

Podsumowanie osiągnięć zawodowych

Podjęcia omiczne rzucające światło na ewolucję i rozwój zwierząt

Dr Guillem Ylla Bou

Kierownik Pracowni Bioinformatyki i Biologii Genomu

Wydział Biochemii, Biofizyki i Biotechnologii

Uniwersytet Jagielloński

Gronostajowa 7, Pokój L.01.40

30-387 Kraków, Polska

Zastrzeżenie: Oryginalny dokument został dostarczony przez autora w języku angielskim. Niniejszy dokument jest tłumaczeniem na język polski sporządzonym przez tłumacza doświadczonego w tej dziedzinie.

Spis treści

| | | |
|-------|---|----|
| 1. | Imię i nazwisko..... | 3 |
| 2. | Posiadane dyplomy, stopnie naukowe lub artystyczne | 3 |
| 3. | Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych..... | 3 |
| 4. | Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce | 3 |
| 4.1 | Tytuł..... | 4 |
| 4.2 | Wykaz artykułów naukowych powiązanych tematycznie należących do głównego osiągnięcia naukowego..... | 4 |
| 4.3 | Opis dorobku naukowego | 7 |
| 4.3.1 | Wprowadzenie | 7 |
| 4.3.2 | Cele badawcze | 12 |
| 4.3.3 | Najważniejsze osiągnięcia naukowe | 12 |
| 4.3.4 | Dalsze kierunki badań | 22 |
| 4.3.5 | Bibliografia..... | 24 |
| 5. | Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej. | 28 |
| 6. | Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę..... | 29 |
| 6.1 | Osiągnięcia organizacyjne | 29 |
| 6.2 | Osiągnięcia dydaktyczne i mentorskie..... | 29 |
| 6.2.1 | Wykładowca na Harvardzie GENED-1004: | 29 |
| 6.2.2 | Wykłady gościnne | 30 |
| 6.2.3 | Mentoring naukowy..... | 30 |
| 6.2.4 | Wykładowca kursu online | 30 |
| 6.3 | Popularyzacja nauki | 31 |
| 7. | Oprócz kwestii wymienionych w pkt. 1-6, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej..... | 32 |

1. Imię i nazwisko

Guillem Ylla Bou

2. Posiadane dyplomy, stopnie naukowe lub artystyczne

2018 Doktorat w dziedzinie biomedycyny (Biomedicine), Institute of Evolutionary Biology, CSIC - Pompeu Fabra University, Barcelona.

- Tytuł rozprawy doktorskiej: Comparative transcriptomics of hemimetabolan and holometabolan metamorphosis.
- Promotorzy: Xavier Belles, Ph.D. & M. Dolors Piulachs, Ph.D.

2014 Magister analizy danych omicznych (M.Sc. in Omics Data Analysis), University of Vic, Katalonia, Hiszpania

2012 Licencjat w dziedzinie biotechnologii (B.Sc. in Biotechnology), University of Vic, Katalonia, Hiszpania

3. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych

2022- Kierownik zespołu, Pracownia Bioinformatyki i Biologii Genomu. Uniwersytet Jagielloński, Kraków, Polska

2019-2021 Staż podoktorski (postdoc), Insects Genomics and Evolution, Organismic and Evolutionary Biology, Harvard University, Cambridge, MA.

Faculty Advisor: Cassandra Extavour, Ph.D.

Projekt: Annotation and comparison of Cricket genomes.

2018-2019 Staż podoktorski (postdoc), Bioinformatics, Microbiology & Cell Science, University of Florida, Gainesville, FL.

Faculty Advisor: Ana Conesa, Ph.D.

Projekt: The regulation of alternative splicing.

4. Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce

Prezentowany dorobek naukowy obejmuje **6 powiązanych tematycznie publikacji naukowych w recenzowanych czasopismach**. Suma *Impact Factor* (IF) publikacji w 2022 roku wynosi 38,5, a łączna liczba cytowań przekracza 50, mimo że 4 publikacje ukazały się w ciągu zaledwie ostatnich 2 lat. Według punktów przyznawanych czasopismom przez

Polskie Ministerstwo Nauki i Szkolnictwa Wyższego, te 6 publikacji zgromadziło łącznie 660 punktów.

4.1 Tytuł

Podejścia omiczne rzucające światło na ewolucję i rozwój zwierząt

4.2 Wykaz artykułów naukowych powiązanych tematycznie należących do głównego osiągnięcia naukowego

Przegląd 6 publikacji składających się na główne osiągnięcia naukowe przedstawiono w **tabeli 1**. Dalsze dane bibliograficzne dotyczące każdej publikacji oraz opis mojego wkładu przedstawiono poniżej. Oznaczenia przy nazwach w opisie publikacji oznaczają: *Równy wkład merytoryczny, ☒ autorzy korespondencyjni, § członkowie Ylla-lab.

Tabela 1: Przegląd 6 wybranych publikacji, w tym: tytuł, czasopismo, podejście badawcze, rok wydania, rodzaj znaczącego wkładu (1 = pierwszy autor, Cor. = autor korespondencyjny), *impact factor* (w 2020 r.), punkty ministerialne i liczba cytowań.

| Nr | Tytuł | Czasopismo | Podejście badawcze | Rok | Rola | IF (2022) | Punkty ministerialne (12/2021) | Cytowania (04/2023) |
|--------------|---|--|-------------------------------------|------|--------|-----------|--------------------------------|---------------------|
| #1 | Insights into the genomic evolution of insects from cricket genomes | Communications Biology | Genomika | 2021 | 1+Cor. | 6,55 | 20 | 41 |
| #2 | Reduction of embryonic E93 expression as a key factor for the evolution of insect metamorphosis | PNAS | mRNA | 2023 | Cor | 12,78 | 200 | 1 |
| #3 | Diversity of piRNA expression patterns during the ontogeny of the German cockroach | Journal of Experimental Zoology, Part B | piRNA | 2018 | 1 | 2,37 | 70 | 8 |
| #4 | siRNA enrichment in Argonaute 2-depleted <i>Blattella germanica</i> | Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms | siRNA | 2021 | Cor | 6,3 | 100 | 4 |
| #5 | Distinct gene expression dynamics in germ line and somatic tissue during ovariole morphogenesis in <i>Drosophila melanogaster</i> . | G3: Genomes, Genes, Genetics | mRNA | 2021 | 1 | 3,54 | 70 | 1 |
| #6 | MirCure: A tool for quality control, filter, and curation of microRNAs of animals and plants. | Bioinformatics | miRNA+ Tworzenie oprogramowania. | 2020 | 1 | 6,93 | 200 | 3 |
| Suma: | | | | | | 38,5 | 660 | 58 |

Publikacja nr 1

G. Ylla✉, T. Nakamura, T. Itoh, R. Kajitani, A. Toyoda, S. Tomonari, T. Bando, Y. Ishimaru, T. Watanabe, M. Fuketa, Y. Matsuoka, A. Burnett, S. Noji, T. Mito✉, & C. G. Extavour✉ (2021). [Insights into the genomic evolution of insects from cricket genomes](#). *Nature Communications Biology*.

Jak opisano w sekcji dotyczącej wkładu autora artykułu, zaprojektowałem doświadczenia i ich analizę, uzyskałem wyniki, napisałem pierwszą wersję artykułu i wykonałem wszystkie ilustracje. Oznacza to, że opracowałem narzędzie do adnotacji genomu, którego użyłem do wykonania adnotacji *de novo* dwóch genomów. Obejmowało to identyfikację powtarzających się elementów, mapowanie wszystkich dostępnych danych sekwencji RNA do genomu, wykonywanie modeli genów *ab initio* i generowanie modeli genów opartych na dowodach. Oprócz adnotacji genomu przeprowadziłem całą porównawczą analizę genomową i uzyskałem wszystkie wyniki zaprezentowane na opublikowanych ilustracjach. Biorąc pod uwagę znaczący wkład w ten artykuł, jestem pierwszym autorem i współautorem korespondencyjnym.

Publikacja nr 2

A. Fernandez-Nicolas*, G. Machaj*§, A. Ventos-Alfonso*, V. Pagone, T. Minemura, T. Ohde, T. Daimon, **G. Ylla**✉ & X. Belles✉ (2023) [Reduction of embryonic E93 expression as a key factor for the evolution of insect metamorphosis](#). *Proceedings of the National Academy of Sciences*

Mój wkład w powstanie tego artykułu rozpoczął się na etapie koncepcji projektu, który opierał się na moich wcześniejszych wynikach. Na ich podstawie przyczyniłem się do stworzenia nowego projektu badawczego, którego zwińczeniem była niniejsza publikacja. Po zaprojektowaniu eksperymentów i etapów analizy danych brałem udział w analizie danych, a także pełniłem funkcję mentora i przewodnika dla studentów i doktorantów zaangażowanych w różne aspekty projektu. Ponadto brałem udział w interpretacji i dyskusji wyników, w przygotowaniu ilustracji i tworzeniu manuskryptu. Ze względu na moją wiodącą rolę w całym projekcie jestem współautorem korespondencyjnym tego artykułu.

Publikacja nr 3

N. Llonga*, **G. Ylla***, J. Bau, X. Belles & MD. Piulachs✉ (2018): [Diversity of piRNA expression patterns during the ontogeny of the German cockroach](#). *Journal of Experimental Zoology, Part B*.

W ramach tego projektu brałem udział w opracowywaniu jego koncepcji oraz projektowaniu pozyskiwania danych eksperymentalnych i analizie danych bioinformatycznych. Nadzorowałem, prowadziłem i byłem mentorem studenta, który przeprowadził zaprojektowaną przeze mnie analizę. Uczestniczyłem również w uzyskaniu i dyskusji wyników, przygotowaniu ilustracji i pisaniu pierwszej wersji manuskryptu. Z uwagi na mój istotny wkład jestem pierwszym współautorem artykułu.

Publikacja nr 4

J.C. Montañés, C. Rojano, **G. Ylla**✉, MD. Piulachs✉, L. Maestro✉ (2021): [siRNA enrichment in Argonaute 2-depleted *Blattella germanica*](#). *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*.

Projekt zrodził się z moich wcześniejszych obserwacji dotyczących analizy siRNA przy użyciu sekwencjonowania małych RNA. W oparciu o te obserwacje brałem udział w opracowaniu nowego projektu badawczego, skupiającego się na badaniu biogenezy siRNA, łączącego eksperymenty biologii molekularnej i analizę danych z sekwencjonowania małych RNA. Oprócz udziału w projektowaniu eksperymentu, określiłem również strategię analizy danych oraz kierowałem i nadzorowałem studenta, który ją stosował. Brałem udział w przygotowaniu ilustracji i redagowaniu artykułu. Z uwagi na mój znaczący wkład jestem współautorem korespondencyjnym artykułu.

Publikacja nr 5

H. Tarikere*, **G. Ylla***, C. Extavour✉ (2021). [Distinct gene expression dynamics in germ line and somatic tissue during ovariole morphogenesis in *Drosophila melanogaster*](#). *G3: Genomes, Genes, Genetics*.

Mój główny wkład w ten projekt polegał na przeprowadzeniu kompletnej analizy danych z dużego zbioru danych transkryptomicznych, na której opiera się artykuł. Oprócz zaprojektowania i wdrożenia analizy danych brałem udział w dyskusji wyników, pisaniu pierwszej wersji artykułu oraz tworzeniu ilustracji od 3 do 5. Z uwagi na ten znaczący wkład jestem pierwszym współautorem artykułu.

Publikacja nr 6

G. Ylla, T. Liu & A. Conesa✉ (2020): MirCure: [A tool for quality control, filter, and curation of microRNAs of animals and plants](#). *Bioinformatics*.

