikacja kolistych RNA oraz białek uczestniczących w ich biogenezie u modelowej rośliny, *Arabidopsis thaliana*” przyznanego w ramach konkursu SONATA 8 Narodowego Centrum Nauki, panel: NZ2 i realizowanego jako kierownik od 2015-07-21 do 2019-07-20. Podstawowym celem badań sformułowanym przez autorkę, związanym z tematem osiągnięcia, było opracowanie zoptymalizowanego protokołu służącego do identyfikacji, a także analizy ilościowej i jakościowej kolistego RNA (circRNA). Kolejnym krokiem była weryfikacja opracowanej metody w różnych układach biologicznych: układzie roślinnym na przykładzie rzodkiewnika pospolitego - w siewce i w organach: korzeniu, liściu, kwiecie oraz u człowieka na przykładzie choroby mięśni o podłożu genetycznym o nazwie dystrofia miotoniczną typu I. Metoda została użyta do identyfikacji czynników genetycznych potencjalnie uczestniczących w procesie biogenezy circRNA w układzie roślinnym oraz przetestowana w tkance mięśniowej człowieka, a zaproponowany protokół postępowania służący do identyfikacji oraz analizy ilościowej i jakościowej circRNA okazał się wg autorki uniwersalny u organizmów roślinnych i zwierzęcych z szerokim polem stosowania niezależnie od testowanego układu biologicznego. Dalej Dr Anna Philips opisuje bardziej szczegółowo realizację swojego nowatorskiego pomysłu na zoptymalizowany protokół i analizę ilościową oraz jakościową circRNA opublikowaną w czterech publikacjach składających się na oceniane osiągnięcie naukowe. Pierwszym bardziej szczegółowym celem pracy była wiarygodna identyfikacja i charakterystyka circRNA w siewkach i organach (korzeniach, liściach, kwiatach) *A. thaliana* typu dzikiego, ekotyp Columbia (Col-0). Publikacje naukowców poświęcone circRNA w *A. thaliana* i innych roślinach wg autorki zawierały w dużej mierze wyniki niespójne oraz brak standaryzacji metod, gdy zapoznawała się z nimi na początku badań. Wymienia, że stosowano wówczas różne sposoby uzyskiwania danych z sekwencjonowania: poddając próbki liniowych transkryptów trawieniu RNazą R bądź nie, a podczas analiz bioinformatycznych z góry uznawano circRNA za zidentyfikowany po zaledwie dwóch odczytach NGS mapujących do miejsca tzw. „kolistego” splicingu (tzw. odczyty typu *back-spliced*), a także, że nie brano często pod uwagę głębokości biblioteki RNA-seq, na skutek czego liczba potencjalnych circRNA znacząco się między badaniami różniła, a zidentyfikowane na podstawie danych RNA-seq circRNA tylko w nielicznych przypadkach były poddawane walidacji eksperymentalnej za pomocą ilościowej reakcja łańcuchowa polimerazy (qPCR). Rozpoznając problemy innych badaczy, habilitantka na początku podjęła próbę oceny i optymalizacji protokołów do identyfikacji i analizy ilościowej circRNA, jeżeli chodzi o przygotowanie bibliotek RNA-seq, jaki i późniejszą analizę bioinformatyczną danych z sekwencjonowania, co opisała w pierwszej publikacji osiągnięcia: *Expression Landscape of circRNAs in Arabidopsis thaliana Seedlings and Adult Tissues*. Philips A. i wsp. *Frontiers in Plant Sciences*. 2020, 11: 576581. doi: 10.3389/fpls.2020.576581. Przetestowane zatem zostały dwa protokoły przygotowywania bibliotek NGS do badań circRNA: odpowiednio przy traktowaniu RNazą R i nietraktowaniu, a także kilka sposobów

normalizacji danych RNA-seq. Badane były circRNA w siewkach i organach *A. thaliana* typu dzikiego, z których izolowano całkowity RNA, by kolejno przygotować biblioteki RNA-seq przy użyciu przytoczonych dwóch protokołów. Odpowiednio do podejścia pierwszego, użyto RNA całkowitego po deplecji rRNA, a próbki zawierały liniowe i koliste transkrypty. Zgodnie z podejściem drugim, użyto RNA całkowite, zubożone o rRNA i potraktowano RNazą R trawiącą transkrypty liniowe w wyniku czego próbki zawierały circRNA oraz śladowe ilości liniowego RNA. Eksperymenty RNA-seq wykonane były w czterech powtórzeniach. Porównanie wyników z obu sposobów wykazało, że stosowanie RNazy R pozwala na identyfikację circRNA występujących rzadziej, które są gubione w standardowych bibliotekach z powodu dużej liczby transkryptów liniowych. Mimo, że biblioteki po trawieniu RNazą R posiadały średnią liczbę odczytów około połowę mniejszą niż biblioteki bez trawienia RNazą R to liczba zidentyfikowanych circRNA była trzykrotnie wyższa, co udowodniło, że stosowanie RNazy R istotnie rozdzielczość analizy zwiększyło. Z drugiej strony autorka zaobserwowała, że niektóre circRNA zidentyfikowane w czterech powtórzeniach biologicznych w bibliotekach nie trawionych RNazą R, ale sprawdzone za pomocą RT-PCR, nie zostały odnalezione w bibliotekach wzbogacanych w circRNA. Odpowiadać wg habilitantki za to miała prawdopodobnie nieenzymatyczna hydroliza niektórych transkryptów kolistych, co przekształciło je w liniowe, podatne na trawienie RNazą R. Ostatecznie zatem stwierdziła, że traktowanie próbek RNazą R wzbogaca circRNA, ale zarazem ta procedura może także powodować utratę informacji o niektórych circRNA. Podsumowując autorka stwierdziła, że podejście bez trawienia RNazą R jest lepsze do identyfikacji i ilościowej oceny całkowitej puli circRNA.

W celu porównania poziomów występowania circRNA w różnych częściach roślin, wymagane jest zastosowanie przeznaczonej dla circRNA metody normalizacji danych RNA-seq do czego autorka używała metodę ddPCR (ang. *droplet digital PCR*), aby wyznaczyć wartości referencyjne akumulacji wybranych circRNA i ustalić optymalny sposób normalizacji danych RNA-seq, a najczęściej stosowana metoda normalizacji przez ustalanie wielkości biblioteki okazała się zawodzić w przypadku analiz circRNA u roślin ponieważ dane były słabo skorelowane z wynikami ddPCR. W kolejnym kroku przetestowane zostały zatem cztery alternatywne metody normalizacji: poprzez liczbę odczytów mapujących do miejsca splicingowego typowego dla circRNA, poprzez współczynnik normalizacji oszacowany przez DESeq2, przez wielkość biblioteki bez odczytów rRNA i chloroplastowego RNA oraz poprzez poziom ekspresji genu ACT2. Zastosowanie tych sposobów umożliwiło osiągnięcie wartości współczynnika korelacji Pearsona pomiędzy wynikami ddPCR i RNA-seq w zakresie 0,76-0,77. Jedno-organowe analizy ilościowe circRNA dawały jeszcze wyższą wartość równą 0,92 dla liści. Wyniki przeprowadzonych badań potwierdziły pogląd, że skład transkryptomu ustalony na podstawie RNA-seq, a zaobserwowany w różnych organach rośliny ma wpływ na analizę ilościową rzadkich gatunków RNA, uwzględniając w tym circRNA. Uzyskane wg habilitantki wyniki wskazują na wysoką wiarygodność ilościowej charakterystyki circRNA w oparciu o RNA-seq w danym organie, przy założeniu użycia adekwatnej metody normalizacji.

