

Recenzja Rozprawy Doktorskiej

Tytuł: Modelowanie komputerowe struktury trójwymiarowej chromatyny na podstawie danych genomicznych i mikroskopii wysokorozdzielczej

Autorka: mgr Zofia Tojek

Promotor: prof. dr hab. Dariusz Plewczyński

Tematyka badawcza

W swojej pracy doktorskiej mgr Zofia Tojek rozwiązuje problemy dotyczące trójwymiarowej struktury genomu. Łączy przy tym podejścia eksperymentalne z obliczeniowymi, co umożliwia uzyskanie wysokiej precyzji wyników oraz ich interpretowalność na gruncie genomiki. Autorka w szczególności zajmuje się modelowaniem trzeciorzędowej struktury chromatyny na podstawie danych eksperymentalnych pochodzących z mikroskopii super rozdzielczej, mikroskopii świetlnej oraz mikroskopii konfokalnej. Podejmuje również temat automatyzacji procesu przetwarzania danych otrzymywanych z HiChIP, ChIP-Seq (immunoprecypitacja chromatyny w połączeniu z sekwencjonowaniem) oraz NGS (sekwencjonowanie nowej generacji). Eksploracja przestrzennej struktury genomu w pracy przebiega dwutorowo: z jednej strony Doktorantka bada konformacje DNA na poziomie pojedynczych komórek, z drugiej analizuje dane genomiczne na poziomie populacji. Ta integracja podejść pozwala na kompleksowe badanie architektury chromatyny oraz dynamicznych interakcji występujących w obrębie genomu.

Układ rozprawy doktorskiej (w tym informacje o jej poszczególnych częściach składowych)

Rozprawa doktorska ma formę monografii naukowej. Składa się z ośmiu rozdziałów poprzedzonych streszczeniem w języku angielskim i jego odpowiednikiem w języku polskim. Streszczenie zarysowuje obszar badawczy doktoratu oraz sygnalizuje jakie są główne osiągnięcia Autorki. Pierwszy rozdział zawiera wprowadzenie w tematykę pracy, dość

szczegółowo przedstawiając problemy biologii komórkowej, a także eksperymentalne i obliczeniowe metody związane z modelowaniem i analizą struktury chromatyny. W drugim i trzecim rozdziale Autorka prezentuje swoje wyniki badawcze, przeplatając prezentację wyników szczegółowymi opisami kolejnych technik eksperymentalnych. W rozdziale drugim skupia się na modelowaniu struktury chromatyny na podstawie zdjęć mikroskopowych, natomiast w rozdziale trzecim – na analizie tej struktury w oparciu o dane sekwencyjne. Rozdziały te są opatrzone bardzo ładnymi ilustracjami (w sumie jest ich 35). W czwartym rozdziale znajdujemy wnioski oraz proponowane kierunki dalszych prac. Rozdział piąty zawiera deklaracje dotyczące wkładu Doktorantki w badania przedstawione w rozdziałach 3 i 4, a dodatkowo także w dwa współautorskie artykuły naukowe, których przedruki znajdują się na końcu rozprawy. Dalej (w rozdziale szóstym) zaprezentowano inne osiągnięcia naukowe Doktorantki, na które składają się prezentacje na pięciu konferencjach międzynarodowych, udział w dwóch grantach badawczych (NCN i FNP), otrzymane stypendia naukowe (brak informacji o ich liczbie) oraz współautorstwo artykułu opublikowanego w 2020 roku w *Nature Communications* (aktualny IF = 14,7; 200 pkt). Rozdział siódmy zawiera bibliografię, natomiast rozdział ósmy – przedruki dwóch artykułów, których współautorką jest mgr Zofia Tojek, z czego jeden został opublikowany w 2022 roku w *Scientific Reports* (IF = 3,8; 140 pkt), a drugi został w maju 2024 zgłoszony do *Nature Communications* i – w wersji zgłoszonej – jest dostępny w repozytorium bioRxiv.

