

Gliwice, 12.08.2023

Recenzja rozprawy doktorskiej mgr inż. Michała Własnowolskiego
pt. *Computational Modelling and Analysis of the Three-Dimensional Structure*
of Human Genome at the Population Scale

Ukończonej na Wydziale Matematyki i Nauk Informacyjnych
Uniwersytetu Warszawskiego

Pod opieką promotora prof. dr hab. Dariusza Plewczyńskiego

Zawartość pracy

Trójwymiarowa architektura chromatyny w jądrze komórkowym jest ważnym czynnikiem regulacji i funkcjonowania genomu. Zaawansowane technologie mapowania całego genomu takie jak Hi-C czy ChIA-PET pozwalają eksperymentalnie badać interakcje chromatyny i dostarczają danych, które można wykorzystać do rekonstrukcji trójwymiarowej struktury genomu. W szczególności wynikiem tych wysokoprzepustowych technologii są dane w postaci częstotliwości interakcji parami między loci genomowymi, które muszą zostać przekształcone we względne odległości fizyczne. Ponieważ te techniki eksperymentalne dostarczają danych na temat kontaktów chromatynowych, a nie bezpośrednich informacji strukturalnych, istnieje potrzeba tworzenia algorytmów oraz narzędzi obliczeniowych potrafiących wykorzystać informacje na temat częstotliwości kontaktów w celu zamodelowania struktury 3D. Możliwość modelowania struktury trójwymiarowej chromatyny jest szczególnie ważna w przypadku próby zrozumienia jak zmienność strukturalna związana z występowaniem delecji, insercji, duplikacji oraz inwersji wpływa na przestrzenną organizację genomu, co ma w konsekwencji wpływ na mechanizmy regulacji transkrypcji genów. W tym kontekście szczególnie interesująca jest możliwość analizy przestrzennych interakcji chromatyny związanych z regionami *enhancerów* i *promotorów* genów, które odgrywają kluczową rolę w regulacji transkrypcji genów.

Przedstawiona praca doktorska została złożona w postaci zbioru opublikowanych i powiązanych tematycznie artykułów naukowych. Na cykl publikacji składają się cztery artykuły naukowe, trzy z nich zostały opublikowane w bardzo dobrych czasopismach naukowych, czwarty znajduje się w recenzji. W dwóch opublikowanych artykułach i w trzecim, który jest w recenzji, doktorant jest pierwszym autorem.

Pierwszy z artykułów [P1] pt. „*Spatial chromatin architecture alteration by structural variations in human genomes at the population scale*” opublikowany został w czasopiśmie *Genome Biology* (IF 18.01, 200 pkt MNiSW). W pracy tej przeanalizowano po raz pierwszy w skali populacyjnej częstość występowania zmienności wariantów strukturalnych i ich wpływ na organizację chromatyny oraz zmienność jej topologii. W celu stworzenia modeli 3D chromatyny wykorzystana jest autorska metoda 3D-GNOME, która modyfikuje referencyjne interakcje genomowe i domeny topologiczne uzyskane na podstawie danych eksperymentalnych wprowadzając informacje o zmienności strukturalnej wynikającej ze zmienności genetycznej. W pracy umieszczono też szeroką populacyjną analizę wariantów strukturalnych.

Drugi artykuł [P2] składający się na cykl publikacji pt. *3D-GNOME 2.0: a three-dimensional genome modeling engine for predicting structural variation-driven alterations of chromatin spatial structure in the human genome* opublikowany został w czasopiśmie *Nucleic Acid Research* (IF=19.16, 200 pkt. MNiSW), sekcja *Web Server Issue*. W pracy tej przedstawiono narzędzie implementujące algorytm 3D-GNOME w wersji 2.0. Algorytm ten szczegółowo został opisany w pracy P1 a w połączeniu z opublikowanym narzędziem umożliwia użytkownikom przewidywanie zmian w trójwymiarowej konformacji chromatyny dla wybranego regionu ludzkiego genomu. Użytkownik może również przesłać własną listę wariantów strukturalnych w celu ich zamodelowania.

Kolejny artykuł [P3] składający się na cykl publikacji opublikowany został w formie nierecenzowanego preprintu na serwerach bioRxiv oraz został przesłany do recenzji w czasopiśmie *Bioinformatics* (IF 6.931, 200 pkt MNiSW). Jest to praca zatytułowana „*cudaMMC - GPU-enhanced Multiscale Monte Carlo Chromatin 3D Modelling*”. W pracy tej przedstawiono modyfikację algorytmu 3D-GNOME w celu przyspieszenia obliczeń. Przyspieszenie uzyskano poprzez zrównoleglenie obliczeń z wykorzystaniem architektury GPU.

Ostatni, czwarty artykuł [P4] składający się na cykl publikacji, pt. *3D-GNOME 3.0: a three-dimensional genome modelling engine for analysing changes of promoter-enhancer contacts in the human genome* opublikowany został w czasopiśmie *Nucleic Acid Research* (IF=19.16, 200 pkt MNiSW), sekcja *Web Server Issue*. Praca ta przedstawia kolejną wersję narzędzia 3D-GNOME. Narzędzie rozszerzone zostało o nowy, znacznie większy zbiór danych interakcji chromatynowych, a także zaktualizowane warianty strukturalne do najnowszej wersji genomu ludzkiego GRCh38. Istniejący silnik zastąpiono algorytmem *cudaMMC* oraz zintegrowano serwis z klastrem obliczeniowym co umożliwiło do 25x przyspieszenie obliczeń. Serwis został też wzbogacony o narzędzia do wizualizacji oraz porównywania wyników.