Zastosowanie zoptymalizowanej przez autorkę strategii umożliwiło zidentyfikowanie i zbadanie akumulacji kolistych RNA w siewkach i trzech różnych organach - korzeniach, liściach i kwiatach *A. thaliana* typu dzikiego. Dodatkowo przeprowadzone zostało porównanie zawartości circRNA z różnych organów i porównanie ich nagromadzenia ze zbiorowiskiem odpowiadających im liniowych transkryptów. Umożliwiło to identyfikację circRNA powszechnych w całej roślinie, jak i specyficznych dla organów. Przeprowadzenie analizy umożliwiło identyfikację aż 5235 różnych circRNA opartych

przynajmniej na dwóch odczytach *back-spliced*. Już na początku analizy danych RNA-seq autorka zauważyła słabą powtarzalność circRNA w powtórzeniach biologicznych w przedziale zaledwie od 1,3 do 5,0%. Uznała to za dość zaskakujące odkrycie, ponieważ brak było wówczas publikacji naukowych na temat podobnej charakterystyki danych o roślinnych circRNA. Wyeliminowała ewentualny błąd eksperymentalny poprzez wprowadzenie dwóch dodatkowych powtórzeń biologicznych. Przeprowadzone także zostały analogiczne analizy circRNA w ludzkich liniach komórkowych Huh-7.5 oraz HeLa i te takiej zmienności nie wykazały. Dodatkowo przeprowadzona analiza publicznie dostępnych danych RNA-seq dla circRNA w *A. thaliana* także potwierdziła tę zaskakującą obserwację. Uzyskane wyniki wskazywały wyraźnie, że pula circRNA w *A. thaliana* zawierać musi wiele generowanych losowo cząsteczek, których funkcja fizjologiczna jest z pewnością nieistotna. Przeważająca część wykrytych circRNA (4297 circRNA z 5235 co daje 82,1%) została poparta przez 2 do 5 odczytów *back-spliced*. Autorka, jednakże scharakteryzowała także 39 kolistych RNA wspieranych przez liczbę większą niż 100 odczytów *back-spliced*. Wśród tych wszystkich prawdopodobnych circRNA 127 zostały wyszukane przez dwa niezależne programy, CIRI2 oraz find_circ oraz poparte przez przynajmniej 10 znormalizowanych odczytów *back-splice* RNA-seq, co odpowiadało około 50 kopiom circRNA na mikrogram RNA całkowitego. Tak restrykcyjna selekcja, wg autorki, pozwoliła na wyodrębnienie frakcji circRNA występujących w roślinie w sposób powtarzalny, w ilościach pozwalających na pełnienie przez te cząsteczki jakichś ról w komórce. Zaproponowana przez Dr Annę Philips metoda umożliwia, wobec tego nie tylko na identyfikację i analizę ilościową, ale także na wyodrębnienie circRNA produkowanych w sposób stochastyczny, nie pełniących funkcji, od tych produkowanych w sposób powtarzalny, o prawdopodobnej funkcji biologicznej.

Liczba powtarzalnych circRNA różniła się w badanych organach roślinnych, od 87 w korzeniach, 79 w liściach, 99 w kwiatach do 109 circRNA w siewkach, a 127 circRNA zostało zidentyfikowanych jako produkowanych w sposób powtarzalny, 52 (40,9%) z nich było obecnych we wszystkich badanych organach rośliny, 112 (88,2%) z nich zostało już znalezionych w badaniach wcześniejszych i umieszczonych w bazie danych PlantcircBase (Chu et al. 2017), a 67 (52,8%) w bazie AtCircDB (Ye et al. 2017). Analiza ontologii genów (ang. *gene ontology*, GO) wykryła, że geny dające początek powtarzalnym circRNA cechują się znacznym występowaniem takich, które są zaangażowane w fotosyntezę i odpowiedź na stres. Korzenie wykazywały największą liczbę unikatowych circRNA (6) i najwięcej circRNA o dużej akumulacji ($p \leq 0,05$) porównując z innymi organami. Autorka najmniejszą liczbę powtarzalnych circRNA o zwiększonym poziomie stwierdziła w siewkach. Dalsze jej analizy koncentrowały się na circRNA obecnych w co najmniej dwóch organach. Porównanie par organów pokazało specyficzny dla nich wzór nagromadzenia circRNA. Największa liczba circRNA o zwiększonym poziomie ($p \leq 0,05$) w porównaniu z innymi tkankami, charakteryzowała korzenie (16 circRNA vs siewki, 25 vs liście i 9 vs kwiaty). Najmniejsza liczba circRNA o podniesionym poziomie została stwierdzona w siewkach (1 vs korzenie, 3 vs liście i 1 vs kwiaty). Analizy wykazały także obecność circRNA pochodzących z genomu chloroplastowego w siewkach (24), korzeniach (2), liściach (12) i kwiatach (14). W innych cytowanych przez autorkę publikacjach chloroplasty zostały zidentyfikowane jako „hot spoty” produkcji circRNA. Uzyskane przez dr Annę Philips wyniki wykazały, że w liściach faktycznie przeważały odczyty typu *back-spliced* wspierające circRNA pochodzące z genomu chloroplastów. Autorka dalej stwierdza, że „nadal nie wyjaśniono, w jaki sposób tak zasadniczo różne procesy, jak splicing jądrowy i organellarny, mogłyby prowadzić do tworzenia circRNA. Kwestię tę komplikuje jeszcze fakt, że niektóre organellarne mRNA są produktami trans splicingu, w którym eksony dwóch pre-mRNA zostają połączone. Niektóre odczyty pochodzące

z takich mRNA mogą być błędnie zaklasyfikowane jako odczyty typu back-spliced. W związku z powyższym chloroplastowe circRNA zidentyfikowane w danych RNA-seq powinny być traktowane z dużą dozą ostrożności a ich obecność powinna zostać potwierdzona przy pomocy innych metod". Przeanalizowanie ekspresji mRNA klarownie dowiodło, że nie ma bezpośredniego powiązania między akumulacją circRNA a akumulacją ich odpowiedników liniowych. Dalej autorka stwierdza „Wybrane przykłady, poparte metodą ddPCR pokazały, że poziom niektórych circRNA wzrasta wraz ze wzrostem ekspresji ich liniowych odpowiedników, jednak ogólnie nie ma korelacji między akumulacją tych dwóch frakcji transkryptomu (współczynnik korelacji Pearsona: 0,14). To odkrycie wskazało, na duże prawdopodobieństwo istnienia specyficznego mechanizmu regulującego powstawanie circRNA u *A. thaliana*. Co więcej, poziom niektórych circRNA był zwiększony w poszczególnych tkankach, podczas gdy poziomy ich liniowych odpowiedników pozostawały względnie stabilne w całej roślinie. Można zatem spekulować, że biogenezę tych circRNA reguluje specyficzny tkankowo czynnik molekularny, który zmienia „wzór” splicingu.”