W pełni zaprojektowałem i wdrożyłem narzędzie MirCure, które zostało opublikowane w tym artykule. Byłem również mentorem studenta, który dołączył do projektu celem przeprowadzenia testów porównawczych i walidacji narzędzia. Napisałem pierwszą wersję manuskryptu i przygotowałem ilustracje. Tym samym, biorąc pod uwagę moją główną rolę w projekcie, jestem pierwszym autorem niniejszej publikacji.

4.3 Opis dorobku naukowego

4.3.1 Wprowadzenie

4.3.1.1 Historyczny przegląd badań ewolucyjnych

Dziedzina biologii ewolucyjnej na długo poprzedza Karola Darwina i Alfreda Russella Wallace'a. Pośród cywilizacji w całej historii i na całym świecie istnieją legendy i religie próbujące wyjaśnić pochodzenie ludzi i innych istot zwierzęcych. Wiele z nich zawierało już elementy ewolucji, rozumianej jako zmiany form zwierzęcych w czasie. Faktycznie, niektóre z tych pomysłów były bardzo żywe w XIX-wiecznej wiktoriańskiej Anglii i całej Europie. Lekarz i filozof przyrody Erazm Darwin, dziadek Karola Darwina, napisał nawet wiersz o ewolucji zatytułowany „Świątynia natury” (ang. „*The temple of nature*”), opublikowany pośmiertnie w 1803 r. Wiersz ten zawiera już niektóre elementy teorii *ewolucji drogą doboru naturalnego* - lub jak wolał to ująć Wallace: *ewolucji przez przetrwanie najlepiej przystosowanych* (Darwin Correspondence Project, 2023), którą pół wieku później, w 1859 roku, C. Darwin i A. R. Wallace przedstawią w eseju na spotkaniu w Linnean Society (Darwin i Wallace, 1858), ustanawiając kamień węgielny dla współczesnej biologii ewolucyjnej.

Następna wielka fala znaczącego postępu w dziedzinie biologii ewolucyjnej nastąpiła, gdy C. Darwin, metaforycznie, spotkał się z G. Mendelem. Choć żyli w tych samych czasach, nigdy się nie spotkali i upłynął cały wiek, zanim ich prace zostały zunifikowane tworząc tak zwaną „współczesną syntezę ewolucyjną” na początku XX wieku (Berry i Browne, 2022). Ta nowa synteza powstała ręką biologów ewolucyjnych, genetyków i matematyków, którzy połączyli teorię ewolucji C. Darwina i A. R. Wallace'a o przetrwaniu najlepiej przystosowanych z obserwacjami G. Mendla dotyczącymi dziedziczenia cech. Dzięki współczesnej syntezie biologia ewolucyjna stworzyła ramy umożliwiające powiązanie genetyki z ewolucją, wywołując rewolucję naukową w dziedzinie biologii i sposobie, w jaki ją rozumiemy.

Prawdopodobnie znajdujemy się obecnie w trzeciej fali postępu w biologii ewolucyjnej, a także w wielu innych dziedzinach biologicznych, wspomaganych przez rozwój technologii sekwencjonowania (Marx, 2013). Po raz pierwszy w historii co roku publikowane są tysiące nowych genomów. W tym samym czasie technologie sekwencjonowania RNA i sekwencjonowania RNA pojedynczej komórki pozwalają nam uzyskiwać statyczny obraz stanu transkryptomycznego komórek i tkanek na różnych etapach lub w różnych warunkach, które mogą być również zaburzone wykorzystując techniki biologii molekularnej, a następnie określić, w jaki sposób perturbacje te wpłynęły na krajobraz transkryptomiczny. Niemniej jednak wąskim gardłem jest obecnie przekształcanie tej ogromnej ilości danych w informacje za pomocą analizy bioinformatycznej. Dodatkowa trudność polega na tym, że każdy rodzaj danych można badać i analizować na wiele sposobów, w zależności od pytania naukowego lub hipotezy do sprawdzenia, dlatego często nie ma jeszcze dostępnych narzędzi służących do

zadawania wszelkich możliwych pytań. Z tego powodu tworzenie nowych narzędzi (**publikacja nr 6**) często stanowi ważne zadanie dla laboratoriów bioinformatycznych, takich jak moje laboratorium.

4.3.1.2 Owady jako modele doświadczalne

Badania nad owadami były kluczem do rozwoju współczesnej syntezy. Laureat Nagrody Nobla, Thomas Hunt Morgan, swoimi badaniami nad muszką owocową *Drosophila melanogaster* położył podwaliny pod nowoczesną genetykę i stworzył fundamenty współczesnej syntezy ewolucyjnej. Inni podążyli za Morganem w kwestii wykorzystania *D. melanogaster* i innych owadów jako organizmów badawczych do badania genetyki i ewolucji, w tym kilku twórców współczesnej syntezy, takich jak Theodosius Dobzhansky, „Henry” Ford i Ernest Mayr.

Wykorzystanie *D. melanogaster* jako organizmu badawczego stało się filarem genetyki i badań ewolucyjnych przez resztę XX wieku aż do czasów obecnych, czyniąc z niego *par excellence* model badań genetycznych. W 2000 roku genom *D. melanogaster* był pierwszym opublikowanym genomem owada (Adams i in., 2000). Od tego czasu, jak wyjaśniamy w niedawnym przeglądzie (Mito i in., 2022), liczba dostępnych genomów wzrosła wykładniczo, przekraczając 700 genomów na początku 2022 r. Podczas gdy liczba genomów gatunków owadów wzrosła wykładniczo, liczba rzędów owadów z co najmniej jednym zsekwencjonowanym genomem nie rosła w tym samym tempie. Innymi słowy, 87,5% genomów należy do gatunków holometabolicznych (owadów z przeobrażeniem zupełnym), a w ich obrębie 89,3% genomów należy tylko do 3 rzędów, którymi są Diptera, Lepidoptera i Hymenoptera. Oznacza to, że nasza wiedza o biologii i ewolucji owadów jest bardzo zniekształcona w kierunku kilku rzędów obejmujących *D. melanogaster* i kilku innych bardziej współczesnych organizmów modelowych, takich jak trojszyk gryzący *Tribolium castaneum*, pszczoła miodna *Apis mellifera* czy jedwabnik morwowy *Bombyx mori*.

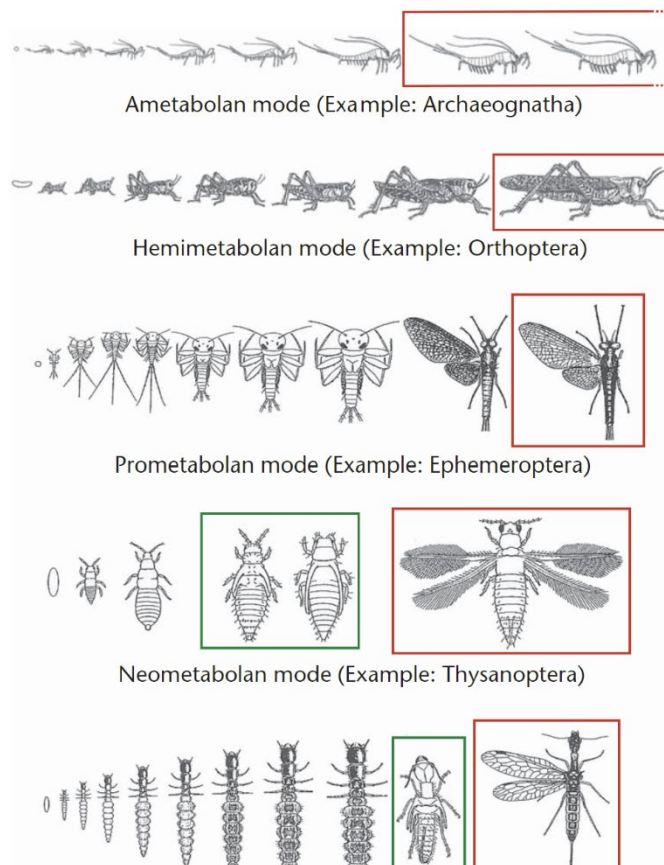
Podczas gdy skupienie się na kilku organizmach modelowych było kluczem do ogromnego postępu w genetyce w XX wieku, teraz w XXI wieku, musimy poszerzyć ich zakres, jeśli chcemy mieć jaśniejszy obraz różnorodności biologicznej świata i procesów ewolucyjnych, które go stworzyły. Owady to najliczniejsza i najbardziej zróżnicowana linia zwierząt na Ziemi, z ponad milionem opisanych gatunków (Grimaldi i Engel, 2005; Ylla, 2017). Owady były jednymi z pierwszych linii zwierząt, które opuściły środowisko wodne i zamieszkały na lądzie, i od tego czasu podbiły właściwie wszystkie ekosystemy lądowe (Engel i Grimaldi, 2004). Ich długa historia ewolucyjna, ogromna liczba gatunków oraz niezliczone formy i style życia sprawiają, że owady są jednym z najlepszych modeli do badania ewolucji i różnorodności biologicznej. Jednak, jak wspomniano wcześniej, nasze dane dotyczące owadów są dość stronicze, dlatego przez całą moją karierę naukową pracowałem nad

generowaniem genomów i danych genomowych dla zazwyczaj niedoreprezentowanych linii owadów, takich jak świerszcze (**publikacja nr 1**) i karaluchy (Harrison i in., 2018).

4.3.1.3 Metamorfoza owadów, czyli jak jeden genom wytwarza różne fenotypy

Berlese w 1913 r. (Berlese, 1913) zaproponował system klasyfikacji owadów oparty na sposobie, w jaki osiągają dorosłość, czyli innymi słowy na podstawie typu metamorfozy (przeobrażenia). Klasyfikacja ta jest nadal powszechnie stosowana (Belles, 2020). System klasyfikacji Berlese (**Ilustracja 1**) definiuje cztery główne grupy owadów: Ametabolous, Hemimetabolous, Holometabolous i Neometabolous.

Owady ametaboliczne to owady, które nie przechodzą metamorfozy (przeobrażenia). Lęgi owadów są morfologicznie praktycznie identyczne z ich formą dorosłą, ale mają mniejszy rozmiar. Podczas stadiów młodzieńczych mają serię linień, w których każda wylinka jest nieco większa niż poprzednia. Na podstawie informacji kopalnych uważa się, że jest to sposób rozwoju przodków współczesnych owadów. Niektóre obecnie żyjące linie owadów, takie jak rybiki cukrowe i skoczogonki, są przykładami rozwoju ametabolicznego



Ilustracja 1: Pięć rodzajów przeobrażenia według klasyfikacji Berlese 1913 z rysunkami przedstawiającymi rozwój jednego reprezentatywnego gatunku. Z Belles 2011, będącym modyfikacją za zgodą Sehnalet al. (1996).

Owady hemimetaboliczne to parafiletyczna grupa owadów. Charakteryzują się one młodymi stadiami morfologicznie podobnymi do dorosłych, ale z pewnymi cechami, takimi jak skrzydła i dojrzałe struktury rozrodcze, pojawiającymi się dopiero w stadium dorosłym. Przykładami tego trybu rozwoju są świerszcze i karaluchy.