Układ pracy mgr Zofii Tojek jest poprawny i w ogólności zgodny z powszechnie przyjętym schematem rozpraw doktorskich mających postać monografii. Nietypowe jest umieszczenie w takiej pracy przedruków artykułów naukowych oraz deklaracji wkładu Doktorantki – nie spotkałam się dotychczas z takimi suplementami w monografii i według mnie są one nadmiarowe. Uważam również, że praca zyskałaby, gdyby wszystkie podstawy teoretyczne były umieszczone w jednym miejscu (we Wprowadzeniu), a nie rozbite na trzy różne rozdziały. W rozprawie wyraźnie powinna zaznaczać się część, w której przedstawione są wyłącznie wyniki badawcze uzyskane podczas realizacji pracy doktorskiej. Przeplatanie wyników bardzo szczegółowymi opisami technik eksperymentalnych, z których pochodzą dane sprawia, iż odnosi się wrażenie, że samych osiągnięć nie ma zbyt wiele i trzeba ten fakt ukryć.

Zastosowane piśmiennictwo

Zastosowane piśmiennictwo jest ściśle związane z tematyką badań. Bibliografia zawiera 203 pozycje literatury, z których większość stanowią artykuły opublikowane w ostatnim dziesięcioleciu w wysoko punktowanych czasopismach z dziedziny nauk o życiu oraz bioinformatyki, m.in. *Bioinformatics*, *Briefings in Bioinformatics*, *Cell*, *Genome Research*, *Genes and Development*, *Nature*, *Nature Communications*, *Nature Biotechnology*, *Nature Genetics*, *Nature Protocols*, *Nature Reviews Genetics*, *Nucleic Acids Research*, *PLoS Computational Biology*, *Science*. Lista zawiera też nieliczne materiały internetowe. Większość publikacji przedstawia przede wszystkim metody oraz wyniki badań danych genomowych. Jest również kilka pozycji z dziedziny nauk ścisłych i technicznych. Spis jest uporządkowany zgodnie z kolejnością cytowania artykułów w treści rozprawy. Dobór bibliografii nie budzi zastrzeżeń. Dowodzi, iż Doktorantka posiada doskonałe rozeznanie w tematyce podjętej w rozprawie doktorskiej oraz zna aktualny stan wiedzy w swoim obszarze badawczym.

W spisie literatury znalazłam nieliczne drobne usterki i niespójności, które nie mają wpływu na pozytywny odbiór i czytelność pracy. Przykłady: tytuł książki pisany zwykłą czcionką a nie kursywą (poz. 14), nazwa tego samego czasopisma pisana w różny sposób w różnych miejscach (poz. 27 i 54, 60 i 62), wielocłonowa nazwa czasopisma z członami pisanymi z małej litery (poz. 99, 162), czasami wymienieni są wszyscy autorzy publikacji (np. 77, 89, 118, 134, 143), a innym razem tylko pierwszy z dopiskiem „et al.” (np. 25, 28, 41, 51, 69).

Cel pracy oraz zastosowane metody badawcze

Autorka nie zdefiniowała celu swojej pracy badawczej. Na podstawie przedstawionych wyników sędzę, że można go określić jako opracowanie metod bioinformatycznych do modelowania i analizy danych genomiki trójwymiarowej oraz ich zastosowanie do wybranych problemów praktycznych – tak zdefiniowany cel można uznać za osiągnięty. W samej rozprawie Doktorantka skupiła się w sposób szczególny na detalicznych opisach przeróżnych technik eksperymentalnych służących do pozyskiwania danych genomowych, co zapewne miało związek ze wskazaną przez nią intencją odmalowania jak najszerszego obrazu osiągnięć współczesnej genomiki trójwymiarowej. Ten cel został w rozprawie osiągnięty.

W pierwszej części swoich badań mgr Zofia Tojek skupiła się na przestrzennej strukturze chromatyny, w szczególności kładąc nacisk na modelowanie pętli chromatynowych. Warsztat badawczy obejmował w tym przypadku metody eksperymentalne, takie jak iPALM