Następnym etapem badań habilitantki było poszukiwanie czynników uczestniczących w biogenezie circRNA. Wcześniej wykazano, że circRNA u roślin są generowane w sposób specyficzny dla komórek organów roślinnych, a poziomy akumulacji nie są skorelowane z poziomami odpowiadających mRNA, umożliwiło postawienie hipotezy, że występują nieznane czynniki, które promują wytwarzanie circRNA. W celu identyfikacji takich czynników użyte zostały licznie dostępne mutanty *A. thaliana* z dobrze scharakteryzowanymi nokautami pojedynczego genu w zestawieniu z typem dzikim Col-0. Autorka dokonała selekcji wariantów, tak aby każdy nosił mutację typu knockout w genie kodującym białko biorące udział w splicingu na różnych etapach tego procesu. Według *Arabidopsis Splicing Related Data Base Genes* (ASRG), białka związane ze splicingiem u *A. thaliana* występują u 395, ale zestaw mutantów został zawężony do tych z tłem Col-0, tak by było możliwe uzyskanie wspólnego odniesienia dla wszystkich analizowanych roślin. Wykluczone zostały heterozygoty i mutacje letalne. Wstępna selekcja wariantów umożliwiła opracowanie listy 86 pozycji ze zmianami w genach związanych ze splicingiem, a 23 z nich miało dokładnie opisany fenotyp. Ostatecznie możliwe stało się uzyskanie listy 18 mutantów *A. thaliana* wymienionych w Tabeli S1 z drugiej publikacji osiągnięcia (*Arabidopsis thaliana cbp80, c2h2, and flk Knockout Mutants Accumulate Increased Amounts of Circular RNAs*, Phillips A. i wsp. Cells. 2020, 9: 1937. doi: 10.3390/cells909193). Skład transkryptomów wszystkich tych roślin i Col-0 został określony za pomocą RNA-seq, a identyfikacja circRNA przeprowadzona przy pomocy opracowanego wcześniej przez autorkę protokołu.

Dr Anna Phillips zaobserwowała, że produkcja circRNA u trzech mutantów różniła się od obserwowanej w typie dzikim i pozostałych mutantach. Pierwszy wariant nie miał funkcjonalnego genu *cbp80*, którego produktem jest element kompleksu wiążącego czapczkę na końcu 5' mRNA. Mutant drugi, *c2h2*, nie posiadał genu kodującego białko U4/U6.U5 tri-snRNP – składnika wstępnie złożonego kompleksu spliceosomowego, ważnego w tworzeniu katalitycznie aktywnego spliceosomu. Mutant trzeci, *flk*, nie miał genu produkującego białko z grupy jądrowych rybonukleoprotein (hnRNP), które, jeśli są związane z cząsteczkami pre-mRNA, informują, że pre-mRNA nie jest jeszcze w pełni dojrzały i gotowy do eksportu do cytoplazmy. Cechą wspólną tych trzech mutantów było zwiększone gromadzenie kolistych transkryptów w porównaniu z typem dzikim i innymi mutantami, której nie towarzyszyły odpowiednie zmiany w schemacie ekspresji genów. Zaobserwowana zwiększona akumulacja circRNA mogła być spowodowana częstszym back-splicingiem lub większą stabilnością circRNA u tych mutantów. Autorka zaobserwowała wzrost liczby wystąpień alternatywnego splicingu w rzeczonych trzech mutantach, co sugeruje, że nokaut genu miał wpływ na splicing.

Mutant *cbp80* wytwarzał bardzo dużą liczbę 129 unikatowych circRNA, a spośród nich, tylko 11 było wspólnych z Col-0 na 35 circRNA zidentyfikowanej w tym typie. Białko CBP80 wraz z CBP20 tworzy kompleks zaangażowany w splicing oraz biogenezę miRNA, zatem mutanty *A. thaliana* pozbawione tego kompleksu lub jego składników kumulują częściowo złożone transkrypty i pri-miRNA. Ponieważ braki w kompleksie wiążącym czapeczkę, a w szczególności nieobecność CBP80, wpływają na powstawanie zjawiska alternatywnego splicingu, szczególnie na końcu 5' pierwszego intronu, badania autorki po raz pierwszy dostarczyły dowodów na to, że brak CBP80 nie tylko zaburza produkcję liniowych izoform transkryptów, ale także znacząco zwiększa wytwarzanie różnych circRNA. Większość circRNA specyficznych dla mutantu *cbp80* zawierała pierwszy ekson, podczas gdy circRNA powstałe w innych mutantach i Col-0 zostały wyprodukowane głównie z drugiego lub kolejnych eksonów i obserwacja ta jest zgodna z wcześniejszymi odkryciami, że CBP80 preferencyjnie wywiera swój wpływ na splicing w mRNA w okolicy końca 5', a białko CBP80 bierze także udział w transdukcji sygnału kwasu abscysynowego (ABA) i szlaku kwitnienia. Jak pisze dalej autorka: „Mutanty pozbawione CBP80 wykazują wywołane przez ABA podwyższenie cytozolowego poziomu wapnia w komórkach, prowadzące do lepszego zamykania aparatów szparkowych zapewniając przez to większą tolerancję na suszę. Te rośliny wykazują również fenotyp wczesnego kwitnienia, wynikający z wadliwego splicingu intronów FLOWING LOCUS C (FLC). Niestety nie wiadomo czy circRNA generowane w mutantach *cbp80* również przyczyniają się do zaobserwowanych fenotypów. Repertuar circRNA specyficznych dla nokautów *cbp80* musi zostać zbadany bardziej szczegółowo w celu przedstawienia ostatecznych wniosków dotyczących tego aspektu. CircRNA mogą być generowane w sposób kontrolowany lub stanowić produkty nieprawidłowego splicingu.” W oparciu o te wyniki badań stawia hipotezę, że „ogólnie niezakłócone funkcjonowanie czynników białkowych zaangażowany w splicing sprzyja tworzeniu liniowych transkryptów”. Dodatkowo stwierdza, że „otrzymane wyniki sugerują, że zmiana kompleksów białkowych związanych ze splicingiem może być łatwym mechanizmem uruchamiającym/wzmacniającym produkcję circRNA”. Dynamiczny skład kompleksu białek wiążących czapeczkę w *A. thaliana* jest dobrze udokumentowany i ulega zmianie w trakcie wzrostu rośliny, a CBP80 jest jednym z białek, które pozostają związane z czapeczką mRNA zarówno w komórkach proliferujących, jak i w spoczynku, ale ulegają zmianie lokalizacja subkomórkowej. Pisząc dalej: „Na wczesnych etapach cyklu wzrostu białko to lokalizuje się preferencyjnie w cytoplazmie, a później gromadzi się w jądrze. Z kolei jego partner, CBP20, najpierw lokalizuje się zarówno w jądrze, jak i w cytoplazmie, a następnie koncentruje się w jądrze” stawia hipotezę, że „niektóre wystąpienia niekanonicznego splicingu na początku cyklu wzrostu rośliny mogą być spowodowane niedoborem CBP80”. Stwierdzając „Takie zjawisko potencjalnie może aktywować tworzenie circRNA. Idąc dalej, kontrola występowania różnorodnych czynników związanych ze splicingiem może stanowić potencjalny mechanizm regulacji ekspresji genów w roślinach. Co ciekawe, pokazano już, że ekspresja circRNA wzrasta, gdy działanie maszynarii odpowiedzialnej za splicing pre-mRNA u *Drosophila* jest upośledzone.” autorka sugeruje zakonserwowanie takiego mechanizmu w różnych królestwach.