Owady holometaboliczne to grupa monofiletyczna, która wyłoniła się z jednej z gałęzi owadów hemimetabolicznych, u których stadium młodociane wyraźnie różni się od postaci dorosłej. Ponadto gatunki te przechodzą etap pośredni, zwany poczwarką, między formami młodocianymi a dorosłymi. Linia ta obejmuje między innymi motyle, muchy i pszczoły

Ostatnia grupa, owady neometaboliczne, jest grupą parafiletyczną obejmującą gatunki co najmniej z rzędu Sternorrhyncha i Thysanoptera (Belles, 2011). Oznacza to, że metamorfoza neometaboliczna pojawiła się niezależnie w co najmniej dwóch różnych rzędach z grupy hemimetabolicznej. Owady z tej linii wytwarzają stadia młodociane, które znacznie różnią się od form dorosłych oraz mają nieruchome stadia młodociane, które przypominają poczwarki holometaboliczne. Według Berlese'a (Berlese, 1913) owady neometaboliczne są formami pośrednimi między hemimetabolicznymi a holometabolicznymi. Do tej grupy należą czerwce i wełnowcowate (Belles, 2020).

Prawdopodobnie jednym z najlepszych przykładów wykorzystania pojedynczego genomu do wytworzenia radykalnie różnych fenotypów jest przeobrażenie holometaboliczne. Kluczem do takiego procesu jest zatem sposób interpretacji genomu, a mianowicie regulacja ekspresji genów. Jednym z głównych sposobów badania regulacji genomu jest badanie różnych poziomów regulacji ekspresji genomu. Jest to, oprócz zainteresowania naukowego metamorfozą, powodem, dla którego w moich projektach badawczych często wykorzystuję przeobrażenie owadów jako model do badania ewolucji i regulacji sieci ekspresji genów

4.3.1.4 Ewolucja i rozwój

Oprócz zrozumienia, w jaki sposób proces jest regulowany (lub w niektórych przypadkach właśnie w celu zrozumienia, w jaki sposób proces jest regulowany) badamy również ewolucyjny aspekt tego procesu. Ten kierunek badań łączący ewolucję i rozwój jest często nazywany „evo-devo”. Dlatego w moich badaniach zawsze mam na uwadze ewolucyjną perspektywę biologicznego fenomenu, który chcemy badać. Na przykład w przypadku przeobrażenia owadów w **publikacji nr 2** opublikowanej niedawno w PNAS proponujemy nową teorię dotyczącą pochodzenia metamorfozy holometabolicznej przez utratę czynnika transkrypcyjnego w zarodku. Ta proponowana teoria, poparta obszernymi dowodami przedstawionymi w artykule, stanowi zmianę paradygmatu w badaniu przeobrażenia owadów.

Przeobrażenie jest często definiowane jako rozwój postembrionalny i, podobnie jak w rozwoju embrionalnym, sieci regulacyjne odgrywają kluczową rolę w wielu innych procesach rozwojowych i ich ewolucji. Jednym z takich procesów jest określanie zdolności reprodukcyjnej zwierząt. Różne zwierzęta wyewoluowały różne strategie reprodukcyjne, niektóre produkują niewielką liczbę potomstwa, podczas gdy inne wytwarzają niezliczone ilości młodych. U *Drosophila* szczytową zdolność reprodukcyjną obserwuje się podczas stadium larwalnego, w którym tworzą się jajniki. W **publikacji nr 5** opisujemy podstawy ekspresji genów rozwoju jajników *D. melanogaster*. Praca ta nie tylko dostarczyła nam istotnych informacji dotyczących samego procesu morfogenezy jajników podczas rozwoju, ale pozwoliła nam ustanowić nowy kierunek badań, który doprowadził nas do odkrycia nowych ról szlaku Hippo związanych ze zdolnościami rozrodczymi (artykuł w przygotowaniu) oraz zaangażowanie nowego czynnika w ten dobrze znany szlak proliferacji komórek.

4.3.1.5 Dane omiczne do badania regulacji ekspresji genów

Przyrostek „omika”, który wyewoluował z przyrostka „ome”, jest często używany do wskazania dziedziny badań grupy cząsteczek biologicznych na dużą skalę. Na przykład z pojęć genomu lub transkryptomu wyłoniły się terminy genomika i transkryptomika, odnoszące się odpowiednio do badania całych genomów i całych transkryptomów. Obecnie prawie wszystkie grupy cząsteczek biologicznych posiadają odrębne pole „omiczne”. Na przykład, oprócz wcześniej wspomnianej genomiki i transkryptomiki, mamy proteomikę, epigenomikę, lipidomikę, glikomikę i inne. Badanie interakcji jednej lub więcej grup cząsteczek jest często nazywane interaktomiką.

Proces regulacji ekspresji genów, a mianowicie sposób, w jaki jeden gen wytwarza dokładną ilość produktu białkowego w danej komórce, jest niezwykle złożonym procesem, w który zaangażowane są niezliczone cząsteczki biologiczne. Często jednak genom, transkryptom i epigenom są uważane za główne czynniki w procesie regulacyjnym. Stąd też przez całą swoją karierę publikowałem artykuły związane z różnymi dziedzinami omiki: genomiką (Harrison i in., 2018; Ylla i in., 2021), transkryptomiką (Fernandez-Nicolas i in., 2023; Llonga i in., 2018; Montañés i in., 2021; Tarikere i in., 2021; Ventós-Alfonso i in., 2019; Ylla, 2017; Ylla i in., 2016, 2017, 2018) oraz epigenomiką (Ventós-Alfonso i in., 2020b).

Co więcej, te pola „omiki” często można dalej podzielić w oparciu o konkretną badaną cząsteczkę. Na przykład w ramach transkryptomiki można badać różne typy RNA, często klasyfikowane według ich wielkości i/lub funkcji. W tym miejscu, w dziedzinie transkryptomiki, zwrócę uwagę na moją pracę nad różnymi poddziedzinami transkryptomiki, które dotyczą mRNA (**publikacje nr 2 i 5**) oraz różnych małych RNA, takich jak piRNA (**publikacja nr 3**) i siRNA (**publikacja nr 4**) (**Tabela 1**). Dziedzina genomiki jest również dogłębnie eksplorowana w **publikacji nr 1**. W publikacji tej zsekwencjonowałem i opisałem

pierwszy genom świerszcza. Opisałem w niej też drugi genom świerszcza i przeprowadziłem porównawczą analizę genomyczną z 14 innymi genomami owadów, aby uzyskać wgląd w regulację genomu i ewolucję owadów.

Praca z nowymi typami danych o gatunkach niemodelowych często oznacza nowe i nieoczekiwane problemy. Aby przezwyciężyć takie problemy, czasami musimy opracować narzędzia, z których mogą później korzystać inni badacze borykający się z podobnymi trudnościami. W tym przypadku moje narzędzia to zazwyczaj narzędzia bioinformatyczne, czyli innymi słowy oprogramowanie. Przykład takiego podejścia można znaleźć w (**publikacji nr 6**) wraz z opracowaniem MirCure. Kolejne takie narzędzie jest obecnie opracowywane w ramach przyznanego projektu grantowego SONATA 17.

4.3.2 Cele badawcze

Nadrzędnym celem moich badań jest zrozumienie ewolucji genowych sieci regulacyjnych. Mianowicie, w jaki sposób genomy determinują fenotypy i jak te sieci ewoluowały w czasie.

W obrębie tego szeroko zakrojonego długoterminowego celu, istnieje szereg celów szczegółowych, które omówię w niniejszym dokumencie.

- Wykorzystanie nowych genomów do zrozumienia ewolucji sieci regulacji genów (omówione w **publikacji nr 1**).
- Określenie biogenezy, funkcji i ról różnych małych RNA, takich jak piRNA i siRNA (omówione w **publikacjach nr 3 i 4**).
- Identyfikacja mechanizmów ekspresji genów, które podczas rozwoju determinują zdolność reprodukcyjną wykorzystując *Drosophila* jako organizm modelowy (omówione w **publikacji nr 5**).
- Opracowanie narzędzia bioinformatycznego niezbędnego do uzyskiwania wysokiej jakości adnotacji miRNA z niemodelowych gatunków zwierząt i roślin (omówione w **publikacji nr 6**).

4.3.3 Najważniejsze osiągnięcia naukowe

Mój główny wkład naukowy, jak wskazano w tytule „Podejścia omiczne rzucające światło na ewolucję i rozwój zwierząt” dotyczy połączenia różnych podejść omicznych w badaniach evo-devo. W tym celu i z powodów wyszczególnionych powyżej większość moich badań została wykonana przy użyciu danych „omicznych” dotyczących gatunków owadów.

W tej sekcji skupię się na moim głównym wkładzie w **1)** dziedzinę genomiki, **2)** różne elementy świata transkryptomiki (mRNA, piRNA, siRNA) oraz **3)** rozwój narzędzi bioinformatycznych do analizy danych omicznych.

4.3.3.1 Istotny wkład w **badania genomiczne** w celu zbadania regulacji i ewolucji genomu
Mój pierwszy wkład w dziedzinę genomiki miał miejsce podczas mojego doktoratu, kiedy brałem udział w składaniu genomu i adnotacji pierwszych genomów karaluchów, które opublikowaliśmy w *Nature Ecology and Devolution* (Harrison i in., 2018). W artykule dostarczyliśmy nowych wskazówek dotyczących ewolucji eusocjalności u owadów hemimetabolicznych, porównując nowo zsekwencjonowany genom żyjącego samotnie karaczana prusaka *Blattella germanica* ze społecznymi termitami, które są siostrzaną grupą karaluchów.

Mój istotny wkład w tę dziedzinę odpowiada opisowi **publikacji nr 1**. Artykuł opublikowany w czasopiśmie *Communications Biology* z Nature Publishing Group jest wynikiem międzynarodowej współpracy, którą kierowałem. W projekcie tym zsekwencjonowaliśmy i opisaliśmy 1,6-Gb genom świerszcza *Gryllus bimaculatus*.

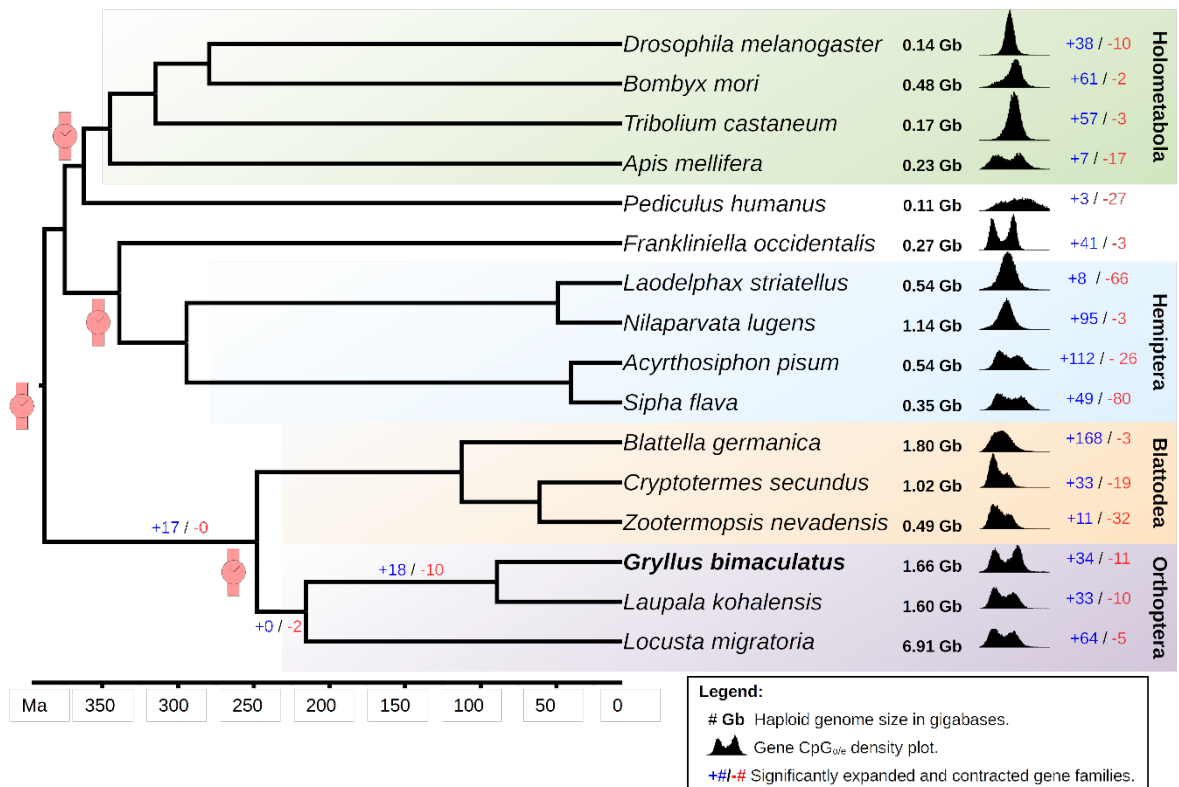
Świerszcz *Gryllus bimaculatus* ma długą historię jako model eksperymentalny w różnych dziedzinach, takich jak neurobiologia, rozwój i ewolucja (Mito i in., 2022). Zaproponowano również różne gatunki świerszczy jako źródło białka zwierzęcego (uzyskiwanego zgodnie z zasadami zrównoważonego rozwoju) do spożycia przez ludzi, biorąc pod uwagę, że są one już tradycyjnie spożywane w wielu częściach świata (van Huis i in., 2013). Pomimo zainteresowania nim w różnych dziedzinach badań i w przemyśle spożywczym, nie było dostępnych genomów z adnotacjami dla gatunków świerszczy.

