

Mikołaj Danielewski

Aspekty metodyczne zastosowania metylomu i transkryptomu w badaniu

chorób cywilizacyjnych

Cele pracy: Celem rozprawy doktorskiej jest analiza wybranych aspektów metodycznych zastosowania metylomu i transkryptomu w badaniu chorób cywilizacyjnych. W przypadku proliferacji metylacji postawiono hipotezę, że ich wybrane części wyróżnia wysoka stabilność w czasie, a w odniesieniu do profilów ekspresji badano hipotezę o niezależności ekspresji genów w analizie korelacyjnej. Celem pracy przeglądowej było zebranie zestawu baz danych, które można by stosować w przyszłych badaniach prowadzonych *in silico* lub w połączeniu z danymi uzyskanymi lokalnie.

Materiały i metody: Przeprowadzono badania *in silico* przy wykorzystaniu zestawów danych udostępnionych w biologicznych bazach danych. Pierwsza praca obejmowała trzy publicznie dostępne zestawy danych metylacyjnych otrzymanych przy użyciu mikromacierzy metylacyjnych Illumina 450k lub EPIC w trzech grupach: dzieci przed i po okresie dojrzewania, dorosłych z otyłością, oraz osób starszych. W każdej z grup metylom został oznaczony dwukrotnie w istotnym odstępie czasu. Do analizy danych użyte zostały narzędzia bioinformatyczne, w tym *minfi*, *sva::ComBat*, *GapHunter*, jak również skrypty własne. Druga praca oryginalna wykorzystała duży i również publicznie dostępny zestaw profilów ekspresji genów otrzymanych w projekcie GTEx. Wykorzystano w tej pracy także zewnętrzne dane do walidacji otrzymanych wyników: cztery zbiory danych dla krwi pełnej (pacjenci z migrenowymi bólami głowy, infekcją SARS-CoV-2 i nieswoistymi zapaleniami jelit) oraz dwa dla błony śluzowej końcowego odcinka jelita krętego (pacjenci z nieswoistymi zapaleniami jelit). Do analizy danych w drugiej pracy zastosowano korelację *bicor* (*biweight midcorrelation*, z pakietu *WGCNA*). Przeprowadzono również analizę literaturową oraz dostępności różnorodnych biologicznych baz danych przydatnych podczas analiz metylomicznych czy transkryptomicznych, co jest przedmiotem pracy przeglądowej.

Wyniki: W pierwszej pracy zademonstrowano istnienie miejsc w metylomie, których poziom metylacji pozostaje niezmienny wraz z upływem czasu niezależnie od badanej grupy wiekowej czy odstępu czasowego. Łącznie znaleziono 239 wysoce stabilnych sond, z czego 172 nie było powiązanych z wpływem czynnika genetycznego. Były one położone m.in. w genach *ENTPD2*, *TESSP1*, *WDR27*, *XAGE3*. W drugiej pracy oryginalnej wykryto istnienie genów których ekspresja jest w największym stopniu niezależna od ekspresji innych genów. Do genów o niezależnej ekspresji we krwi, potwierdzonych w danych zewnętrznych, zaliczono m.in.: *HIST1H2AD*, *TMEM176B*, *GPX1*, *PI3*, *TMEM176A*, *HIST1H2BD*. W publikacji przeglądowej opisano łącznie 50 baz danych z różnorodnych dziedzin, wiele z których może być potencjalnie przydatna w badaniach chorób cywilizacyjnych i/lub metylomów oraz transkryptomów.

Wnioski: W profilu metylacji DNA krwi pełnej człowieka znajdują się nieliczne miejsca wysoce stabilne. Sondy, których poziom metylacji zależy od genotypu, są z reguły bardziej niezmiennie. W transkryptomie tkanek człowieka obecne są geny, których ekspresja nie koreluje z ekspresją innych genów, co może sugerować jej stochastyczność lub niezmienną, jak również istnienie indywidualnych mechanizmów regulacyjnych i wskazywać na ważną rolę zmiennych proporcji w mieszaninach komórkowych. Wyniki wskazują na cechy biologiczne będące źródłem istotnych wyzwań metodycznych w badaniu chorób cywilizacyjnych z wykorzystaniem metylomu i transkryptomu. Dotyczy to w szczególności odkrywania biomarkerów. Sporządzona lista baz danych wskazuje ogrom publicznie dostępnych informacji biologicznych, a w konsekwencji podkreśla potencjał ich wykorzystania jako źródeł w badaniach naukowych.

Danielewski Mikołaj