(wysokorozdzielcze obrazowanie pętli chromatynowej zlokalizowanej w chromosomie 14 zrealizowane przez współpracujący z Doktorantką zespół z Jackson Laboratory) i 3D-EMISH (trójwymiarowa mikroskopia elektronowa z hybrydyzacją *in situ*), oraz metody obliczeniowe do przetwarzania danych eksperymentalnych (PeakSelector) i modelowania trzeciorzędowej struktury pętli chromatynowej na podstawie tych danych. Opracowanie tych ostatnich (narzędzia bioinformatyczne ChromoLooping i DensityDots) oraz ich aplikacja do danych eksperymentalnych były zadaniem Doktorantki. W implementacji narzędzi obliczeniowych Autorka wykorzystwała m.in. modele i metody grafowe, algorytm zachłanny dla problemu komiwojażera z funkcją optymalizacji, metody statystyczne (np. skalowanie wielowymiarowe), interpolację splajnami kubicznymi, miarę dystansową RMSD, indeks Jaccarda. Druga część pracy doktorskiej dotyczyła przetwarzania danych z sekwencjonowania nowej generacji pochodzących z eksperymentów Hi-C, Micro-C, Capture-C i HiChIP. Doktorantka opracowała potok obliczeniowy nf-HiChIP, w ramach którego wykorzystwała m.in. moduł obliczeniowy Luigi (do koordynowania złożonych zadań obliczeniowych), Nextflow (do tworzenia skalowanych, reprodukowalnych potoków obliczeniowych), algorytm Burrowsa-Wheelera (do mapowania sekwencji genomu referencyjnego), SAMtools rmdup (do usuwania duplikatów), BamCoverage (do wyznaczania pokrycia), IGV (do wizualizacji danych), systemy MAPS i Juicer (do analizy danych eksperymentalnych HiChIP). Ponadto mgr Zofia Tojek stosowała podczas pracy doktorskiej szereg innych narzędzi pomocniczych, m.in. SciPy (do wyznaczania otoczki wypukłej), Chimera (do wizualizacji strukturalnej), Stripenn i gStripe (do analizy wykrywania pasków chromatynowych), LoopSage (do symulacji ekstruzji pętli), HiCUP (do analizy odczytów w plikach .fastq), CHiCANE i CHiCAGO (do analizy danych z eksperymentów Capture-C).

Ogólnie rzecz biorąc, w swoich badaniach Doktorantka wykorzystwała metodologię powszechnie stosowaną we współczesnej bioinformatyce, która łączy przetwarzanie i modelowanie danych biologicznych, analizy statystyczne, analizy dużych zbiorów danych, algorytmikę, symulacje komputerowe oraz różnorodne techniki programistyczne. Uważam, że zastosowane metody badawcze są adekwatne do rozwiązywanego problemu i świadczą o dobrej znajomości przez Autorkę rozprawy nowoczesnych, efektywnych metod i technologii wykorzystywanych w naukach o życiu i naukach obliczeniowych.

Wyniki badań oraz ich praktyczne zastosowanie

W pracy doktorskiej mgr Zofia Tojek przeprowadziła szereg analiz różnego rodzaju danych eksperymentalnych, zapoznała się z problemami pojawiającymi się podczas tych analiz z wykorzystaniem istniejących metod, a następnie, na podstawie zebranych doświadczeń opracowała analityczne narzędzia obliczeniowe i przetestowała ich działanie w zastosowaniu dla różnych problemów współczesnej genomiki na dostępnych zestawach danych eksperymentalnych. Za jej najważniejsze osiągnięcia badawcze uznają stworzone metody bioinformatyczne, jak również wyniki ich zastosowań. Pierwszą metodą opracowaną przez Doktorantkę jest ChromoLooping – narzędzie do modelowania trójwymiarowej struktury chromatyny na podstawie danych z mikroskopii super-rozdzielczej, które pozwala na rekonstrukcję pętli chromatynowych z obrazów pojedynczej pętli oraz z wielodomenowych zdjęć mikroskopii świetlnej. Umożliwia ono również statystyczną analizę danych (m.in. liczby pików, jasności zdjęć, parametru sigma, grupowania pików). W wyniku kolejnych eksperymentów system był wielokrotnie ulepszany i wzbogacany o nową funkcjonalność. W celu ewaluacji został przetestowany i porównany z jedną z konkurencyjnych metod dostępnych na rynku. Jako swego rodzaju suplement do narzędzia ChromoLooping Doktorantka zaimplementowała algorytm DensityDots wspomagający odwzorowanie rzeczywistych pozycji zagęszczenia chromatyny. Następnie dokonała modelowania struktury pojedynczej pętli chromatynowej w wysokiej rozdzielczości, a także zaprezentowała metodę symulacji zdjęć mikroskopowych pętli chromatynowej na podstawie populacyjnych danych genomicznych. Trzecim narzędziem opracowanym w ramach doktoratu jest potok obliczeniowy nf-HiChIP, który służy do automatycznej kompleksowej analizy danych eksperymentalnych HiChIP przy uwzględnieniu braku dodatkowych niezależnych danych. Umożliwia on równoczesne przetwarzanie wielu próbek i integrację danych z różnych replikatów, co wpływa na znaczne usprawnienie analiz. W potoku nf-HiChIP łączą się tradycyjne techniki przetwarzania danych ChIP-seq (mapowanie, filtrowanie, identyfikacja pików i obliczanie pokrycia) i specjalizowane analizy HiChIP wykorzystujące funkcjonalność systemu MAPS. Narzędzie nf-HiChIP zostało zastosowane m.in. do analizy struktury przestrzennej genomu ludzkiego w skali populacji w ramach przetwarzania i eksperymentalnej weryfikacji limfoblastów wybranych rodzin z projektu 1000 Genomów. Dane przeanalizowane i wyselekcjonowane w ramach tego projektu stworzyły zbiór, który jest obecnie wykorzystywany przez metody uczenia maszynowego. Służy również jako źródło