Oprócz sekwencjonowania i adnotacji genomu *G. bimaculatus*, dokonaliśmy również adnotacji wstępnego genomu świerszcza *Laupala kohalensis*. Mając dwa genomy z adnotacjami porównujemy je z genomami 14 innych gatunków owadów (**Ilustracja 2**). Te genomowe porównania dostarczają bardzo interesujących informacji dotyczących ewolucji genomowej owadów, jak podsumowano poniżej.

Owady wykazują duże różnice pod względem wielkości genomu, z niektórymi liniami charakteryzującymi się wielkością genomu poniżej 1 Gb (Holometabola), podczas gdy inne zazwyczaj przekraczają 1 Gb (Gregory, 2002). W obrębie rodzaju Orthoptera, który obejmuje świerszcze, koniki polne i szarańczę, istnieje ogromna różnorodność rozmiarów genomów, od mniej niż 1 Gb do 15 Gb. Nie było jasne, jaka jest przyczyna takiego rozrostu genomu u niektórych gatunków ani w którym momencie ich ewolucji nastąpiła. Czy pierwotnie występował duży genom i uległ on skurczeniu w niektórych liniach? A może genom przodków był niewielki i niezależnie ulegał rozrostowi w niektórych liniach?

Nasze wyniki z porównania genomów świerszczy sugerują, że genom przodków był mniejszy i rozrastał się niezależnie w każdym z gatunków świerszczy ze względu na aktywność różnych elementów transpozycyjnych. Wyniki te uzyskano z porównania ich powtarzalnych sekwencji. W skrócie, dwa świerszcze mają bardzo podobny rozmiar genomu

(1,66 i 1,60 Gb) i podobną zawartość elementów transpozycyjnych (TE) (29% i 35%). Jednak skład powtarzalnych sekwencji różni się między nimi. U *L. kohalensis* długie, przeplatane elementy jądrowe (LINEs) stanowią blisko 60% całkowitej zawartości TE, podczas gdy u *G. bimaculatus* LINE stanowią jedynie 31% TE. Wyniki te, wraz z dogłębnymi porównaniami innych rodzin TE, sugerują, że genom przodków był mniejszy i rozszerzał się niezależnie dzięki aktywności różnych TE.



Ilustracja 2: Drzewo filogenetyczne z Ylla i in. 2021, z 2 genomami świerszcza i 14 dodatkowymi genomami owadów. Skala czasu drzewa została skalibrowana w 4 punktach czasowych (czerwone zegarki). Dla każdego gatunku rozmiar genomu jest pokazany na czarno. Liczba rozszerzonych i skurczonych rodzin genów w każdym gatunku i gałęziach jest pokazana odpowiednio na niebiesko i czerwono. Wykresy gęstości pokazują rozkład CpG_{0/e} dla każdego gatunku.

Byliśmy również w stanie wykorzystać sekwencje genomów, aby uzyskać wgląd w ewolucję metylacji DNA w tych genomach. Metylacja DNA zwykle zachodzi w cytozynie, po której następuje guanina (CpG). Metylacja C zwiększa częstość mutacji z C do T (Bird, 1980). Oznacza to, że regiony metylowane DNA w komórkach rozrodczych będą miały tendencję do utraty CpG. To zmniejszenie częstości CpG nazywane jest wyczerpaniem CpG i można je badać, porównując obserwowane CpG w danym regionie z oczekiwaną liczbą CpG na podstawie częstości C i G, co nazywa się nadmierną obserwacją CpG w stosunku do oczekiwanej (CpG_{0/e}).

Obliczyłem CpG_{0/e} dla wszystkich genów z 16 genomów i wykreśliłem wyniki jako gęstości na **ilustracji 2**. Na pierwszy rzut oka można zauważyć, że niektóre genomy

charakteryzuje pojedynczy pik, podczas gdy w innych obserwuje się dwa piki. Pojedynczy pik CpG_{o/e}, którego wartość wynosi około 1, wskazuje, że geny mają oczekiwaną ilość CpG, a tym samym sugeruje brak metylacji DNA. Natomiast dwa piki, jeden w okolicy 1 i drugi poniżej wartości 1, wskazują na dwie grupy genów, z których jedna jest zazwyczaj metylowana, a druga jest zazwyczaj niemetylowana.

Ze względu na brak metylacji DNA u większości typowych owadów modelowych (takich jak muchy, pszczoły i chrząszcze), występowanie metylacji DNA u owadów od wielu lat budzi kontrowersje. Obecnie wyraźnie widać, że brak metylacji DNA nie jest ogólną cechą owadów, ale tylko niektórych gatunków holometabolicznych i Hemiptera. Rola metylacji DNA u owadów pozostaje jednak nieznana. W publikacji tej udało nam się uzyskać kilka interesujących informacji dotyczących potencjalnych funkcji metylacji DNA u owadów.

4.3.3.2 Istotny wkład w wykorzystanie transkryptomiki do badania metamorfozy owadów

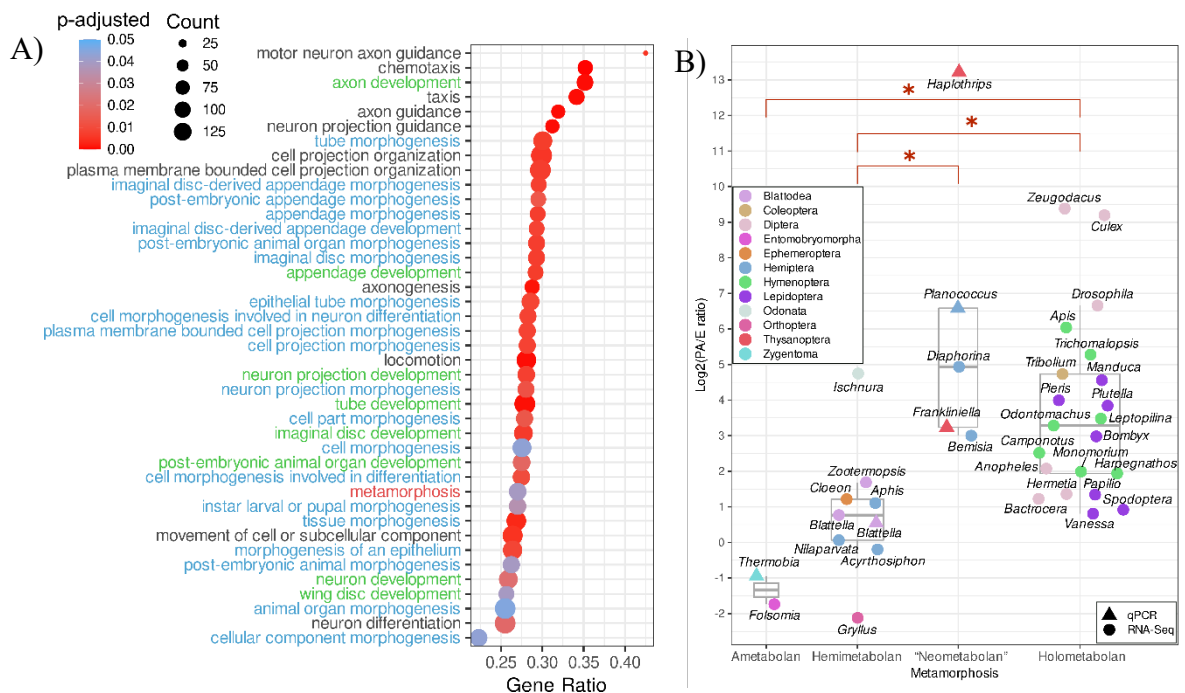
Podczas mojego doktoratu zacząłem badać transkryptomiczne podstawy regulacji, pochodzenia i ewolucji metamorfozy owadów. Jest to temat, na który opublikowałem kilka artykułów (Belles & Ylla, 2015; Ventos-Alfonso i in., 2019; Ventós-Alfonso i in., 2020a; Ylla i in., 2018). Wszystkie te artykuły dotyczą różnych aspektów regulacji metamorfozy na poziomie transkryptomiki. W jednej z tych publikacji (Ylla i in., 2018) zaobserwowałem, że gen o nazwie E93 ulegał silnej ekspresji w zarodku karalucha *B. germanica* i był nieobecny w zarodku *D. melanogaster*. Obserwacja ta zaowocowała czterema latami dalszych badań, których wyniki stały się jednym z moich najważniejszych wkładów w dziedzinę transkryptomicznej regulacji metamorfozy owadów i zostały opublikowane w 2023 roku w *Proceedings of the National Academy of Science (PNAS)* i przedstawione tutaj jako **publikacja nr 2**.

E93 jest genem, który inicjuje kształtowanie dorosłych owadów hemimetabolicznych i holometabolicznych w stadiach młodocianych (Urena i in., 2014). Moja wstępna obserwacja, oparta na danych z sekwencjonowania RNA, wskazywała, że E93 ulegał silnej ekspresji w zarodku owada hemimetabolicznego *B. germanica*, a nie tylko podczas oczekiwanego stadium przed dojrzałością. Natomiast holometaboliczny *D. melanogaster* cechował się jedynie oczekiwanym szczytem ekspresji w stadium młodocianym (poczwarka). Doprowadziło to do hipotezy, że E93 w zarodku owadów hemimetabolicznych może być wyznacznikiem wytworzenia stadium młodocianego, które wygląda jak dorosły, a jego brak może pozwolić na rozwój różnych młodocianych morfologii, takich jak larwy holometaboliczne.

Hipoteza ta wymagała jednak dalszej weryfikacji. Aby uzyskać jej potwierdzenie, stworzyliśmy nowy projekt badawczy, który łączył eksperymenty biologii molekularnej i analizę dużych danych obliczeniowych. Część eksperymentalna polegająca na uzyskaniu danych została przeprowadzona w laboratoriach naszych współpracowników w Barcelonie i

Kioto, natomiast w moim laboratorium na UJ zajęliśmy się analizą danych, uzyskaniem wyników i wyciągnięciem wniosków.