Opracowana w pracy pierwszej cyklu metoda analizy danych RNA-seq do identyfikacji i oceny ilościowej circRNA w roślinach została użyta do powtórnego przebadania wszystkich publicznie dostępnych zbiorów danych RNA-seq dla *A. thaliana*, które spełniały założone przez autorkę kryteria opracowanego protokołu. Mimo, że circRNA u *A. thaliana* oceniano w wielu badaniach, porównanie wyników wykazywało na rozbieżności przez co analiza porównawcza nie była możliwa. Główną przyczyną, wg autorki, takiego stanu rzeczy był brak standaryzacji w metodach stosowanych do analiz circRNA. Ujednolicone wyniki otrzymane z analizy

publicznych danych zostały zatem wraz z panią Katarzyną Nowis, doktorantką habilitantki, zaimplementowane i udostępnione w Internecie szerszemu gronu odbiorców poprzez utworzenie nowej bazy danych At-C-RNA (<http://plantcircrna.ibch.poznan.pl/>), a informacja o tym i opis został przedstawiony w publikacji trzeciej cyklu - *At-C-RNA database, a one-stop source for information on circRNAs in Arabidopsis thaliana in a unified format*. Nowis K, Jackowiak P, Figlerowicz M, Philips A. Database (Oxford), 2021, 00, 1–5, doi: 10.1093/database/baab07. Baza ta po raz pierwszy umożliwiła analizę i porównanie danych o circRNA z różnych źródeł. Wg Anny Philips At-C-RNA to obecnie jedyne źródło, w którym dane pochodzące z różnych projektów badawczych zostały ponownie przeanalizowane przy pomocy takiej samej metody, w ten sposób ujednolicone i ustandaryzowane. At-C-RNA jest obecnie jedyną bazą danych dostarczającą informacje o występowaniu circRNA w dzikich oraz zmutowanych roślinach *A. thaliana* i ich organach. Baza została zaprojektowana, aby interfejs był przyjazny użytkownikowi, a informacje o circRNA prezentowane są w tabeli interaktywnej umożliwiającej zdefiniowanie kryteriów przeglądania użytkownikowi, dając możliwość sortowania i filtrowania danych według każdej z kolumn oraz możliwe jest również ręczne usuwanie wybranych rekordów circRNA z tabeli, a także filtrowanie tabeli po kliknięciu na znajdujące się pod nią interaktywne wykresy.

Baza At-C-RNA posiada obecnie zdeponowanych 113 327 circRNA i jest możliwość pobrania ich w całości jako plik *.xlsx formatu Microsoft Excel (dane podane przez autorkę oraz sprawdzone przeze mnie dnia 7.04.2021), a jedynie niewielka część z nich została umieszczona wcześniej w bazach poświęconych roślinnym circRNA: AtCircDB, PlantcircNet, PlantcircBase, odpowiednio 16,2%, 18,9% i 19,7%.

Każdy circRNA w bazie ma przypisany unikalny identyfikator zgodnie z ogólnie przyjętym wzorcem: AT_numer_chromosomu:circRNA_start-circRNA_stop, co umożliwia łatwe porównywanie z wynikami innych baz. Dla każdego circRNA przypisane są informacje takie jak: dane literaturowe, na podstawie których wygenerowano zestawy danych RNA-seq, (ii) linia roślinna i organ, w której zidentyfikowano circRNA, (iii) średnia punktacja At-C-RNA wyliczona dla wszystkich circRNA zgodnie z taką samą procedurą wyrażającą poziom akumulacji circRNA, punktacja At-C-RNA poszczególnych circRNA wchodzących w skład punktacji średniej, informacja o tym, czy dany circRNA został potwierdzony eksperymentalnie oraz informacja na temat „powtarzalności” circRNA. To ostatnie kryterium jest wg autorki szczególnie istotne, ponieważ we wcześniejszych badaniach wykazała, że większość circRNA produkowana jest w *A. thaliana* w sposób przypadkowy i niepowtarzalny, a przypuszczalnie i w innych roślinach także, nie pełniąc prawdopodobnie żadnej funkcji biologicznej. Autorka zakłada, że jedynie circRNA produkowane w sposób powtarzalny mogą mieć potencjał funkcjonalny i są istotnymi obiektami dla dalszego szerokiego zakresu badań.

Baza At-C-DB uznaje, że circRNA jest powtarzalny, jeśli został zidentyfikowany w przynajmniej 4 biologicznych powtórzeniach w określonej linii rośliny i organie, a sumarycznie 655 circRNA jest sklasyfikowanych jako powtarzalne, gdzie 226 jest zidentyfikowanych we wszystkich badanych organach roślinnych (kwiecie, liściu, korzeniu i siewce) oraz w całej roślinie. Najwięcej unikatowych, powtarzalnych circRNA w liczbie 35 znaleziono w liściu, a korzeń i kwiat, charakteryzowały się odpowiednio tylko dwoma i jednym circRNA typowym dla danego organu, w siewce nie zidentyfikowano unikatowego circRNA. 362 genów dających początek powtarzalnym circRNA wytwarza ponad pięć izoform circRNA, a zaledwie 11 genów wytwarza jedną izoformę. Punktacja At-C-DB 91,6% circRNA mieści się w zakresie 1-15, co odpowiada niskiej liczebności, a 8,4% circRNA przekracza średnią ocenę powyżej 15. Lokalizacja powtarzalnych circRNA na chromosomach pokazuje że większość pochodzi z genów chromosomu 1, w zgodzie z obserwacjami dokonanymi

we wcześniejszej pierwszej pracy z cyklu. Analiza danych z różnych źródeł umożliwia ukazanie całego obrazu rozkładu circRNA w *A. thaliana* i jej organach.

Dodatkowym bardzo ważnym i godnym uznania aspektem badań dr Anny Philips, wskazującym na istotną wagę naukową opracowanych rozwiązań, było zastosowanie metody analizy circRNA w *A. thaliana* do zbadania circRNA ludzi chorych na dystrofię miotoniczną typu 1 (DM1), gdzie wyniki zostały opublikowane w czwartej pracy osiągnięcia naukowego: *Global Increase in Circular RNA Levels in Myotonic Dystrophy*. Czubak K, Taylor K, Piasecka A, Sobczak K, Kozłowska K, Philips A, Sedehizadeh S, Brook JD, Wojciechowska M, Kozłowski P. *Frontiers in Genetics*. 2019, 10:649. doi: 10.3389/fgene.2019.00649. DM1 jest wywoływana ekspansją powtórzeń CTG w regionie 3' UTR genu DMPK kodującego enzym kinazę miotoniczną (ang. *myotonin-protein kinase*). Osoby zdrowe cechuje od 5 do ~34 powtórzeń, podczas gdy u chorych z DM1 występują powtórzenia sięgające do setek, a nawet tysiące kopii powtórzeń CTG. Zmutowane transkrypty RNA z wydłużonymi powtórzeniami CTG występujące w dystrofii miotonicznej wiążą się specyficznie z białkami *muscleblind* (MBNL), co zaburza funkcję MBNL w regulacji alternatywnego składania pre-mRNA. MBNL odgrywają także rolę w biogenezie circRNA, zatem, ponieważ ich działanie w DM1 przez wiązanie długich sekwencji CTG jest zaburzone, autorka sformułowała hipotezę, że także poziom niektórych circRNA może być obniżony w tej jednostce chorobowej. W celu falsyfikacji postawionej hipotezy habilitantka skorzystała z publicznie dostępnych zestawów danych RNA-Seq zdeponowanych w bazie danych dystrofii miotonicznej: DMseq (<http://www.dmseq.org/>). Analiza została przeprowadzona jedynie dla danych, które zostały wygenerowane z zastosowaniem takiej samej procedury, gdzie biblioteki były przygotowywane z RNA całkowitego pozbawionego rRNA oraz gdzie stosowano sekwencjonowanie Illumina typu pair-end. Przebadane zostały dwa zestawy danych reprezentowane najliczniej w bazie DMseq. Próbkę do badań pochodziły z dwóch różnych lokalizacji w ramach mięśni szkieletowych człowieka: mięśnia czworogłowego uda (lat. *Musculus quadriceps femoris*, QF; 23 próbki, 11 próbek kontrolnych i 12 próbek DM1) i mięśnia piszczelowego przedniego (lat. *Musculus tibialis anterior*, TA; 27 próbek, 6 próbek kontrolnych i 21 próbek DM1). Opierając się na wynikach uzyskanych wcześniej, poziomy akumulacji circRNA zostały znormalizowane względem liczby odczytów mapujących do genów ACTB i GAPDH.

