

Gwałtowny postęp technologii sekwencjonowania genomów ludzkich pozwolił na analizę porównawczą sekwencji molekularnych pochodzących od różnych dawców. Odsłoniła ona drastyczną różnorodność genetyczną w populacji ludzkiej.

Badania prowadzone we współpracy z Baylor College of Medicine przyczyniły się do zmiany postrzegania genomu jako stabilnie zakodowanej informacji. Identyfikowane w genomach zmiany strukturalne obejmują delecje, duplikacje, inwersje oraz zbalansowane translokacje. Molekularne mechanizmy powstawania tych rearanżacji nie są jeszcze do końca poznane, a badacze poszukują przyczyn, dla których pewne miejsca na chromosomach są wysoce niestabilne. Nasze dotychczasowe badania pozwoliły wykazać, że rearanżacje genomowe mogą być mediowane przez powtarzalne fragmenty genomu. Początkowo wykazano, że taką rolę pełnią segmentalne duplikacje, a następnie, także o wiele krótsze, ruchome elementy genetyczne, czyli transpozony np. z rodziny HERV, czy LINE.

Warianty strukturalne genomu są niezwykle ważne jako narzędzie w analizach ewolucji molekularnej, ale przede wszystkim ze względu na ich potencjalnie patogenny charakter. Zidentyfikowano wiele zespołów chorobowych, których podłożem jest np delecja pewnego fragmentu DNA (badane przez nas w poprzednich projektach to m.in.: autyzm, epilepsja, zaburzenia rozwojowe).

Dzięki współpracy genetyków i informatyków opracowano w ostatnich latach niezwykle efektywne metody identyfikacji zmian strukturalnych zarówno na podstawie danych sekwencyjnych jak też pochodzących z mikromacierzy hybrydyzacji porównawczej. Jednak dużo większym wyzwaniem pozostaje interpretacja wykrytych zmian. Stosunkowo rzadko zdarza się, że rearanżacja narusza określony gen powiązany z daną chorobą. Najczęściej interpretacja zmiany genomu jest bardzo trudna i do jej adekwatnej analizy niezbędna okazuje się współpraca interdyscyplinarnych zespołów składających się z biologów, lekarzy, matematyków i informatyków.

W niniejszym projekcie koncentrujemy się na interpretacji zmian strukturalnych genomu, która integruje informacje pochodzące z różnych eksperymentów. Połączenie wiedzy o wrażliwości na dawkę poszczególnych genów, informacji o położeniu elementów regulatorowych oraz strukturze chromatyny w danym rejonie genomu, pozwoli na zbudowanie modelu umożliwiającego predykcję patogenności danej rearanżacji.

W ścisłej współpracy z genetykami zamierzamy stworzyć narzędzie obliczeniowe wspierające lekarzy w diagnostyce zespołów genetycznych. Zaproponowany przez nas model probabilistyczny pozwoli na wyłonienie najbardziej istotnych cech, które odpowiadają za fenotyp. Opracowane algorytmy zostaną przetestowane na danych o rearanżacjach genomowych kilkudziesięciu tysięcy pacjentów zgromadzonych w Baylor College of Medicine. Dodatkowo planujemy sekwencjonowanie techniką długich odczytów dla wybranej grupy pacjentów, co mamy nadzieję, że pozwoli na wyjaśnienie występowania bardzo zróżnicowanych objawów przy tym samym wariacie genetycznym.