Wykorzystując technikę interferencji matczynego RNA (RNAi), pozbawiliśmy zarodki *B. germanica* transkryptów E93. Eksperyment ten spowodował poważne deformacje embrionów, niezdolnych do formowania kończyn i rozwoju. Następnie przeprowadziliśmy sekwencjonowanie RNA z zarodków pozbawionych E93, jak również z zarodków kontrolnych. Analiza transkryptomyczna ujawniła, że u owadów pozbawionych E93 dochodziło do znacznej redukcji procesów biologicznych zaangażowanych w rozwój i morfogenezę (**rysunek 3A**). Wszystkie te dane potwierdzają naszą hipotezę, że E93 może odgrywać rolę determinującą dorosłość w zarodku owada hemimetabolicznego *B. germanica*.



Ilustracja 3: A) Analiza transkryptomyczna próbek *B. germanica* zubożonych w E93 ujawniła znacząco obniżone warunki GO zaangażowane w rozwój (zielony) i morfogenezę (niebieski) B) Stosunek poziomów E93 w zarodku i w wieku przeddorosłym jest niski u gatunków hemimetabolicznych, co wskazuje na podobne poziomy transkryptu E93 między dwoma etapami. Dla kontrastu, neometabolany i holometabolany charakteryzuje wysoki poziom E93 u form przed-dorosłych i bardzo niski w zarodkach, co jest przeciwieństwem ametabolonów. Rysunek zmodyfikowany z Fernandez-Nicolas et. al. 2023

Na tym etapie trudność polegała jednak na znalezieniu sposobu, w jaki moglibyśmy dalej testować, czy obserwacja poczyniona u *B. germanica* i *Drosophila* dotycząca poziomów E93 w zarodku jest prawdziwa u innych owadów. Potwierdzałoby to naszą hipotezę dotyczącą roli E93 w ewolucji metamorfozy owadów. Postanowiliśmy zastosować podejście *big data*. Zidentyfikowaliśmy wszystkie genomy owadów z dostępnymi adnotacjami genów w NCBI. Następnie, dla wszystkich tych owadów, przyjrzelśmy się metadaniom NCBI-SRA dla

wszystkich dostępnych transkryptomów zarodków (dowolnego stadium) i stadiów przed osiągnięciem dojrzałości.

Zapytania te pozwoliły nam uzyskać 848 transkryptomów z 33 gatunków. Korzystając z klastra obliczeniowego mojego laboratorium, zmapowaliśmy wszystkie te zestawy danych sekwencjonowania RNA z odpowiadającymi im genomami i określiliśmy ilościowo poziom E93 w każdym z nich. Następnie, aby określić, co można uznać za wysoki lub niski poziom ekspresji E93, zdecydowaliśmy się użyć poziomu E93 w stadiach młodocianych jako odniesienia. Podzieliliśmy ekspresję E93 w najwyższym punkcie dla zarodka przez ekspresję dla najwyższego punktu w okresie przeddojrzałym i przekształciliśmy go do skali log2 (**ilustracja 3B**). Wyniki były niezwykle interesujące.

Hemimetabolany miały tendencję do posiadania podobnych jak obserwowane u *B. germanica* poziomów ekspresji E93 w zarodku i wieku przeddorosłym. Owady holometaboliczne, jak obserwowaliśmy u *Drosophila*, miały wysoki poziom E93 w stadium przeddorosłym, lecz bardzo niski lub nieobecny w zarodkach. Neometabolany, które, jak wyjaśniono powyżej, niezależnie wyewoluowały bardzo rozbieżne stadia młodociane, były podobne do owadów holometabolicznych, co również pasowało do naszej hipotezy, że brak E93 pozwoliłby na tworzenie młodych osobników o bardzo odmiennej morfologii. Ostatecznie, 2 ametaboliczne owady, które wylęgają się jako niemal identyczne wersje osobników dorosłych, miały wysoki poziom E93 jedynie w zarodku.

Aby dodatkowo zweryfikować profil ekspresji E93 owadów ametabolicznych, nasi współpracownicy w Kioto wykonali oznaczenie E93 metodą RT-qPCR podczas całego życia ametabolicznego *Thermobia domestica*. Ich profil potwierdził, że E93 charakteryzował pojedynczy pik ekspresji w zarodku.

Podsumowując, dzięki wygenerowanym przez nas transkryptomom oraz setkom publicznie dostępnych transkryptomów dostarczamy bardzo rzetelnych danych potwierdzających naszą hipotezę, że E93 w zarodku determinuje formy morfologicznie zbliżone do form dojrzałych u młodocianych owadów ametabolicznych i hemimetabolicznych. Utrata embrionalnego E93 umożliwiłaby pojawienie się pochodnych osobników młodocianych, takich jak gatunki neometaboliczne i holometaboliczne.

4.3.3.3 Istotny wkład w rolę mniej zbadanych małych RNA: siRNA i piRNA

W ostatnich dziesięcioleciach odkryto, że małe RNA są ważnymi regulatorami ekspresji genów. Większość uwagi skupiała się w szczególności na miRNA. Rzeczywiście, podczas mojego doktoratu dogłębnie przestudiowałem rolę miRNA w ewolucji i rozwoju zwierząt, co zaowocowało wysoce cytowanymi publikacjami (Ylla i in., 2016, 2017). Oprócz miRNA istnieją inne typy małych RNA, których funkcje są jeszcze słabo zbadane. W tej sekcji opiszę

moje dwie istotne zasługi dla lepszego zrozumienia małych RNA. Pierwsza dotyczy roli piRNA, a druga biogenezy siRNA.

Mój **najistotniejszy wkład w dziedzinę piRNA** został opublikowany w *Journal of Experimental Biology* (JEZ) (**publikacja nr 3**). piRNA to jednoniciowe RNA złożone z ok. 28 niekodujących nukleotydów, które zostały opisane głównie jako strażnicy genomu w komórkach rozrodczych (Brennecke i in., 2007; Sytnikova i in., 2014). Ponieważ piRNA zostały dokładnie zbadane wyłącznie w kilku organizmach modelowych (np. *Drosophila*, ludzie i myszy), uważano, że ulegają one ekspresji wyłącznie w komórkach rozrodczych. Jednak ekspresja wyłącznie w komórkach rozrodczych może nie być uniwersalną cechą wszystkich zwierząt. Faktycznie, ostatnio wykazano, że mięczaki wykazują ekspresję somatycznego piRNA (Jehn i in., 2018).

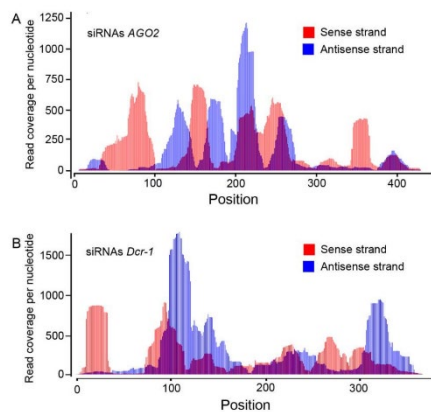
W **publikacji nr 3** ponownie wykorzystaliśmy zestaw 22 małych bibliotek sekwencji RNA *B. germanica*, które wygenerowałem do jednej z publikacji mojego doktoratu (Ylla i in., 2017), aby określić profile ekspresji piRNA w trakcie rozwoju. Wyzwaniem tej publikacji było rzetelne zidentyfikowanie i porównanie piRNA w każdym z 11 etapów rozwojowych, dla których wygenerowaliśmy dane. Ostatecznie wykorzystaliśmy znane cechy piRNA, takie jak pochodzenie z klastrów lub cyklu „ping-pong”, ich typową długość, odchylenie w kierunku uracylu na pierwszej pozycji itp., co pozwoliło nam rzetelnie zidentyfikować piRNA w każdej bibliotece i rozróżnić je pod względem ścieżki biogenezy, która je zapoczątkowała.

Wyniki naszej analizy sugerowały wydatną ekspresję piRNA w komórkach somatycznych przez cały czas rozwoju, z najbardziej wyraźną ekspresją podczas rozwoju zarodka. Ponadto istniały grupy piRNA o ekspresji swoistej dla danego stadium, co sugeruje aktywną i regulowaną transkrypcję różnych piRNA. Co więcej, wszystkie wyniki tego projektu wraz z dalszą literaturą opublikowaną w tym samym czasie wskazywały, że podstawowe linie zwierzęce mogą wykazywać ekspresję somatycznych piRNA i że mogą one odgrywać rolę regulacyjną ekspresji genów, oprócz pełnienia funkcji strażnikami genomu w zarodkach komórki. Hipoteza ta jest kamieniem węgielnym mojego obecnie realizowanego projektu badawczego (SONATA 17).

Kolejny **istotny wkład, jaki wniosłem w dziedzinę siRNA**, został opublikowany w **publikacji nr 4**. siRNA to małe (ok. 21 nt) jednoniciowe RNA wytwarzane z egzogennych lub endogennych dsRNA, które mają zdolność do zużywania komplementarnych jednoniciowych RNA. Technika interferencji RNA (RNAi) wykorzystuje ten mechanizm do wyciszania genów *in vivo* na poziomie posttranskrypcyjnym. Odbywa się to poprzez wstrzyknięcie lub ekspresję dsRNA, który zawiera sekwencję genu będącego obiektem zainteresowania. dsRNA jest przetwarzane przez komórki celem wytworzenia siRNA, które byłyby ukierunkowane na mRNA genu będącego obiektem zainteresowania, wyciszając go

po transkrypcji. Technika ta jest potężnym narzędziem, w szczególności stosowanym w organizmach niemodelowych, dla których techniki edycji genomu nie zostały jeszcze w pełni rozwinięte.

Pomimo tego, że jest to szeroko stosowana technika, istnieje jednak wiele niewiadomych dotyczących molekularnych podstaw wytwarzania i działania tych siRNA. W konsekwencji naukowcy mogą przygotowywać dsRNA, które mają niską skuteczność w wyciszeniu docelowego genu, a zmieniając nieco region wstawiany do dsRNA, wydajność może wzrosnąć lub zmaleć. Oznacza to, że często ma miejsce proces „prób i błędów” celem znalezienia najskuteczniejszego fragmentu genu do wyciszenia danego genu. Lepsze zrozumienie tego mechanizmu mogłoby pozwolić na zaprojektowanie najskuteczniejszego dsRNA *in silico*, zamiast kosztownej metody prób i błędów.



Ilustracja 4: pokrycie odczytu wzdłuż małej sekwencji RNA wstrzykniętego dsRNA. Rysunek z Montanes i in. 2011.

Jednym z najciekawszych wyników analizy był fakt, że pod nieobecność Argonaute-2 nadal powstają siRNA, które gromadzą się w komórce, biorąc pod uwagę, że Argonaute-2 jest nieobecny, więc nie można ich rekrutować i wykorzystać jako szablonu do znajdowania docelowego mRNA, które ulegają degradacji. Analizując odczyty małego RNA zmapowane na sekwencję dsRNA, które odpowiadają odczytom z cząsteczek siRNA, zaobserwowaliśmy, że każdy dsRNA wytwarza dowolny z możliwych siRNA obejmujący jego pełną sekwencję (**Ilustracja 4**). Nie wszystkie produkowane siRNA są jednak wykrywane z tą samą częstotliwością, a raczej niektóre siRNA występują w znacznie większej proporcji niż inne. Wzór jest powtarzalny między powtórzeniami i różnicami między dwiema niemi oraz między różnymi sekwencjami dsRNA (**Ilustracja 4**). Zwiększona zawartość jednego siRNA względem innego również nie zależała od tego, czy siRNA ma cel (target) w komórce, czy nie, ponieważ analizując odczyty z siRNA z dsRNA przeciwko genowi nieobecnemu w danym gatunku, zaobserwowaliśmy również niektóre regiony dsRNA wytwarzające większą ilość siRNA niż inne.