Próbkę z QF zawierały 22816 różnych circRNA, z czego 4168 (18%) zostało sklasyfikowanych jako powtarzalne, potwierdzone przez co najmniej pięć odczytów typu *back-splice* w co najmniej dwóch próbkach, a 152 (0,7%) zaliczone jako powszechne, obecne we wszystkich lub we wszystkich oprócz jednej próbkach kontrolnych DM1. Próbkę z TA posiadały 38403 circRNA, z czego 7537 (20%) powtarzalnych i 403 (1%) powszechnych. Autorka stwierdziła, że „Porównanie globalnego poziomu akumulacji wszystkich circRNA w kontroli i w próbkach DM1, ujawniło istotnie **zwiększony poziom circRNA w DM1**”. Podobne różnice nie zostały zauważone dla ich liniowych odpowiedników. Podwyższony poziom circRNA w próbkach DM1 był również obserwowalny w grupach „powszechnych” circRNA. Znormalizowany został również poziom circRNA w stosunku do odpowiadających im liniowych transkryptów w celu wykluczenia, że obserwowane zmiany poziomu circRNA są wynikiem wzrostu lub spadku transkrypcji prekursorów liniowych. Wyniki wskazywały także **wyższe poziomy circRNA** w próbkach DM1 niż w próbkach kontrolnych. Dodatkowo, przy użyciu wygenerowanych danych, autorka przeanalizowała różnicową akumulację circRNA w DM1. Analiza tą została ograniczona jedynie do zbiorów circRNA występujących powszechnie we wszystkich lub we wszystkich próbkach kontrolnych DM1 (n = 152 w QF i n = 403 w TA). Zmiany poziomu circRNA obliczone w oparciu o normalizację obiema metodami były silnie skorelowane zarówno dla QF, jak i TA, wykazując

brak zależności poziomu circRNA od poziomu ekspresji genów, z których są generowane. Wartości zmian akumulacji okazały się być znacznie przesunięte w kierunku wartości dodatnich, wskazujący na większą ilość circRNA o podwyższonym poziomie w próbkach DM1. Efekt ten jest zgodny ze stwierdzonym wcześniej całkowitym wzrostem poziomu circRNA w DM1, zarówno dla QF, jak i dla TA. Przeprowadzona analiza wykazała zatem, że globalny poziom circRNA w DM1 jest **podwyższony**. Stosując protokół analizy danych, szczególnie, stosując różne metody normalizacji (względem liczby odczytów mapujących do genów ACTB i GAPDH; względem globalnej liczby mapujących odczytów; oraz względem liczby odczytów mapujących do liniowych odpowiedników circRNA) wykazano, że wzrost poziomu circRNA w DM1 nie jest wynikiem zmian globalnego poziomu transkrypcji czy podwyższenia ekspresji ich odpowiedników liniowych. Autorka zatem **sfalsyfikowała wstępnie postawioną hipotezę** i stwierdziła, że otrzymane w badaniu wyniki poddają w wątpliwość rolę MBNL jako czynnika odgrywającego ważną rolę w biogenezie circRNA u człowieka. Zauważyła także, że zwiększony globalny poziom circRNA jak i podniesiony poziom indywidualnych circRNA w DM1 okazał się bardzo dobrze korelować z ostrością przebiegu choroby i w związku z tym daje to możliwość badań w kierunku użycia circRNA jako biomarkerów w DM1. Osobiście uważam za bardzo ciekawe i ważne spostrzeżenie z tych badań.

Habilitantka sugeruje, że jej opracowana metoda analizy danych RNA-seq do badań circRNA u *A. thaliana* może zostać rozszerzona na badania circRNA w innych układach biologicznych, co częściowo sama zaprezentowała w ostatniej pracy na przykładzie DM1 u człowieka. Powtarzalne, specyficzne tkankowo circRNA mogą być badane dalej w bardziej zaawansowanych układach funkcjonalnych, a uzyskane wyniki powinny ułatwić planowanie i prowadzenie kolejnych badań nad circRNA specyficznymi dla narządów, organów czy tkanek roślinnych i zwierzęcych wraz z odkrywaniem ich funkcji, jak i być miejscem startu ku odkryciu mechanizmów sterujących przełączaniem między tworzeniem się transkryptów kolistych i liniowych.

Mutanty *A. thaliana* o zmienionych fenotypach mogłyby być skutecznie używane do badań funkcjonalnych circRNA, a w przypadku niedoboru circRNA u mutantu typu *knockout* wprowadzonego sztucznie możliwe byłyby natychmiastowe obserwacje wywołanych taką ingerencją efektów fenotypowych. Dodatkowo otrzymane wyniki mogą przyczynić się do lepszego zrozumienia i identyfikacji nowych mechanizmów regulacji ekspresji genów, ponieważ rozpoznawanie nieznanymi czynnikami zaangażowanymi w splicing i związanych z produkcją circRNA jest istotne dla lepszego zrozumienia mechanizmów leżących u podstaw niekanonicznego łączenia eksonów.

Baza danych At-C-RNA będzie niewątpliwie bardzo użyteczna dla naukowców z różnych krajów do analizy i wykorzystania wyników badań wygenerowanych przez grupy krajowe i zagraniczne, będąc obecnie jedynym źródłem, w którym dane z różnych eksperymentów zostały ponownie przeanalizowane, ustandaryzowane oraz ujednoczone.

Badanie wpływu niedoborów czynników uważanych za biorące udział w biogenezie niektórych circRNA może przyczynić się do rozwoju dalszych badań oraz sprawdzenia powiązania z produkcją circRNA, w tym również związku generowania tych cząsteczek z białkami z rodziny MBNL w chorobach takich jak DM1 u człowieka.

