

Recenzja

rozprawy doktorskiej mgr Michała Szczęśniaka

pt. „Nowe metody identyfikacji mikro RNA” (ang. *„New methods for identification of microRNAs”*)

dla Rady Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu

Stosunkowo niedawne odkrycie regulacyjnych funkcji małych niekodujących cząsteczek RNA otworzyło nowy rozdział poszukiwań w biologii molekularnej i jest prawdopodobnie dopiero jednym z pierwszych tak ważnych znalezisk w „śmieciowym” DNA. W ostatnich latach wiele uwagi poświęcono zarówno molekularnym jak i bioinformatycznym analizom miRNA, co umożliwiło identyfikację wielu nowych cząsteczek i ujawniło bogactwo regulowanych przez nie funkcji i procesów komórkowych.

Rozprawa doktorska mgr Michała Szczęśniaka poświęcona jest opisaniu współtworzonych przez niego metod bioinformatycznych służących do identyfikacji miRNA.

Pierwszym z opracowanych przez doktoranta narzędzi służących do identyfikacji miRNA jest HuntMi – algorytm oparty o techniki uczenia maszynowego. Należy podkreślić, że algorytmy uczenia maszynowego były już wcześniej używane do tej klasy problemów. Autorzy HuntMi zidentyfikowali słabe punkty opublikowanych do tej pory implementacji, do których należą zarówno zbyt restrykcyjne ograniczenia narzucane tym algorytmom przez samych twórców (np. dotyczące struktury miRNA lub klasy stosowalności), jak i problemy metodologiczne związane z nierównowagą liczby przykładów pozytywnych i negatywnych. Z problemem niezrównoważonego uczenia radzą sobie Autorzy metody wprowadzając procedurę ROC-select polegającej na połączeniu dość standardowego klasyfikatora z krzywymi charakterystyki operacyjnej odbiornika (ROC). Wprowadzenie procedury ROC-select wraz z dodaniem 7 nowych (do 21 znanych) cech bazowych, pozwoliło na znaczące poprawienie czułości klasyfikacji w porównaniu do dotychczas istniejących rozwiązań. HuntMi jest publicznie dostępnym narzędziem z istniejącymi wzorcami dla zwierząt, roślin i wirusów, jednocześnie umożliwiając budowę własnego modelu przez użytkownika.

Drugie ze zbudowanych przez doktoranta narzędzi służy do przewidywania nowych sekwencji potencjalnych miRNA przez poszukiwanie podobieństwa do znanych już sekwencji. Autorzy zbudowali potok analityczny pozwalający na poszukiwanie nowych miRNA, podstawą którego są sekwencje EST znajdujące się jako podobne do znanych miRNA. Większość z reszty wykonywanych w potoku operacji (poza RNAfold) to usuwanie przypadkowo znalezionych lub znanych sekwencji. Uzyskane wskaźniki zarówno specyficzności jak i czułości przewidywania są dobre, a sama baza danych miRNEST wyróżnia się dużą liczbą uwzględnionych gatunków i przewidzianych sekwencji miRNA.

W ostatniej części rozprawy opisuje doktorant bardzo prostą i jednocześnie bardzo ciekawą analizę bioinformatyczną genów kodujących miRNA w celu identyfikacji intronów w tychże genach. Analiza 7 gatunków roślin (MEGABLAST sekwencji EST do pre-miRNA z bazy miRBase, a następnie porównanie z sekwencją genomową) pozwoliła na znalezienie 45 genów kodujących miRNA i zawierających introny (u 5 gatunków roślin). W utworzonej bazie danych ERISdb (zawierającej także dane z innych badań) znalezione introny pokazane są w ciekawej formie ukazującej genomowy kontekst, w którym występuje dany intron.

Rozprawa doktorska mgr Michała Szczepniaka przedstawiona jest w formie 4 wieloautorskich publikacji opatrzonych krótkim omówieniem w języku polskim i angielskim. Trzy, syntetycznie omówione przeze mnie powyżej, publikacje ukazały się w dobrych czasopismach o międzynarodowym zasięgu. Czwarta publikacja doktoranta to praca przeglądowa w języku polskim umieszczona w *Postępie Biochemii*.

We wszystkich 4 pracach mgr Szczepniak jest ich pierwszym autorem, w jednym przypadku jest to miejsce dzielone z mgr Adamem Gudysiem. Częścią rozprawy są: oświadczenie doktoranta dotyczące jego udziału w powstaniu tych prac oraz oświadczenia współautorów. Oświadczenia te są niesprzeczne i pozwalają stwierdzić, że wkład doktoranta w powstanie opisanych wyników jest dominujący.

Rezultaty przedstawione w rozprawie doktorskiej mgr Michała Szczepniaka oraz zadeklarowany wkład w ich uzyskanie pokazują, że doktorant bardzo dobrze opanował wiele technik bioinformatycznych. Zbudowane przez niego dwa narzędzia do identyfikacji miRNA oraz baza ERISdb o unikalnej zawartości będą z pewnością używane przez społeczność naukową. Godny podkreślenia jest także fakt publicznego udostępnienia tych narzędzi bez żadnych ograniczeń.

W podsumowaniu stwierdzam, że rozprawa doktorska mgr Michała Szczęśniaka spełnia z naddatkiem wszelkie ustawowe i zwyczajowe wymogi i stawiam wniosek o dopuszczenie doktoranta do dalszych etapów przewodu. Jednocześnie, biorąc pod uwagę wysoką wartość uzyskanych w pracy wyników, jak i fakt ich opublikowania w renomowanych czasopismach, zwracam się do Rady Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu z prośbą o rozważenie możliwości jej wyróżnienia i/lub wystąpienia o stosowną nagrodę.

A handwritten signature in blue ink, appearing to read 'Szczęśniak'.