Wyniki te pokazują, że nasze zrozumienie sposobu wytwarzania siRNA jest nadal niepełne. Według naszej wiedzy nie ma jasnego wyjaśnienia, dlaczego niektóre siRNA gromadzą się z wyższą częstotliwością niż inne, ani jakie cechy to determinują. Możliwość zaprojektowania dsRNA w regionach, które wytwarzałyby największą ilość siRNA, pomogłaby zredukować koszty i czas poszukiwań najskuteczniejszego z nich metodą prób i błędów. Pomimo tych pierwszych kroków nadal musimy pracować w tej dziedzinie, aby wyjaśnić, jak działa szlak RNAi na poziomie molekularnym.

4.3.3.4 Istotny wkłady w transkryptomikę rozwoju i reprodukcji

Zdolność reprodukcyjna, która ma kluczowe znaczenie dla przetrwania gatunku, jest często określana w trakcie rozwoju. U *Drosophila* zdolność reprodukcyjna jest określana w fazie larwalnej, w której zaczynają tworzyć się jajniki. Morfogeneza jajnika jest złożonym procesem, kierowanym przez interakcję komórek rozrodczych i komórek somatycznych. Jajniki składają się z owariol, które są pojedynczymi jednostkami produkującymi jajeczka złożonymi z komórek somatycznych, z komórką rozrodczą na dnie. Zmiany w równowadze komórek rozrodczych lub somatycznych mogą wpływać na liczbę owariol, a tym samym na zdolność reprodukcyjną rozumianą jako liczba jaj do złożenia (Sarıkaya i Extavour, 2015).

Moim **najistotniejszym wkładem** w tę dziedzinę było odkrycie procesu regulacji ekspresji genów związanego z powstawaniem jajników *Drosophila* zarówno w komórkach rozrodczych, jak i komórkach somatycznych (**publikacja nr 5**). W tym celu wykonaliśmy sekwencjonowanie RNA całych jajników w 3 różnych punktach czasowych podczas ich rozwoju. Ponadto, aby zrozumieć różnice regulacyjne między komórkami somatycznymi i zarodkowymi, wykorzystaliśmy FACS do oddzielenia komórek rozrodczych od komórek somatycznych w tych samych 3 punktach czasowych. Następnie wykonaliśmy sekwencjonowanie RNA komórek rozrodczych i komórek somatycznych w trzech różnych punktach czasowych. Wszystkie badania wykonano w 3 powtórzeniach biologicznych, a tym samym wygenerowano kompleksowy zestaw danych 27 transkryptomów rozwijających się jajników *D. melanogaster*.

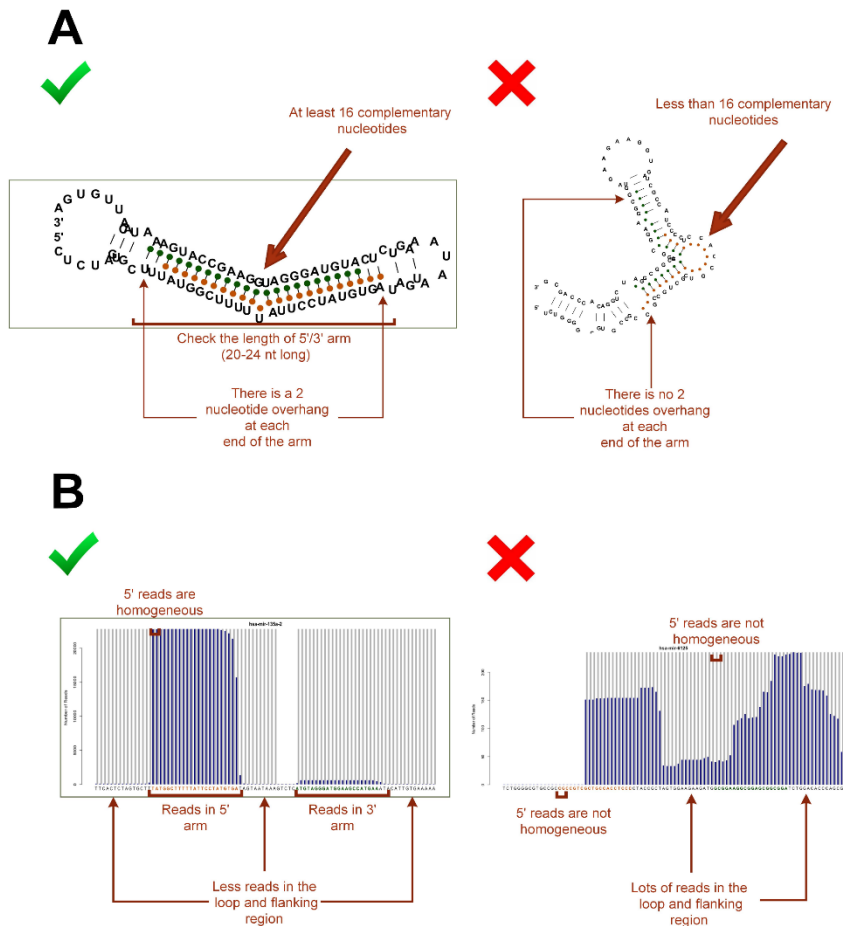
Moja analiza tego złożonego zestawu danych dostarczyła bardzo interesujących spostrzeżeń dotyczących sposobu, w jaki odbywa się morfogeneza jajników, jak opisano w publikacji. Do najciekawszych odkryć należy obserwacja, że komórki somatyczne wykazują największe zmiany transkryptomiczne pod koniec morfogenezy jajnika, podczas gdy komórki rozrodcze wykazują taki wzorzec we wczesnych stadiach. Inną istotną obserwacją było to, że w komórkach somatycznych istnieją różne szlaki zaangażowane w aktywną proliferację komórek, takie jak szlak MAPK i Hippo. Po tym projekcie przystąpiliśmy do badania roli szlaku Hippo w komórkach somatycznych i zarodkowych jajnika jako wyznacznika zdolności rozrodczych. Manuskrypt przedstawiający wyniki tego projektu uzupełniającego jest obecnie w przygotowaniu.

4.3.3.5 Istoty wkład w rozwój narzędzi bioinformatycznych do badań omicznych

W jednym z artykułów podczas mojej pracy doktorskiej (Ylla i in., 2016) chciałem porównać zestaw narzędzi miRNA między 7 gatunkami owadów, aby sprawdzić, czy pojawieniu się metamorfozy holometabolicznej towarzyszył wzrost liczby miRNA. W tym celu musiałem uzyskać wysoce wyselekcjonowane adnotacje miRNA dla każdego interesującego gatunku owada. Ku mojemu zdziwieniu narzędzia do przewidywania miRNA generowały znaczne ilości wyników fałszywie pozytywnych (do 99% niepoprawnych adnotacji) w dużych i nieopisanych genomach owadów. Nie było to również łatwe dla gatunków, dla których autorzy opisali już miRNA. Bazy danych miRNA również były pełne błędnych adnotacji. Kompleksowe badanie miRbase, głównego repozytorium adnotacji miRNA w tamtym czasie, zawierało prawie dwie trzecie adnotacji miRNA, którym brakowało odpowiedniego wsparcia (Fromm i in., 2015).

Oznacza to, że w przypadku tej pracy musiałem zainwestować ogromny wysiłek w kurację adnotacji miRNA zarówno pochodzącego z narzędzi do przewidywania *de novo*, jak i baz danych. W tym celu wygenerowałem wiele skryptów, które sprawdzały różne cechy miRNA, pomagając w procesie kuracji. Tak więc projekt, który chciałem realizować w ramach stażu podoktorskiego (postdoc), polegał na połączeniu wszystkich wygenerowanych skryptów w łatwej do uruchomienia aplikacji bioinformatycznej, która mogłaby korygować prognozy miRNA na podstawie danych o małych sekwencjach RNA i wiedzy, którą posiadamy na temat biogenezy miRNA. Laboratorium wybranym do tego projektu było Laboratory of Genomics of Gene Expression, kierowane przez dr Ana Conesa, znajdujące się wówczas na University of Florida (Gainesville, Floryda USA), które miało duże doświadczenie w opracowywaniu narzędzi bioinformatycznych.

Tak więc moim **najistotniejszym wkładem w rozwój narzędzi bioinformatycznych** było narzędzie o nazwie MirCure (**publikacja nr 6**), które zostało opublikowane w Bioinformatics. W skrócie, to bioinformatyczne narzędzie zaimplementowane w R z graficznym interfejsem użytkownika w Shiny pozwala użytkownikom na selekcjonowanie adnotacji miRNA zwierząt i roślin. Jako dane wejściowe wymaga małych sekwencji RNA zmapowanych do genomu referencyjnego oraz pliku z adnotacjami miRNA do wyselekcjonowania. Następnie MirCure dla każdej adnotacji miRNA sprawdzi różne cechy w oparciu o naszą wiedzę na temat biogenezy miRNA (**ryc. 5**). Na przykład, prawdziwy miRNA musi mieć prekursorowy RNA o długości ok. 100 nt, który składa się w strukturę przypominającą spinkę do włosów, która wytwarza miRNA 5' i 3'. Te 2 ramiona muszą być częściowo komplementarne i wspierane przez małe sekwencje RNA. Na 2 ramionach powinien znajdować się nakładający fragment 2 nukleotydów, co jest sygnaturą rozszczepienia Dicer i Drosha. MirCure zwraca użytkownikowi wyselekcjonowane adnotacje miRNA wraz z wynikiem wskazującym na ich jakość. Ponadto zapewnia również dane wyjściowe w formie wizualnej (**Ilustracja 5**) dla każdego kandydata na miRNA.



Ilustracja 5: Graficzne przedstawienie kryteriów oceny MirCure dla miRNA o wysokiej i niskiej punktacji. (A) Kryteria strukturalne, pokazujące właściwą adnotację miRNA po lewej stronie i niewspierany miRNA po prawej stronie. (B) Dowód oparty na sekwencji małego RNA (po lewej stronie właściwe miRNA, po prawej stronie niewspierana adnotacja miRNA). Rysunek z Ylla i in. 2020.

4.3.4 Dalsze kierunki badań

Mój plan badawczy na nadchodzące lata polega na kontynuowaniu badań różnych elementów sieci regulacyjnej genów, tego jak wchodzi z sobą w interakcje i jak ukształtowały ewolucję.

Znaczna część wysiłków w ciągu najbliższych trzech lat będzie skierowana na realizację celów przyznanego grantu SONATA 17 projektem pt. „Role funkcjonalne somatycznych piRNA i ich ewolucja”. Projekt ten łączy w sobie wiele z omawianych tu elementów. Z jednej strony ma na celu zbadanie roli i ewolucji piRNA przy użyciu owadów jako organizmów doświadczalnych oraz w oparciu o dane wygenerowane przez sekwencjonowanie RNA. Ponadto część tego projektu koncentruje się również na opracowaniu aplikacji bioinformatycznej, która pozwoliłaby nam identyfikować i opisywać piRNA.

W tym roku otrzymaliśmy również grant Polonez Bis 2 dla naukowca ze stopniem doktora, który dołączy do nas w celu zbadania odporności na stres abiotyczny manneczki łąkowej z wykorzystaniem podejścia genomowego i transkryptomowego. W projekcie tym, w którym pełnię rolę mentora, uzyskamy znaczną ilość danych genomowych i transkryptomowych z różnych akcesji manneczki łąkowej, aby znaleźć molekularne podstawy tolerancji na stres wywołany różnymi czynnikami.