Podsumowanie osiągnięcia naukowego

Osiągnięcie naukowe zaprezentowane we wniosku przez dr Annę Philips obejmuje cykl publikacji naukowych, które zostały wykonane w ramach grantu SONATA NCN, którego była kierownikiem. Pierwszym etapem badań była identyfikacja circRNA u *A. thaliana*. Opracowała metodę identyfikacji i standaryzacji danych circRNA, która później została użyta do powtórnego przebadania danych dostępnych publicznie, w wyniku czego utworzona została nowa baza danych (<http://plantcircrna.ibch.poznan.pl/>), umożliwiającą analizę i porównanie danych o circRNA z przeróżnych źródeł. Wykazała także, że większość circRNA u *A. thaliana* jest produkowana losowo i zaproponowała metodę umożliwiającą odróżnienie przypadkowych circRNA od tych generowanych w sposób powtarzalny, przy założeniu, że tylko takie cząsteczki powinny pełnić funkcje biologiczne. Dalej, przy użyciu utworzonych narzędzi, pracowała nad identyfikacją czynników genetycznych uczestniczących w biogenezie circRNA. Wykazała, że mutanty rośliny pozbawione genów *cbp80*, *c2h2* i *flk* gromadzą więcej circRNA, co ma wskazywać na udział produktów tych genów w regulacji ekspresji circRNA. Pokazała także, że niedobór czynnika uważanego za powiązany z powstawaniem circRNA u ludzi - białko z rodziny *muscleblind*, nie koreluje z ograniczeniem poziomu circRNA w warunkach *in vivo* - w tkankach mięśni (poprzecznie prążkowanych) osób chorych na dystrofię miotoniczną typu 1, gdzie poziom circRNA był raczej zwiększony. Uznała podsumowując, że „wiarygodna identyfikacja i oznaczenie ilościowe circRNA u roślin oraz identyfikacja czynników zaangażowanych w ich biogenezę przyczyni się do właściwego zrozumienia uniwersalnych zasad, które rządzą powstawaniem tych RNA w różnych królestwach, oraz znaczenia circRNA w szerokim kontekście ewolucyjnym”.

Podsumowując, w pracy dr Anny Philips uważam za bardzo cenną falsyfikację wstępnie postawionej hipotezy (ostatnia publikacja osiągnięcia) i stwierdzenie, że otrzymane w badaniu wyniki poddają w wątpliwość rolę MBNL jako czynnika odgrywającego ważną rolę w biogenezie circRNA u człowieka oraz zwrócenie uwagi na to, że zwiększony globalny poziom circRNA jak i podniesiony poziom indywidualnych circRNA w DM1 okazuje się bardzo dobrze korelować z ostrością przebiegu DM1, co daje możliwość badań w kierunku użycia circRNA jako biomarkerów w DM1. Podtrzymuję na koniec swoje stwierdzenie, że jest to bardzo ciekawe i ważne spostrzeżenie badań habilitacyjnych autorki w jej osiągnięciu.

Ocena pozostałego dorobku naukowego

Dr Anna Philips w trakcie kariery naukowej zajmowała się badaniami zastosowania i rozwoju różnych narzędzi bioinformatycznych w celu poznawania i zrozumienia zjawisk i procesów biologicznych na poziomie molekularnym. Pracowała między innymi nad stworzeniem metod komputerowych, które umożliwiałyby przewidywanie miejsca wiązania jonów metali i różnych ligandów w strukturach makromolekuł biologicznych takich jak białka oraz RNA. Rozpoczynając przygodę z nauką na studiach magisterskich na wydziale Fizyki UAM, pod kierunkiem pana dr. Leszka Rychlewskiego, opracowała meta-metodę oceniającą różne konformacje przestrzenne kompleksu białko-ligand w oparciu o cząstkowe oceny uzyskiwane z sześciu programów do dokowania. Ostateczna ocena opracowanej meta-metody, nazwanej HarmonyDock dawała lepsze wyniki (mniejszą wartość RMSD do struktury natywnej) niż każdy z programów dokujących osobno. Wyniki badań zostały opisane w publikacji “: *HarmonyDOCK: The structural analysis of poses in protein-ligand docking*”. Plewczynski D, Philips A, Von Grotthuss M, Rychlewski L, Ginalski K. *Journal of Computational Biology*. 2014 21: 247-256.

doi: 10.1089/cmb.2009.0111 cytowanej do tej pory wg baz *Elsevier Scopus* oraz *Clarivate Web of Science* 6 razy bez autocytowań.

Badania nad strukturami makromolekuł, kontynuowała w trakcie studiów doktoranckich na Wydziale Biologii Uniwersytetu Adama Mickiewicza w Poznaniu, pod kierunkiem prof. dr. hab. Janusza Bujnickiego. Uzyskała grant PRELUDIUM 2, panel NZ2 Narodowego Centrum Nauki by realizować jako kierownik projekt „Nowa metoda bioinformatyczna do predykcji miejsc wiązania ligandów w strukturach RNA” od: 2012-09-18 do 2013-09-17. Jako doktorantka opracowała anizotropowy potencjał statystyczny (potencjał anizotropowy bazujący na wzajemnych odległościach atomów RNA i jonu metalu, atomów liganda, jak i na ich wzajemnej orientacji w przestrzeni), charakteryzujący się dobrą pewnością przez zastosowanie parametru odległości między atomami oraz parametru orientacji przestrzennej – kąta. Potencjał zastosowany został dalej w programach *MetalionRNA* oraz *LigandRNA*. Pierwszy z nich jest algorytmem służącym do przewidywania miejsc wiązania jonów metali w strukturach RNA, drugi umożliwia ocenę kompleksów RNA-ligand. Narzędzia umożliwiają badania oddziaływań między RNA a małymi cząsteczkami i mogą być używane do znajdowanie nowych, potencjalnych leków. Wyniki badań uzyskiwane w trakcie studiów doktoranckich zostały upublicznione poza pracą doktorską w postaci czterech publikacji w renomowanych czasopismach naukowych (przed doktoratem: Philips A, Milanowska K, Lach G, Boniecki M, Rother K, Bujnicki JM. *MetalionRNA: computational predictor of metal-binding sites in RNA structures*. *Bioinformatics*. 2012, 28: 198-205. doi: 10.1093/bioinformatics/btr636; Cruz JA i wsp. Philips A 16/32 autor. *RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction*. *RNA*. 2012 18: 610-625. doi: 10.1261/rna.031054.111; Philips A, Milanowska K, Lach G, Bujnicki JM. *LigandRNA: computational predictor of RNA-ligand interactions*. *RNA*. 2013, 19: 1605-1616. doi: 10.1261/rna.039834.113; po doktoracie: Philips A, Lach G, Bujnicki JM. *Computational methods for prediction of RNA interactions with metal ions and small organic ligands*. *Methods in Enzymology*. 2015;553:261-85. doi: 10.1016/bs.mie.2014.10.057). Rozprawa doktorska „Nowe metody bioinformatyczne służące do przewidywania miejsc wiązania jonów metalu i niskocząsteczkowych ligandów w strukturach RNA” została obroniona z wyróżnieniem i nagrodzona przez Polskie Towarzystwo Bioinformatyczne jako najlepszy doktorat z bioinformatyki w 2013 roku, a 22 listopada 2013 roku Rada Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu przyznała stopień naukowy doktora w dziedzinie nauk biologicznych w zakresie biologii, specjalność bioinformatyka.

Pracując w ICHB PAN w Poznaniu jako adiunkt głównym nowym tematem zainteresowań badawczych stały się koliste cząsteczki RNA w roślinach stanowiące przedmiot osiągnięcia naukowego. (Habilitationka niepotrzebnie w części „4. OMÓWIENIE POZOSTAŁYCH OSIĄGNIĘĆ NAUKOWO-BADAWCZYCH”, jeszcze raz zbyt szczegółowo opisują tę część działalności naukowej, opisaną wcześniej wystarczająco w części „3. OMÓWIENIE OSIĄGNIĘCIA autoreferatu.”)