Opinia o rozprawie

Należy podkreślić, iż problem, który podjąć się rozwiązać Doktorant jest niewątpliwie ważny i wpisujący się w trendy najnowszych badań w dziedzinie. Rozwój algorytmów oraz narzędzi do przewidywania oraz wizualizacji struktury trójwymiarowej chromatyny jest pierwszym krokiem do zrozumienia w jaki sposób zmiany struktury przestrzennej chromatyny wpływają na mechanizmy regulacji transkrypcji genów. Analiza zróżnicowania populacyjnego wariantów strukturalnych jest ważna w kontekście poznania patogenego potencjału wariantów strukturalnych zmieniających organizację chromatyny wyższego rzędu.

Przedstawione w ramach pracy doktorskiej artykuły stanowią niewątpliwie spójny cykl powiązanych ze sobą publikacji w tematyce rozprawy doktorskiej.

Dobór artykułów i kolejność prezentacji wyników badań jest spójny i logiczny. Pierwsza praca zawiera analizę częstości występowania zmian strukturalnych chromatyny w populacji i motywację do wprowadzenia narzędzia 3D-GNOME. Następne publikacje przedstawiają kolejne ulepszone wersje algorytmu. Opublikowane z sukcesem wyniki badań publikacje świadczą o tym, iż Doktorant opanował warsztat badawczy i potrafi realizować badania naukowe.

Spis literatury dobrze oddaje aktualny stan wiedzy w zakresie, którego dotyczy rozprawa. Wiele z cytowanych prac to pozycje najnowsze co świadczy o śledzeniu przez Doktoranta na bieżąco literatury przedmiotu.

Przedstawiony w rozprawie cel, który brzmi:

Opracowanie, wdrożenie i współtworzenie narzędzi obliczeniowych służących do generowania i analizy trójwymiarowych modeli chromatyny oraz badania potencjalnego wpływu struktury przestrzennej chromatyny na aktywność genetyczną komórek.

jest jasno sformułowany, a przedstawiony cykl artykułów świadczy o tym, że Doktorantowi udało się go osiągnąć.

Uwagi krytyczne i dyskusyjne

Na początku tej części chciałbym podkreślić, że nie znalazłem w przedstawionych wynikach żadnych zasadniczych błędów merytorycznych czy niewłaściwych rozumowań. Wszystkie poniższe uwagi wynikają z chęci podjęcia dyskusji i dialogu na temat niektórych aspektów pracy. Uwagi te nie obniżają mojej pozytywnej oceny pracy.

Charakter publikacji w czasopiśmie Nucleic Acid Research Web Server Issue, powoduje, że skupiają się one bardziej na przedstawieniu cech narzędzia niż analizie metody. Po pierwsze, chciałabym zaznaczyć, że tego typu publikacje są bardzo ważne, gdyż udostępniają daną metodę szerokiemu środowisku naukowemu

przyczyniając się do powstawania nowych wyników i wniosków nie tylko w grupie badawczej, w której została stworzona metoda. Niemniej jednak pewną wadą tych publikacji jest brak pogłębionych analiz wyników udostępnionych metod. Nie jest więc jasne czy dla nowych wersji algorytmu Doktorant wykonał analizy podobne do analiz przedstawionych w pracy [P1]? Czy wprowadzenie nowych modyfikacji do algorytmu - poza przyspieszeniem obliczeń i aktualizacją do nowej wersji genomu - pozwoliło uzyskać dodatkową wiedzę biologiczną lub nowe wnioski odnośnie wpływu wariantów strukturalnych na strukturę przestrzenną genomu w analizie populacyjnej?

Niedosyt budzi też niezbyt konkretne przedstawienie przez Doktoranta wkładu własnego w powstanie publikacji. W przypadku pracy doktorskiej przedstawionej jako cykl powiązanych tematycznie publikacji jest to niezwykle istotne, ponieważ pozwala ocenić jaka część z opublikowanych prac stanowi oryginalny wkład Doktoranta. Załączono co prawda oświadczenia współautorów, niemniej są one bardzo ogólne (przykładowo jest to koordynacja zespołu, konceptualizacja i metodologia – wspólnie z Promotorem) lub obejmują zadania mocno techniczne takie jak integracja zbioru danych czy formatów danych. Przedstawione publikacje są wieloautorskie, co w przypadku tak złożonych metod i narzędzi jest normalną praktyką, jednak Doktorant powinien w jasny sposób zaakcentować swój twórczy wkład w powstanie publikacji, a w szczególności w rozwój kolejnych wersji algorytmu 3D-GNOME.

Wnioski

Pan mgr inż. Michała Własnowolski przedstawił zbiór opublikowanych i powiązanych tematycznie artykułów naukowych stanowiących oryginalne rozwiązanie problemu naukowego z zakresu trójwymiarowego modelowania chromatyny w kontekście analizy wariantów strukturalnych. Należy podkreślić, iż opracowane przez Doktoranta metody udostępnione zostały środowisku naukowemu w postaci serwisów internetowych.

Poza publikacjami składającymi się na pracę doktorską, Doktorant jest również współautorem trzech publikacji opublikowanych w bardzo dobrych czasopismach naukowych (odpowiednio 100, 140 oraz 100 punktów MNiSW) oraz jednej publikacji konferencyjnej.

Biorąc pod uwagę powyższą ocenę, stwierdzam, że przedstawiona do oceny praca doktorska w pełni odpowiada warunkom określonym w Art. 187 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (tekst jednolity Dz. U. z 2023 r. poz. 742 z późn. zm.) i na tej podstawie wnoszę do Wysokiej Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej o dopuszczenie mgr inż. Michała Własnowolskiego do dalszych etapów przewodu doktorskiego.



dr inż. hab. Aleksandra Gruca