Poza tymi dwoma projektami NCN zamierzam nadal badać możliwości finansowania innych interesujących projektów w dziedzinie danych omicznych, aby badać ewolucję i rozwój. Jeden z takich projektów, nad którymi obecnie pracuję, skupi się na badaniu roli metylacji DNA w embriogenezie owadów. Kolejny projekt, który planuję, będzie dotyczył genomowej adaptacji owadów do środowisk antropicznych.

Jak wspomniano w niniejszym dokumencie, wiele wcześniejszych projektów badawczych doprowadziło do powstania nowych projektów, z których część została opublikowana, a inne nadal są w toku. Podobnie niektóre z obecnie realizowanych projektów mogą poprowadzić mnie w dalszych i nieoczekiwanych kierunkach, czego nie mogę się doczekać. Dlatego mając umysł otwarty na nowe kierunki i metodologie, nie mogę się doczekać przyszłości w badaniach nad bioinformatyką, aby zrozumieć ewolucję i rozwój.

Będąc w uprzywilejowanym położeniu, posiadając nowoczesne techniki generowania danych i techniki bioinformatyczne, będę nadal opierał się na dziedzictwie A. R. Wallace'a i C. Darwina i dalej badał, jak *„od tak prostego początku rozwijały się i nadal rozwijają nieskończone formy, najpiękniejsze i najwspanialsze”* (Darwin, 1859).

4.3.5 Bibliografia

- Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., ... Venter, J. C. (2000). The Genome Sequence of *Drosophila melanogaster*. *Science*, 287(5461), 2185–2195.
<http://science.sciencemag.org/content/287/5461/2185>
- Belles, X. (2011). Origin and Evolution of Insect Metamorphosis. *Encyclopedia of Life Sciences*, 1–11. <https://doi.org/10.1002/9780470015902.a0022854>
- Belles, X. (2020). Insect Metamorphosis. In *Insect Metamorphosis*.
<https://doi.org/10.1017/CBO9781107415324.004>
- Belles, X., & Ylla, G. (2015). Towards understanding the molecular basis of cockroach tergal gland morphogenesis. A transcriptomic approach. *Insect Biochemistry and Molecular Biology*, 63, 104–112. <https://doi.org/10.1016/j.ibmb.2015.06.008>
- Berlese, A. (1913). Intorno alle metamorfosi degli insetti. *Redia*, 9(2), 121–137.
- Berry, A., & Browne, J. (2022). Mendel and Darwin. *Proceedings of the National Academy of Sciences*, 119(30), e2122144119. <https://doi.org/10.1073/PNAS.2122144119>
- Bird, A. P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research*, 8(7), 1499–1504. <https://doi.org/10.1093/nar/8.7.1499>
- Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., & Hannon, G. J. (2007). Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in *Drosophila*. *Cell*, 128(6), 1089–1103.
<https://www.sciencedirect.com/science/article/pii/S0092867407002577?via%3Dihub>
- Darwin, C. (1859). *On the origin of species*.
- Darwin, C., & Wallace, A. (1858). On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection. *Journal of the Proceedings of the Linnean Society of London. Zoology*, 3(9), 45–62.
<https://doi.org/10.1111/J.1096-3642.1858.TB02500.X>
- Darwin Correspondence Project. (2023, January 30). *Letter no. 5140*.

- Engel, M. S., & Grimaldi, D. A. (2004). New light shed on the oldest insect. *Nature*, 427(6975), 627–630. <https://doi.org/10.1038/nature02291>
- Fernandez-Nicolas, A., Machaj, G., Ventos-Alfonso, A., Pagone, V., Minemura, T., Ohde, T., Daimon, T., Ylla, G., & Belles, X. (2023). Reduction of embryonic E93 expression as a hypothetical driver of the evolution of insect metamorphosis. *Proceedings of the National Academy of Sciences*, 120(7), e2216640120. <https://doi.org/10.1073/PNAS.2216640120>
- Fromm, B., Billipp, T., Peck, L. E., Johansen, M., Tarver, J. E., King, B. L., Newcomb, J. M., Sempere, L. F., Flatmark, K., Hovig, E., & Peterson, K. J. (2015). A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annual Review of Genetics*, 49(1), 213–242. <https://doi.org/10.1146/annurev-genet-120213-092023>
- Gregory, T. R. (2002). Genome size and developmental complexity. *Genetica*, 115(1), 131–146. <https://doi.org/10.1023/A:1016032400147>
- Grimaldi, D., & Engel, M. S. (2005). *Evolution of the Insects*. Cambridge University Press. <http://books.google.es/books?id=Ql6Jl6wKb88C>
- Harrison, M. C., Jongepier, E., Robertson, H. M., Arning, N., Bitard-Feildel, T., Chao, H., Childers, C. P., Dinh, H., Doddapaneni, H., Dugan, S., Gowin, J., Greiner, C., Han, Y., Hu, H., Hughes, D. S. T., Huylmans, A. K., Kemena, C., Kremer, L. P. M., Lee, S. L., ... Bornberg-Bauer, E. (2018). Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nature Ecology and Evolution*, 2(3), 557–566. <https://doi.org/10.1038/s41559-017-0459-1>
- Jehn, J., Gebert, D., Pipilescu, F., Stern, S., Kiefer, J. S. T., Hewel, C., & Rosenkranz, D. (2018). PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineage-specific adaptation. *Communications Biology*, 1(1). <https://doi.org/10.1038/s42003-018-0141-4>
- Llonga, N., Ylla, G., Bau, J., Belles, X., & Piulachs, M. D. (2018). Diversity of piRNA expression patterns during the ontogeny of the German cockroach. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 330(5), 288–295. <https://doi.org/10.1002/jez.b.22815>
- Marx, V. (2013). The big challenges of big data. *Nature* 2013 498:7453, 498(7453), 255–260. <https://doi.org/10.1038/498255a>

- Mito, T., Ishimaru, Y., Watanabe, T., Nakamura, T., Ylla, G., Noji, S., & Extavour, C. G. (2022). Cricket: The third domesticated insect. In *Current Topics in Developmental Biology* (Vol. 147, pp. 291–306). Academic Press.
<https://doi.org/10.1016/bs.ctdb.2022.02.003>
- Montañés, J. C., Rojano, C., Ylla, G., Piulachs, M. D., & Maestro, J. L. (2021). siRNA enrichment in Argonaute 2-depleted *Blattella germanica*. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1864(6–7), 194704.
<https://doi.org/10.1016/j.bbagr.2021.194704>
- Nakamura, T., Ylla, G., & Extavour, C. G. (2022). Genomics and genome editing techniques of crickets, an emerging model insect for biology and food science. *Current Opinion in Insect Science*, 50, 100881. <https://doi.org/10.1016/j.cois.2022.100881>
- Sarikaya, D. P., & Extavour, C. G. (2015). The Hippo Pathway Regulates Homeostatic Growth of Stem Cell Niche Precursors in the *Drosophila* Ovary. *PLoS Genetics*, 11(2), 1–28. <https://doi.org/10.1371/journal.pgen.1004962>
- Sytnikova, Y. A., Rahman, R., Chirn, G. W., Clark, J. P., & Lau, N. C. (2014). Transposable element dynamics and PIWI regulation impacts lncRNA and gene expression diversity in *Drosophila* ovarian cell cultures. *Genome Research*.
<https://doi.org/10.1101/gr.178129.114>
- Tarikere, S., Ylla, G., & Extavour, C. G. (2021). Distinct gene expression dynamics in germ line and somatic tissue during ovariole morphogenesis in *Drosophila melanogaster*. *G3 Genes|Genomes|Genetics*. <https://doi.org/10.1093/G3JOURNAL/JKAB305>
- Urena, E., Manjon, C., Franch-Marro, X., & Martin, D. (2014). Transcription factor E93 specifies adult metamorphosis in hemimetabolous and holometabolous insects. *Proceedings of the National Academy of Sciences*, 111(19), 7024–7029.
<https://doi.org/10.1073/pnas.1401478111>
- van Huis, A., van Itterbeeck, J., Klunder, H., Mertens, E., Halloran, A., Muir, G., & Vantomme, P. (2013). *Edible insects: future prospects for food and feed security* (Issue 171). Food and agriculture organization of the United nations (FAO).
- Ventos-Alfonso, A., Ylla, G., & Belles, X. (2019). Zelda and the maternal-to-zygotic transition in cockroaches. *FEBS Journal*, 286(16), 3206–3221.
<https://doi.org/10.1111/febs.14856>

- Ventós-Alfonso, A., Ylla, G., Montañes, J.-C., & Belles, X. (2020a). DNMT1 Promotes Genome Methylation and Early Embryo Development in Cockroaches. *IScience*, 23(12), 101778. <https://doi.org/10.1016/j.isci.2020.101778>
- Ventós-Alfonso, A., Ylla, G., Montañes, J.-C., & Belles, X. (2020b). DNMT1 promotes genome methylation and early embryo development in cockroaches. *IScience*, 101778. <https://doi.org/10.1016/j.isci.2020.101778>
- Ylla, G. (2017). Comparative transcriptomics of hemimetabolan and holometabolan metamorphosis [Pompeu Fabra University]. In *TDX (Tesis Doctorals en Xarxa)*. <https://repositori.upf.edu/handle/10230/34842>
- Ylla, G., Fromm, B., Piulachs, M. D., & Belles, X. (2016). The microRNA toolkit of insects. *Scientific Reports*, 6, 37736. <https://doi.org/10.1038/srep37736>
- Ylla, G., Nakamura, T., Itoh, T., Kajitani, R., Toyoda, A., Tomonari, S., Bando, T., Ishimaru, Y., Watanabe, T., Fuketa, M., Matsuoka, Y., Barnett, A. A., Noji, S., Mito, T., & Extavour, C. G. (2021). Insights into the genomic evolution of insects from cricket genomes. *Communications Biology*, 4(1), 733. <https://doi.org/10.1038/s42003-021-02197-9>
- Ylla, G., Piulachs, M. D., & Belles, X. (2017). Comparative analysis of miRNA expression during the development of insects of different metamorphosis modes and germ-band types. *BMC Genomics*, 18(1), 774. <https://doi.org/10.1186/s12864-017-4177-5>
- Ylla, G., Piulachs, M.-D., & Belles, X. (2018). Comparative transcriptomics in two extreme neopterans reveal general trends in the evolution of modern insects. *IScience*, 4, 164–179. <https://doi.org/10.1016/j.isci.2018.05.017>

5. Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.

Moja kariera akademicka zaprowadziła mnie na **osiem uniwersytetów** w ośmiu różnych miastach na trzech różnych kontynentach. Ukończyłem studia licencjackie z biotechnologii na **University of Vic (UVic)** w mieście Vic w prowincji Barcelona (Katalonia). Tam po raz pierwszy zetknąłem się z badaniami naukowymi jako student-wolontariusz w laboratorium Consol Blanch, badając związki lotne trufli czarnozarodnikowej (*Tuber melanosporum*). Podczas studiów licencjackich miałem również okazję studiować za granicą. Mój pierwszy pobyt za granicą miał miejsce latem 2011 roku, kiedy przez trzy miesiące studiowałem na **Beijing Language and Culture University** w Pekinie (Chiny). Drugie studia za granicą również odbyły się w Chinach, tym razem przez pół roku na **Dalian University of Foreign Languages** w prowincji Liaoning (Chiny).