Innym zainteresowaniem, poza tym w ramach osiągnięcia naukowego, była tematyka mikrobiomów, najpierw w ramach badania składu bakterii w kopalnych szczątkach ludzkich wykazując, że poza występującymi powszechnie bakteriami środowiskowymi, możliwa jest także identyfikacja kopalnych szczepów, typowych dla jamy ustnej człowieka. Dr Anna Philips opracowała metodę identyfikacji tych bakterii, a także oceny ich wieku za pomocą analiz modyfikacji i zniszczeń DNA. Wykazała, że czas przechowywania w magazynie, wiek

czy stanowisko archeologiczne na jakim zostały pobrane nie mają wpływu na zawartość ilościową zachowanych kopalnych próbek.

Porównawcze badania genomów bakteryjnych starożytnych i współczesnych mogą dostarczyć ważnych danych o ewolucji i pochodzeniu mikroorganizmów, w tym informacje o czynnikach zjadliwości i adaptacji do żywiciela, jakim może być człowiek, w tym także przy zmianie diety i nawyków higienicznych. Wyniki z tej tematyki zostały opublikowane w trzech artykułach (wg autoreferatu: N5, N7, N9) oraz w dwóch monografiach (wg autoreferatu: N1, N2). Dodatkowo habilitantka nawiązała współpracę z grupą z Uniwersytetu w Louisville (UofL), Kentucky, USA, w ramach badania patogenów jamy ustnej i dzięki tej kooperacji zrekonstruowała cztery kopalne genomy *Tannerella forsythia* i porównała je ze genomami współczesnych *T. forsythia*, co wykazało różnice sekwencji niektórych genów związanych z wirulencją tej bakterii (wg autoreferatu publikacja N5).

Obecnie pracuje jako kierownik Pracowni Bioinformatyki w ICHB PAN i dalej prowadzi badania nad mikrobiomami. Dodatkowo wspólnie z partnerem biznesowym, firmą Ardigen SA, otrzymała grant Narodowego Centrum Badań i Rozwoju „Mapa mikrobiomu Polski” na opracowanie referencyjnego zbioru mikrobiomów przewodu pokarmowego Polaków. Repozytorium to ma być startem do badań nad zależnościami między składem bakteryjnym jelit a stylem życia, dietą, chorobami oraz odpowiedzią pacjentów na immunoterapię w onkologii.

Utworzenie tego repozytorium dającego perspektywę poznania sieci powiązań pomiędzy mikrobiomem a cechami fenotypowymi i środowiskowymi takimi jak styl życia, a także chorobami i skutecznością terapii immunologicznej może ułatwić zrozumienie wpływu mikrobiomu jelit na organizm człowieka. Habilitantka w autoreferacie podaje, że metodyka z tych badań została opracowana w postaci zgłoszenia patentowego („Sposób izolowania kwasów nukleinowych drobnoustrojów z kału”, patent krajowy zgłoszony dnia 26.02.2021 (nr zgłoszenia P.437097), współdzielony z ICHB PAN: Luiza Handschuh, Katarzyna Tomela, Magdalena Rakoczy, Anna Samelak-Czajka, UP: Marcin Schmidt, Ardigen SA: Emilia Strycharz-Angrecka, Kaja Milanowska-Zabel, Jan Majta, Agata Szymanek, Bożena Augustyn) oraz jak podaje dalej wyniki mają być udostępnione w postaci dwóch publikacji, z których jedna znajduje się w recenzji czasopisma *Scientific Reports*, a druga jest w przygotowaniu.

Prowadząc badania interdyscyplinarne we współpracy z chemikami, fizykami, biologami i lekarzami z kraju i z zagranicy zaangażowała się dodatkowo w kilka projektów badawczych w których służy wiedzą i doświadczeniem w wyborze, projektowaniu i interpretacji wyników analiz bioinformatycznych. Współpracując z grupą z Wydziału Biologii Uniwersytetu im. A. Mickiewicza w Poznaniu opublikowane zostały dwie publikacje (wg autoreferatu: N4, N6), w których przeanalizowany został bakteryjny skład ścieków przed, w trakcie i po oczyszczaniu dzięki czemu udało się wykazać różnice w trakcie oczyszczania. Badania te wykazały z jednej strony, że oczyszczanie ścieków zmniejsza całkowitą liczbę bakterii, ale z drugiej w ściekach po oczyszczeniu znajduje się proporcjonalnie więcej szczepów antybiotykoopornych i że poprzez oczyszczanie ścieków może dochodzić do transmisji genów związanych z antybiotykoopornością. Jest to kolejna bardzo ważna obserwacja z badań habilitantki skłaniająca do dalszych pytań, czy jest możliwe zapobieganie temu zjawisku, w sytuacji dostawania się masowo stosowanych antybiotyków i ich metabolitów poprzez odpady komunalne z miejsc zamieszkania ludności stosującej antybiotyki, szpitali być może także innych instytucji.

W ramach uzupełnienia do „Informacji o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej można wskazać”, że:

- Przeprowadzała badania w ramach trzech polskich jednostek badawczych: Uniwersytet im. Adama Mickiewicza w Poznaniu, Międzynarodowy Instytut Biologii Molekularnej i Komórkowej w Warszawie oraz ICHB PAN Poznań, gdzie od 2014 roku pracuje.
- Od 2018 roku współpracuje z grupą prof. Jana Potempy z *Department of Oral Immunity and Infectious Diseases, University of Louisville School of Dentistry, USA*, gdzie jak podaje planowany był staż, jednak ze względu na pandemię COVID-19 jest przekładany na okres późniejszy.
- Jako wykonawca interdyscyplinarnego grantu SYMFONIA NCN (kierownik prof. dr hab. M. Figlerowicz, ICHB PAN – lider) współpracowała z pozostałymi jednostkami wchodzącymi w skład konsorcjum realizującego grant, tj. Instytutem Antropologii UAM oraz grupą prof. Jasińskiego z Wydziału Historii UAM. Prowadzone badania zaowocowały szeregiem wspólnych publikacji (wg autoreferatu: N1, N2, N5, N7, N9), wystąpień konferencyjnych oraz wspólną organizacją konferencji naukowych. Czynny wkład w realizację grantu polegał na zastosowaniu nowoczesnych metod obliczeniowych do uzyskania odpowiedzi na zagadnienia historyczne takie jak badania kopalnego DNA w celu udzielenia odpowiedzi na pytanie o pochodzenie genetyczne populacji z terenów Polski, a zastosowane metody wniosły nową jakość badań interdyscyplinarnych do badań antropologicznych i historycznych na polskich uczelniach.
- Od 2019 roku współpracuje stale z grupą pani prof. dr hab. Joanny Mokrackiej (Wydział Biologii UAM), a w wyniku tej współpracy opublikowane zostały dwa artykuły naukowe (wg autoreferatu: N4 i N6) w ramach analizy składu bakteryjnego ścieków przed, w trakcie i po oczyszczaniu. Dr Nicoletta Makowska, pierwsza autorka obu prac po obronie swojej pracy doktorskiej na UAM uzyskała grant SONATINA NCN, który realizuje w zespole habilitantki.