wiarygodnych danych w badaniach pasków chromatynowych, domen chromatynowych, analizach wiązań białka CTCF, ekspresji genów itp.

Wszystkie narzędzia opracowane przez mgr Zofię Tojek są dostępne na platformie GitHub wraz z samouczkami i gotowymi przykładami użycia.

Wyniki badawcze uzyskane przez Doktorantkę w pracy doktorskiej zasługują na pozytywną ocenę. Niewątpliwym atutem rozprawy jest fakt, że mgr Zofia Tojek posiada już dwa dobrze opublikowane artykuły w wysoko punktowanych czasopismach naukowych. Według Web of Science praca opublikowana w 2022r. w *Scientific Reports* ma obecnie 12 cytowań, natomiast artykuł z 2020r. opublikowany w *Nature Communications* był cytowany 28 razy. Tym samym można stwierdzić, że mgr Zofia Tojek posiada H-indeks = 2, a jej publikacje były cytowane łącznie 40 razy. Wskaźniki te świadczą o tym, że w/w artykuły dotyczą aktualnych problemów badawczych, trafiają do właściwego grona odbiorców i mają wpływ na badania innych zespołów zajmujących się podobną tematyką.

Nieprawidłowości i braki w ocenianej rozprawie doktorskiej

Rozprawa doktorska mgr Zofii Tojek jest napisana ładnym i zrozumiałym językiem. W poszczególnych akapitach zachowana jest ciągłość wywodu. Opisy przeróżnych technik eksperymentalnych, eksperymentów, z których pochodziły dane do analizy (jak również tych, które nie zostały wzięte pod uwagę – np. ORCA), uzyskanych w pracy wyników są bardzo szczegółowe, co przydaje pracy walorów edukacyjnych. W pracy zauważam jednak pewne usterki, które komentuję poniżej:

- 1) Autorka nie podała *explicite* celu badań, ani kamieni milowych jakie zdefiniowała rozpoczynając pracę nad doktoratem. Choć nie jest to element konieczny rozprawy, to jednak zazwyczaj doktoranci formułują cele badawcze, co bardzo ułatwia czytelnikowi podążanie za przedstawioną historią badań oraz natychmiastowe uchwycenie sensu takiej a nie innej organizacji pracy i wybór warsztatu badawczego. W rozdziale podsumowującym znajduje się co prawda zdanie „Celem pracy było zaprezentowanie osiągnięć w dziedzinie genomiki trójwymiarowej na podstawie zarówno zdjęć mikroskopowych, jak i danych populacyjnych.” jednak odnosi się ono – jak mierniam – do samej monografii (a w zasadzie jej dużej części) a nie do pracy nad doktoratem. Celem doktoratu raczej nie jest popularyzacja wiedzy o cudzych osiągnięciach.