Po studiach w Chinach przenieśliem się do Barcelony, gdzie rozpocząłem pracę jako technik bioinformatyk w **Institute of Evolutionary Biology (IBE)**, będącym wspólnym ośrodkiem **Pompeu Fabra University** i **Spanish Research Council**. Pracując jako technik w IBE, zdobyłem tytuł magistra analizy danych omicznych (*Master in Omics Data Analysis*) na kampusie **UVic** w Barcelonie.

Po ukończeniu studiów magisterskich kontynuowałem naukę w **IBE**, tym razem jako doktorant w pracowniach Xaviera Bellesa i Marii Dolors Piulachs. Przez cały czas spędzony w IBE jako magistrant i doktorant, byłem autorem 5 publikacji. Również w tym czasie jako doktorant, uzyskałem stypendium na realizację letniego projektu badawczego na Tajwanie. Z tego powodu latem 2015 roku przez trzy miesiące dołączyłem do Computational Systems Biology Laboratory na **National Yang-Ming University** w Taipei (Tajwan). Pracowałem tam pod opieką Hsuan-Cheng Huanga, badając docelowe sieci mikroRNA.

Po uzyskaniu doktoratu przenieśliem się na **University of Florida** w Gainesville (Floryda, USA) jako stażysta podoktorski (postdoc) w laboratorium Genomics of Gene Expression prowadzonym przez Ane Conesa. W ciągu roku spędzonego w tym laboratorium opracowałem narzędzie do przeprowadzania kontroli jakości i kuracji miRNA opisane w części dotyczącej osiągnięć (**publikacja nr 6**).

Po ukończeniu projektu na Florydzie przenieśliem się na drugi staż podoktorski (postdoc) na **Harvard University** w Bostonie (Massachusetts, USA). Pracowałem tam przez trzy lata w laboratorium Cassandry Extavour, skupiając się na ewolucji i rozwoju stawonogów. W tym czasie byłem autorem dwóch artykułów naukowych (Tarikere i in., 2021; Ylla i in., 2021), recenzji (Nakamura i in., 2022) oraz rozdziału w książce (Mito i in., 2022).

Po trzech latach stażu podoktorskiego (postdoc) na Harvardzie, w listopadzie 2021 objąłem obecne stanowisko kierownika laboratorium na **Uniwersytecie Jagiellońskim** w Krakowie. W ciągu tych pierwszych lat od rozpoczęcia pracy na UJ byłem autorem 4 opublikowanych artykułów, 1 preprintu i uzyskałem grant SONATA na łączną wartość 1,3 mln zł.

6. Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę.

6.1 Osiągnięcia organizacyjne

W 2015 roku stworzyłem bioinformatyczny start-up o nazwie **Biocloud Services**, którym kierowałem od 2015 do 2018 roku. W ciągu tych trzech lat działalności pełniłem funkcję CEO, podpisując dziesiątki umów z prywatnymi firmami, ośrodkami badawczymi i uczelniami. Wyniosłem z tego doświadczenia bardzo ważne umiejętności zarządzania ludźmi i zasobami, zarządzania ekonomicznego organizacją oraz umiejętności negocjacyjne.

Założenie nowego laboratorium od podstaw w obcym kraju, tak jak to robię od czasu przeprowadzki do Krakowa w listopadzie 2021 r., również było trudnym, intensywnym, ale satysfakcjonującym doświadczeniem. Przekształcenie pustego pokoju w laboratorium ze zmotywowanymi ludźmi i niezbędnym do prowadzenia badań sprzętem wymagało dużego nakładu pracy, wiele planowania, pracy z dokumentami i cierpliwości. Niemniej jednak, po roku od rozpoczęcia, sądząc po jakości osób, które dołączyły, ilości otrzymanych przeze mnie i członków laboratorium grantów oraz opublikowanych pracach, oceniam ten proces bardzo pozytywnie.

6.2 Osiągnięcia dydaktyczne i mentorskie

Naszym obowiązkiem jako naukowców jest nauczanie młodszego pokolenia, celem zapewnienia ciągłości wysoko wykwalifikowanych naukowców. Mając to na uwadze, mimo że nigdy nie miałem takiego obowiązku, angażowałem się w wiele inicjatyw dydaktycznych i mentorskich.

6.2.1 Wykładowca na Harvardzie GENED-1004:

Będąc na stażu podoktorskim (postdoc) na Harvardzie, jesienią 2019 r. byłem wykładowcą na kursie GENED-1004 zatytułowanym „Zrozumieć Darwinizm” (ang. „*Understanding Darwinism*”). Kurs ten, koordynowany przez dr Andrew Berry'ego (Department of Organismic and Evolutionary Biology) oraz dr Janet Browne (Department of the History of Science), był kursem otwartym dla wszystkich studentów studiów licencjackich Uniwersytetu Harvarda, w ramach którego zapoznaliśmy ich z tematem ewolucji. Kurs łączył elementy historii nauki i biologii molekularnej, by nauczyć studentów, jak funkcjonuje ewolucja.

Moje obowiązki na tym kursie obejmowały przygotowanie i nauczanie jednej 75 minutowej sekcji tygodniowo dla ok. 20 studentów, promowanie ich zaangażowania i debaty podczas

zajęć. Miałem również cotygodniowe godziny konsultacji, aby odpowiadać na indywidualne pytania studentów. Byłem również odpowiedzialny za ocenianie zestawów zadań i egzaminów.

Za tę rolę, na podstawie ocen wystawionych przez studentów, otrzymałem „Harwardzki Certyfikat Wyróżnienia w Nauczaniu” (ang. „*Harvard Certificate of Distinction in Teaching*”).

6.2.2 Wykłady gościnne

Kilkakrotnie byłem zapraszany do wygłaszania gościnnych wykładów na różnych kursach i studiach magisterskich na różnych uczelniach. Wybrane z tych wykładów gościnnych opisano poniżej:

- Winter school of the Faculty of Biochemistry, Biophysics, and Biotechnology “*small RNAs: when the small rule the bigger*”. Uniwersytet Jagielloński, luty 2022.
- Master in Omics Data Analysis. Tytuł wykładu: “*A scientific journey unveiling evolution with bioinformatics*”. University of Vic, 10 lutego 2022.
- Master in Bioinformatics and Biostatistics. “*Introduction to Git and GitHub*”. Open University of Catalonia-University of Barcelona, 4 marca 2021.
- Master in Omics Data Analysis. Tytuł wykładu: “*From UVic to the USA: a scientific journey unveiling evolution using bioinformatics*”. University of Vic, 3 lutego 2021.
- HSCI S-118: Darwin, Evolution and Society in the 19th and 20th centuries. Wykład “*Modern Synthesis*”. Harvard University, 25 lipca 2020.

6.2.3 Mentoring naukowy

Od 2017 roku jestem „współpracownikiem dydaktycznym” studiów magisterskich z bioinformatyki i biostatystyki, organizowanych wspólnie przez Open University of Catalonia i University of Barcelona. Jestem mentorem i prowadzę studentów studiów magisterskich w ich projektach końcowych prac magisterskich. Od rozpoczęcia pracy w 2017 roku nadzorowałem 14 projektów magisterskich. Ponadto jestem członkiem komisji oceniającej prace magisterskie innych studentów studiów magisterskich. W tym charakterze brałem udział w ocenie ponad 40 prac magisterskich.

W swoim laboratorium, na UJ, w ubiegłym roku byłem mentorem 1 licencjanta i 1 magistranta. Obecnie w laboratorium opiekuję się 2 studentami licencjackimi i 2 magisterskimi oraz jednym stażystą podoktorskim (postdoc). Dodatkowo w przyznanym mi grantie SONATA 17 posiadam środki na przyjęcie 1 doktoranta.

6.2.4 Wykładowca kursu online

W 2016 roku opracowałem kurs bioinformatyczny „Next-Generation Sequencing data analysis on cloud”, który był oferowany poprzez Bioclud. Mój wkład w powstanie tego kursu

polegał na opracowaniu sylabusów, przygotowaniu materiałów, ocenianiu ćwiczeń i odpowiadaniu na pytania studentów. Łącznie na zajęcia uczęszczało 92 studentów.

6.3 Popularyzacja nauki

Jako naukowcy mamy obowiązek raportowania naszych działań i wyników naszym interesariuszom, czyli społeczeństwu. Dlatego bardzo ważne jest, abyśmy pełnili aktywną rolę w popularyzacji nauki i upowszechnianiu naszych wyników. Z powodu tego silnego przekonania zaangażowałem się w niezliczone działania, by nauka i wiedza naukowa były dostępne dla ogółu społeczeństwa. Niektóre z najistotniejszych działań naukowych są wyszczególnione poniżej:

- Pisanie artykułów popularyzujących naukę w lokalnych czasopismach niespecjalistycznych takich jak artykuł, który opublikowałem w kwartalniku mojego rodzinnego miasta w kwietniu 2021 roku.
- Rozpowszechnianie wyników badań naukowych w gazetach i agencjach informacyjnych. W przypadku moich najważniejszych publikacji napisaliśmy komunikaty prasowe, które zostały wysłane do agencji prasowych i dziennikarzy, którzy publikowali w różnych mediach. Na przykład nasza najnowsza publikacja w PNAS była rozpowszechniana nie tylko na głównych portalach UJ i WBBiB, ale także w mediach w Polsce, takich jak TVP3, Forum Akademickie, Radio Opole, Nauka w Polsce; zaś w Hiszpanii w dużych portalach medialnych takich jak La Sexta TV czy La Vanguardia.
- Wikipedia jest głównym źródłem wiedzy dla wielu ludzi na całym świecie. Jako naukowiec przyczyniam się do tego, aby Wikipedia była aktualna i zgodna z wiedzą naukową w różnych językach, którymi biegle władam (kataloński, angielski i hiszpański). Piszę również nowe artykuły z mojej dziedziny i tłumaczę artykuły między trzema językami, którymi się posługuję.
- Odwiedzanie szkół i wyjaśnianie uczniom, co to znaczy być naukowcem i czym się zajmuje naukowiec, to dobry sposób na rozwijanie zainteresowań naukowych wśród młodych uczniów. Z tego powodu brałem udział w działaniach takich jak „Nauka w czasach szkolnych”, organizowanych przez Catalan Research Foundation.
- W różnych ośrodkach naukowych i na uczelniach, na których pracowałem, brałem udział w „dniach otwartych”, podczas których pokazujemy szerokiej publiczności obiekty badawcze i to, co w nich robimy.

7. Oprócz kwestii wymienionych w pkt. 1-6, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej

Poza 6 najważniejszymi publikacjami opisanymi powyżej, jestem autorem 12 innych artykułów. 18 publikacji zgromadziło w sumie 399 cytowań według stanu na luty 2023 r., a łączna suma *impact factor* wynosi 113 (**tabela 2**).

Tabela 2: Statystyki zbiorcze autorskich artykułów naukowych.

| | Podczas doktoratu | | | Po doktoracie | | | Wszystie | | |
|------------------------------------|-------------------|----|-----------|---------------|----|-----------|-------------------------------|-----|------------|
| | Liczba | IF | Cytowania | Liczba | IF | Cytowania | Liczba | IF | Cytowania |
| Badania | 5 | 34 | 298 | 11 | 68 | 96 | 16 | 102 | 394 |
| Przeglądowe / Rozdziały | - | - | - | 2 | 11 | 5 | 2 | 11 | 5 |
| Wszystkie | 5 | 34 | 298 | 13 | 79 | 101 | 18 | 113 | 399 |
| | | | | | | | Cytowania Łącznie: | | 399 |
| | | | | | | | Wyłączenie autocytowań | | 386 |
| | | | | | | | H-Index: | | 9 |

..... (Podpis wnioskodawcy)