

Tytuł: Pary LINE-LINE jako substraty nie-allelicznej homologicznej rekombinacji

Promotorzy:

prof. Anna Gambin, Instytut Informatyki, Uniwersytet Warszawski
tel.: +48 22 5544212; e-mail: aniag@mimuw.edu.pl

prof. Paweł Stankiewicz, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA
tel.: 001-713-798-5370; e-mail: pawels@bcm.edu

dr Michał Startek (promotor pomocniczy), Instytut Informatyki, Uniwersytet Warszawski
tel.: +48 609401659 e-mail: mist@mimuw.edu.pl

Badania przeprowadzone w ramach projektu mają obejmować analizę całego genomu pod kątem możliwości mediowania rearanzacji przez pary transpozonów typu LINE (Long Interspersed Elements), mechanizmów powstawania tych rearanzacji i ich wpływu na zmienność genetyczną oraz występowanie chorób genetycznych. W projekcie zostaną wykorzystane dane pochodzące ze źródeł dostępnych publicznie oraz prywatnych wyników doświadczeń wielkoskalowych opartych na sekwencjonowaniu następnej generacji oraz genomowej hybrydyzacji porównawczej do mikromacierzy (aCGH - array Comparative Genomic Hybridization). Dane te zostaną poddane analizie statystycznej oraz będą użyte w symulacjach komputerowych i algorytmach uczenia maszynowego.

W ostatnich latach wysiłki naukowców skupiają się na wyjaśnieniu mechanizmów leżących u podłoża rearanzacji genomowych. Najnowsze badania wskazują, że niealleliczna homologiczna rekombinacja (NAHR) może być powodowana nie tylko przez pary LCRów (Low-Copy Repeats), ale również przez pary transpozonów typu LINE. Znaczna ilość genomu ludzkiego może zatem być podatna na rearanzacje mediowane przez pary LINE-LINE ze względu dużą liczbę tych transpozonów w genomie (transpozony LINE typu L1 stanowią około 17% genomu referencyjnego) oraz bardzo duże wzajemne podobieństwo.

Pary LINE-LINE mogą być substratami nie-allelicznej homologicznej rekombinacji (NAHR) zarówno pomiędzy chromosomami (translokacje) jak i w obrębie jednego chromosomu (powodując inwersje, delecje lub duplikacje). Wystąpienie zdarzenia NAHR powodującego polimorfizm zróżnicowania liczby kopii może być wykryte z wykorzystaniem mikromacierzy DNA. Pozostałe typy wariantów strukturalnych (inwersje, translokacje) wymagają opracowania innych metod ich detekcji.

Symulacje przeprowadzone wstępnie będą polegały na identyfikacji par LINE-LINE mogących mediować delecje lub duplikacje. Pary transpozonów typu LINE posiadające odpowiednie właściwości (np. wysokie podobieństwo obydwu sekwencji, zgodne zorientowanie) można znaleźć w bazie danych dostępnej publicznie. Polimorfizmy zróżnicowania liczby kopii u pacjentów zostaną zidentyfikowane na podstawie analizy oligonukleotydowych chromosomowych mikromacierzy DNA w Medical Genetics Laboratories w Baylor College of Medicine przez grupę prof. Pawła Stankiewicza.

Do analizy wariantów strukturalnych mediowanych przez pary LINE-LINE wykorzystane dane, doświadczenia i narzędzia wypracowane w Instytucie Informatyki Uniwersytetu Warszawskiego oraz w Department of Molecular and Human Genetics w Baylor College of Medicine.

Przeprowadzone badania pozwolą na dokładniejsze poznanie procesów rekombinacji do których dochodzi w komórkach, zarówno rekombinacji nieallelicznych jak i występujących podczas zwyczajnych procesów mejozy, a także czynników które im sprzyjają. Rzuci ona również nowe światło na związek pomiędzy transpozonami jako czynnikami roznoszącymi „hotspoty” rearanżacji a ewolucją organizmów. Zdobyta w ten sposób wiedza będzie miała szerokie konsekwencje, w tak różnorodnych dziedzinach jak genetyka człowieka, medycyna chorób genetycznych, modelowanie ewolucji organizmów a także powstanie mechanizmu rozmnażania płciowego.

Interdyscyplinarność projektu wynika z konieczności połączenia wiedzy biologicznej, dotyczącej mechanizmów transpozycji, rekombinacji oraz zaburzeń tych mechanizmów z zaawansowanymi metodami matematycznymi oraz informatycznymi pozwalającymi na modelowanie tych zjawisk. Konieczne będzie także wykonanie porównawczej analizy genomów ssaków naczelnych przy użyciu wyrafinowanych metod bioinformatycznych oraz statystycznych w celu wyszukania ewolucyjnych śladów takich zjawisk.

Główne zadania badawcze stojące przed doktorantem to:

- analiza całego genomu pod kątem możliwości mediowania rearanżacji przez pary LINE-LINE
- opracowanie metod identyfikacji rearanżacji genomu powodujących CNV mediowanych przez pary LINE-LINE na podstawie danych aCGH oraz pochodzących z sekwencjonowania następnej generacji
- opracowanie metod identyfikacji rearanżacji mediowanych przez pary LINEów nie powodujących zmian w ilości materiału genetycznego
- stworzenie narzędzia informatycznego pozwalającego na rekonstrukcję historii ewolucji molekularnej LINEów
- wykonanie filogenomicznej analizy porównawczej genomów ssaków naczelnych oraz sformułowanie hipotez dotyczących wpływu rekombinacji historii ewolucji

Współpraca naukowa z zespołem kierowanym przez profesora Stankiewicza pozwoli na opracowanie metod identyfikacji rearanżacji genomu na podstawie analiz danych z aCGH i sekwencjonowania następnej generacji oraz jego wiedza biologiczna umoliwi na weryfikację hipotez wynikających z analizy porównawczej genomów ssaków naczelnych. Praca pod kierunkiem profesor Anny Gambin pozwoli na zaprojektowaniu efektywnych algorytmów identyfikacji rearanżacji genomów oraz matematycznych modeli ewolucji LINEów.

Chwilowo nie są dostępne żadne zewnętrzne środki z grantów zapewniające finansowanie projektu.