- 2) W rozprawie położono nacisk na dokładność i kompletność opisu przeróżnych technik eksperymentalnych. Brakuje tego samego w odniesieniu do narzędzi bioinformatycznych opracowanych przez Doktorantkę a będących jej głównymi osiągnięciami kwalifikującymi tę pracę doktorską do dziedziny nauk inżynieryjno-technicznych. Odczuwam niedosyt jeśli chodzi o opis narzędzia ChromoLooping. Bardzo lakonicznie przedstawiono, co to narzędzie robi, w związku z czym trudno je w pełni docenić. Jego opis jest rozrzucony w różnych miejscach pracy. Na przykład zaimplementowany w tym narzędziu algorytm zachłanny został z nieznanymi powodów przedstawiony w rozdziale pierwszym (w sekcji 1.4.2.4 zatytułowanej „Problem komiwojażera”), zamiast w rozdziale drugim, który jest poświęcony wynikom uzyskanym w doktoracie. Chciałabym, aby Autorka przedstawiła to narzędzie pokazując wszystkie jego możliwości oraz zarys działania krok po kroku. Warto byłoby je porównać z innymi podobnymi narzędziami o ile istnieją. Tego również trudno się doszukać w rozprawie. Choć w rozdziale, w którym wspomniano o testowaniu ChromoLooping w skali wielodomenowej (2.2.1) znajduje się wzmianka o porównaniu narzędzia z innym podejściem, nie ma w tym rozdziale szerszej analizy ilościowej i jakościowej wyników tego eksperymentu obliczeniowego.
- 3) Działanie algorytmu dobrze jest przedstawiać w postaci schematu i uszczegóławiać w komentarzu do rysunku. Sam opis jest zwykle mniej czytelny i mało atrakcyjny dla czytelnika. Na przykład algorytm BWA (str. 30) lub algorytm głównej pętli przetwarzającej dane w narzędziu ChromoLooping (str. 50) można było przedstawić właśnie w postaci schematu ukazującego poszczególne etapy obliczeniowe i przepływ danych. Znacząco poprawiłoby to klarowność przekazu.
- 4) W rozdziale 2 bardzo przydałby się schemat przepływu danych ukazujący w sposób skondensowany kolejne kroki podjęte przez Doktorantkę. Czytając opis trudno było się zorientować co po kolei było wykonywane w pracy, odniosłam wrażenie, że w niektórych miejscach Autorka wraca do kroków opisanych wcześniej. Przy tak szczegółowym opisie różnych eksperymentów rozmywa się ogólny zarys postępowania w toku badań.
- 5) Szczegółowe opisy technik eksperymentalnych powinny się znajdować w rozdziale pierwszym, natomiast rozdział drugi i trzeci powinny przedstawiać wyłącznie przeprowadzone badania oraz ich wyniki. Tymczasem w tych dwóch ostatnich rozdziałach Autorka wplata opisy eksperymentów w akapity prezentujące sam przebieg badań. Działa

to na niekorzyść pracy, gdyż powoduje, że czytelnik gubi się w opisach oraz trudno mu wyłuskać co dokładnie zrobiono w ramach doktoratu i jak przebiegał proces badawczy.

- 6) Odnośnie algorytmu zachłannego (str. 37) – warto byłoby dopisać, że jest to heurystyka, gdyż bez tej informacji niezorientowany czytelnik nie wie, do której z wymienionych wcześniej kategorii należy podejście zachłanne. Ponadto opis algorytmu zachłannego podany w punktach 1-3 jest nieprecyzyjny: (a) czym jest „kawałek ścieżki”? (b) w punkcie 3 napisano „powtarzaj” bez podania, które kroki są powtarzane (powinno być: powtarzaj kroki 2-3). Skoro akurat ten algorytm został wykorzystany w pracy, to powinien on być przedstawiony w sposób kompletny i dokładny. Wreszcie: należałoby podać jaka jest złożoność pierwotnej wersji algorytmu, zamiast dość ogólnikowo stwierdzać, że znajduje on rozwiązanie w kilka minut (zwłaszcza, że dalej podana jest złożoność funkcji optymalizującej).
- 7) Odnośnie RMSD napisano: „Zlicza uśredniony dystans pomiędzy dwoma odpowiadającymi sobie atomami w dwóch strukturach.” Jest to dość mylące sformułowanie. Powinno być raczej: „liczy uśredniony dystans między dwoma strukturami” lub dokładniej: „uśrednia sumę dystansów między wszystkimi odpowiadającymi sobie atomami w dwóch strukturach”.
- 8) Autorka podaje, że opracowane przez nią narzędzia obsługują dane w formacie PDB. Warto pamiętać, że od 2014 roku PDB nie jest już podstawowym formatem do przechowywania współrzędnych atomowych. Dlaczego zdecydowano się dedykować programy zaimplementowane w pracy do obsługi formatu PDB a nie CIF?
- 9) W niektórych miejscach znalazłam literówki, brakujące części zdania, powtórzenia, nieprawidłową odmianę wyrazów itd. W niektórych dłuższych fragmentach brakowało przecinków. Te usterki nie mają wpływu na klarowność przekazu. Przykłady:
 - W pracy nie ma rysunku z numerem 20, są natomiast dwa rysunki z numerem 21.
 - str. 9: „W projekcie uzyskać precyzję lokalizacji aż do 2 nm”
 - str. 13: „idące za nim konsekwencja (..) wzbudza”
 - str. 13: „z wieloma chorobami ludzkim”
 - str. 16: „w utrzymania stabilności genomu”
 - str. 16: „powiązane zosyały”
 - str. 18: „chromatynowej[Rys.2A]” (brak spacji)