Brała udział w 7 (1 poza krajem) konferencjach naukowych po uzyskaniu stopnia doktora i w 11 (2 poza krajem) konferencjach przed uzyskaniem stopnia doktora.

Punkt: „Informacja o odbytych stażach w instytucjach naukowych lub artystycznych, w tym zagranicznych, z podaniem miejsca, terminu, czasu trwania stażu i jego charakteru” zawiera informacje o aktywności związanej z wizytami naukowymi krajowymi i zagranicznymi przed uzyskaniem stopnia naukowego doktora:

„2010-2012 Międzynarodowy Instytut Biologii Molekularnej i Komórkowej, Warszawa. Realizacja grantu FNP & OP IE 1.1.2 TEAM/2009-4/2 „Modeling of RNA and protein-RNA complexes: from sequence to structure to function”.

2007 - 2008 BioInfoBank Institute, Warszawa. Staż, realizowany projekt badawczy: „The structural analysis of poses in protein-ligand docking”.

2006 - 2007 Vrije Universiteit, Amsterdam, Holandia - Wyjazd w ramach program Socrates-Erasmus.

03-09.2006 ITTI e-technologie i biznes, Poznań. Asystent. Staż.

07.2005 Joint Institute Of Nuclear Research (JINR), Dubna, Russia - praktyki.
12.2003 Joint Institute Of Nuclear Research (JINR), Dubna, Russia - praktyki."

Niestety brak jest takich aktywności po uzyskaniu stopnia doktora.

Autorka podaje sumaryczną liczbę cytowań wg bazy *Clarivate Web of Science Core Collection* (bez autocytowań) z dn. 19.11.2021 na 253, do dnia 4.04.2022 ta wartość wzrosła do 284 dla 16 prac, choć indeks Hirsha o wartości 7 pozostał bez zmian. Baza *Elsevier Scopus* notuje obecnie ponad 300 cytowań i indeks Hirsha 8 dla 16 prac.

3. Ocena działalności dydaktycznej i organizacyjnej

Po uzyskaniu stopnia doktora w latach 2014-2021:

Habilitantka opiekowała się dwoma doktorantami. Osoby te były zatrudnione podczas realizacji wymienionego wcześniej grantu SONATA NCN, gdzie była kierownikiem, a po uzyskaniu stopnia doktora habilitowanego zakłada, że zostanie promotorem tych osób. Była także promotorem dwóch prac magisterskich: temat pracy: „Identyfikacja mikroorganizmów towarzyszących kopalnym materiałom kostnym”, UAM (współautorska publikacja w *GigaScience*, wg autoreferatu N9), oraz temat pracy: „Porównywanie dostępnych metod do identyfikacji kolistego RNA na podstawie danych z RNA-Seq”. Obie osoby studiowały kierunek Bioinformatyka, 2016 i planują studia doktoranckie. Poza tym w ramach działalności dydaktycznej recenzowała także jeden licencjat: na kierunku Bioinformatyka, UAM, 2016 i opiekowała się dwoma stażystami.

Przed uzyskaniem stopnia doktora opiekowała się dwoma projektami licencjackimi UAM na kierunku Bioinformatyka, 2011.

Działalność organizacyjna obejmowała:

- Od 2020 roku kierowanie Pracownią Bioinformatyki w ICHB PAN.
- Kierowanie projektami finansowanymi ze źródeł zewnętrznych wymienionych wcześniej.
- W 2021 uczestniczenie w zespołach przygotowujących materiały do oceny instytutów PAN i ewaluacji jednostek naukowych, prowadzonej przez MEiN (habilitantka podaje MNiSW) w celu ewaluacji.
- W 2018 uczestniczenie w komisji oceniającej pracowników ICHB PAN.
- Współorganizowanie czterech konferencji: „*Fascinating World of Bioorganic Chemistry*”, konferencja międzynarodowa dedykowana prof. A. Legockiemu z okazji 80 lecia, organizowana przez ICHB PAN, Poznań, 12-13.11.2019; „*Piast & Přemyslovci Meeting*”, Poznań, 29-30.05.2018, konferencja międzynarodowa organizowana przez ICHB PAN, UAM (Wydział Historyczny) oraz Instytut Archeologii Czeskiej Akademii Nauk w Pradze; „DNA – język życia”, Poznań, 27-28.10.2016, konferencja krajowa organizowana przez ICHB PAN i PP.; „Tradycje i nowoczesność - początki państwa polskiego na tle środkowoeuropejskim w badaniach interdyscyplinarnych”, Poznań, 10-12.06.2015 - konferencja międzynarodowa, organizowana przez Instytut Historii i Instytut Prahistorii UAM oraz ICHB PAN.

4. Ocena działalności popularyzującej naukę
Habilitantka nie podała żadnych aktywności w ramach działalności popularyzującej naukę.

5. Wnioski końcowe

Podsumowując całkowity dorobek naukowy habilitantki (osiągnięcie naukowe w postaci 4 naukowych publikacji oryginalnych jak i pozostała aktywność) jest on niewątpliwie zdecydowanie dobry na obecnym etapie kariery, wskazując dodatkowo na liczne zainteresowania i umiejętności badawcze w zakresie nie tylko bioinformatyki oraz jej zastosowania w badaniach biologii molekularnej, ale także w aspekcie interdyscyplinarnym szeroko pojętych nauk przyrodniczych prowadzonych na całym ziemskim globie, co ma także odzwierciedlenie w postaci nieco poniżej 300 cytowań przez innych naukowców oraz w wartości indeksu Hirsha 8 w ciągu 10 lat okresu publikacyjnego badaczki (uwzględniając 16 recenzowanych artykułów naukowych w bazach *Scopus* i *Web of Science*).

O ile habilitantka wykazywała aktywność związaną z wyjazdami zagranicznymi w celach staży naukowych praktycznie w całym okresie studenckim, o tyle niestety brak jest takiej aktywności po uzyskaniu stopnia doktora.

Dotychczasowe doświadczenie w ramach działalności dydaktycznej uznaję za wystarczające w ramach współpracy z UAM, przy uwzględnieniu pracy w Polskiej Akademii Nauk, po osiągnięciu stopnia doktora, instytucji bardziej nakierowanej na aktywność naukową niż dydaktyczną. Nie mam zastrzeżeń do zgłoszonej działalności organizacyjnej habilitantki. Pewien niedosyt sprawia brak podania aktywności w ramach działalności popularyzującej naukę, którą bardzo często można powiązać z działalnością dydaktyczną i organizacyjną. Zauważyłem kilka literówek w autoreferacie, ale w żaden sposób nie uznaję tego jako coś, co mogłoby umniejszyć moją ocenę dorobku habilitantki.

Biorąc pod uwagę pozytywną ocenę osiągnięcia naukowego - cykl czterech publikacji naukowych, oraz pozostałego dorobku naukowego, a także dotychczasową działalność dydaktyczną i działalność organizacyjną, stwierdzam, iż w mojej ocenie Pani dr Anna Philips spełnia ustawowe wymogi stawiane kandydatom do stopnia naukowego doktora habilitowanego (art. 221 ust. 4 i 5 ustawy z dn. 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2020 r. poz 85. z późn. zm.)). Wnioskuje zatem o dopuszczenie do dalszych etapów zmierzających do nadania stopnia doktora habilitowanego.

Łódź, 27 kwietnia 2022


Michał Ponczek