- str. 19: „Kohezyna i Ekstruzja” (ekstruzja powinna być pisana z małej litery)
- str. 20: „Ich rola w regulacji kluczowych procesów (...) powodują”
- str. 20: „technologie długich odczytu”
- str. 25: „ilość materiału (komórek) potrzebnych”
- str. 26: „jest jednym z celi”
- str. 26: „białkiem.Finalnie” (brak spacji)
- str. 26: „co umożliwiło stworzenia modeli ułożenia (...) i zbudowania”
- str. 29: „Prowadzi do do powstania”
- str. 29: „mapowanie na z wykorzystaniem”
- str. 31: „w skali całego genom.”
- str. 31: „systematyczne błędy systematyczne”
- str. 32: „mechanizmy regulacyjne regulujące”
- str. 34: „mapa kontaktów z eksperymentów Hi-C, HiChIP lub ChIA-PET reprezentujące”
- str. 35: „Problem komiwojażera (...) jest klasycznym problemem optymalizacyjnym. Jej nazwa”
- str. 36: „minimalnego drzewa rozpinającego na miastach” (chodzi raczej o minimalne drzewo rozpinające w grafie, którego wierzchołki reprezentują miasta)
- str. 38: „rotacje i przesuwanie struktur” (lepiej „rotacje i translacje struktur” lub „obracanie i przesuwanie struktur”)
- str. 44: „W przestrzeni między genowej”
- str. 44: „Top 5 najsilniejszych pętli” (wystarczy napisać „5 najsilniejszych pętli”)
- str. 49: „maksymalnej liczby następujących po sobie klatek nie posiadających sygnału przypisanego do danej grupy dzielących sygnały przypisane do tej samej grupy” (brak przecinków ułatwiających zrozumienie tego fragmentu)
- str. 60: „Zlokalizowanie brakujących fragmentów z zdjęciach”
- str. 60: „nie jest możliwe ponieważ eksperyment nie pozwala na stwierdzenie które sondy skutecznie związały się do chromatyny” (brak przecinków)
- pipeline = potok obliczeniowy (sugeruję używać polskich określeń jeśli istnieją)
- str. 73: „in situ” powinno być pisane kursywą

- str. 82 „algorytmu problemu komiwojażera” (dość dziwne i raczej nieużywane sformułowanie; problem komiwojażera to jedno, a algorytm, który go rozwiązuje to drugie; algorytmów rozwiązujących TSP jest wiele, więc lepiej pisać konkretnie o który algorytm chodzi)
- str. 82: „ścieżki.Wprowadzenie” (brak spacji)
- str. 88: „skupiam się głównie na analizie wyników sekwencjonowanie”
- str. 89: „zawsze mamy dostęp (lub samodzielnie wyprodukujemy) dane”
- str. 93: „W przypadku eksperymentach”
- str. 97: „wykorzystana zostaną została wersja BWA-MEM”
- str. 102: „Druga część, zwana MAPS, wykonuje podział (...) oraz identyfikuję”
- str. 109: „HiChIP. najmniejszą”
- str. 112: „Teza ta był jednym z ważniejszych”
- str. 115-116: „Wyniki potwierdzają wcześniejsze również wcześniejsze doniesienia”
- str. 120: „pasków chromatinowych”
- str. 122: „chromatyny.Podsumowując” (brak spacji)
- W wielu miejscach Autorka używa w opisie badań czasu przyszłego. Niezupełnie rozumiem, czy oznacza to, że przedstawia jakieś przyszłe badania lub możliwości, czy po prostu błędnie stosuje czas przyszły zamiast przeszłego. Przykład „Natalia Zawrotna przeprowadziła eksperymenty, stosując metody takie jak in situ Hi-C22, spike ATAC-seq, RNA-seq oraz CHIP-seq. Dzięki temu będziemy mogli zmierzyć wpływ SIRT6 na dostępność genomu i organizację przestrzenną, ekspresję genów oraz położenie białek na DNA. Moją rolą w projekcie jest zaprojektowanie oraz przeprowadzenie analiz na danych Hi-C oraz CHIP-seq oraz połączenie wyników wykonanych analiz z danymi ATAC-seq. Dodatkowo na etapie przygotowywania próbek Hi-C brałam udział ...” (w pracy znalazłam wiele podobnych fragmentów, gdzie w następujących po sobie zdaniach czas przeszły miesza się z przyszłym i/lub teraźniejszym).

Dodatkowo, kilka drobnych uwag odnoszących się do wykazu literatury zawarłam w tej części recenzji, która dotyczy piśmiennictwa.

Wnioski końcowe

Czytając rozprawę doktorską mgr Zofii Tojek odniosłam wrażenie, że doktorat ten był wyjątkowo trudny. Wymagał od Doktorantki ogromu żmudnej i systematycznej pracy, która

niejednokrotnie nie przynosiła oczekiwanych lub satysfakcjonujących rezultatów, w związku z czym nie mogła być podsumowana publikacją. Do jego realizacji potrzebne było zapoznanie się z licznymi technikami eksperymentalnymi na poziomie, który umożliwiał analizę, walidację oraz interpretację otrzymywanych danych, a następnie również ich modelowanie *in silico*. Podziwiam i doceniam wytrwałość Doktorantki, która przebija między wierszami recenzowanej monografii. Jej Autorka wykazała się zdolnością do prowadzenia badań naukowych, poprawnej analizy danych i formułowania wartościowych wniosków badawczych. Udowodniła, że bardzo dobrze poznała aktualny stan wiedzy w zakresie tematyki swojej pracy doktorskiej. Posiada solidną wiedzę teoretyczną z zakresu biologii molekularnej, genomiki, bioinformatyki, badań operacyjnych oraz technologii stosowanych w analizach danych genomicznych. Wykazuje umiejętność samodzielnego prowadzenia badań naukowych, pracy w zespole badawczym, a także stosowania zdobytej wiedzy w rozwiązywaniu problemów praktycznych.

Recenzowana rozprawa zawiera oryginalne rozwiązanie problemu naukowego. Duża część uzyskanych przez Autorkę wyników badawczych stanowiących trzon rozprawy doktorskiej została przedstawiona w wieloautorskich artykułach naukowych, z których dwa zostały opublikowane w bardzo dobrych czasopismach dziedzinowych, a trzeci został w tym roku zgłoszony do publikacji i jest w fazie recenzji. Mgr Zofia Tojek jest pierwszą autorką dwóch z tych prac (w jednym przypadku pierwsze autorstwo jest współdzielone).

Pozytywnie oceniam pracę mgr Zofii Tojek pt. „Modelowanie komputerowe struktury trójwymiarowej chromatyny na podstawie danych genomicznych i mikroskopii wysokorozdzielczej” i stwierdzam, że spełnia ona wymagania stawiane rozprawom doktorskim określone w art. 13.1 Ustawy o stopniach naukowych i tytule naukowym z dnia 14.03.2003 oraz stanowi oryginalne rozwiązanie przez autorkę zagadnienia naukowego. Wnoszę o dopuszczenie mgr Zofii Tojek do kolejnego etapu postępowania kwalifikacyjnego w celu nadania stopnia naukowego doktora w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja.

.....
prof. dr hab. inż. Marta Szachniuk