

Rozwój baz danych MODOMICS i NACDDB oraz ich zastosowanie do badania wpływu modyfikacji RNA na jego strukturę

Streszczenie

Modyfikacje posttranskrypcyjne kształtują strukturę i funkcję RNA we wszystkich organizmach. Te zmiany chemiczne wpływają na tworzenie się struktury, oddziaływania molekularne i stabilność, wpływając tym samym na wydajność i dokładność biosyntezy białek. Jednak ich dystrybucja, zachowawczość w ewolucji i wpływ strukturalny są niekompletnie scharakteryzowane. Oprócz modyfikacji naturalnych, do kwasów nukleinowych wprowadzono szeroką gamę sztucznych modyfikacji chemicznych, które wykorzystuje się m.in. w badaniach biofizycznych, w chemicznym sondowaniu struktur i w zastosowaniach terapeutycznych. Główna część pracy badawczej przedstawionej w tej rozprawie obejmuje rozwój baz danych, które gromadzą i organizują różne typy danych związanych z modyfikacjami kwasów nukleinowych. Po pierwsze, przeprowadzono znaczną rozbudowę bazy danych MODOMICS, skupiając się na modyfikacjach RNA i ich kontekście strukturalnym. Po drugie, opracowano nową bazę danych, NACDDB, do gromadzenia i analizy widm dichroizmu kołowego (CD) kwasów nukleinowych, obejmujących zarówno cząsteczki występujące naturalnie, jak i syntetycznie zmodyfikowane. Oprócz rozwoju tych narzędzi bioinformatycznych, dane w nich zawarte zostały przeanalizowane w celu zbadania, w jaki sposób modyfikacje chemiczne wywierają zależne od położenia i kontekstu efekty na strukturę RNA. Przeprowadzono analizę porównawczą przyrównań sekwencji RNA i struktur trójwymiarowych w MODOMICS w celu zidentyfikowania sygnatur ewolucyjnych i preferencji strukturalnych modyfikowanych nukleotydów. Analiza rybosomalnego rRNA wykazała wzory zachowania kluczowych modyfikacji w procesie ewolucji dywergentnej sekwencji rRNA, szczególnie w obrębie centrum peptydylotransferazy. Konteksty strukturalne urydyny i jej zmodyfikowanych pochodnych pseudourydyny (Ψ) i dihydrourydyny (D) zostały zbadane pod kątem oddziaływania warstwowego z rybonukleotydami sąsiadującymi w sekwencji RNA. Analiza wykazała, że reszty dihydrourydyny zazwyczaj występują w kontekście pozbawionym oddziaływań warstwowych, podczas gdy reszty pseudourydyny częściej niż niemodyfikowane reszty urydyny biorą udział w tworzeniu tych oddziaływań, szczególnie z rybonukleotydami sąsiadującymi z nimi od strony 5'. Wyniki te wskazują na związek między zmienionymi preferencjami Ψ i D względem oddziaływań warstwowych a ich preferencjami do znajdowania się w takim kontekście strukturalnym, w którym obecność lub brak oddziaływań warstwowych jest ważna dla poprawnego utworzenia struktury RNA. Kolejnym elementem prac była analiza danych dichroizmu kołowego z NACDDB pod kątem oceny wpływu strukturalnego modyfikacji m¹A i m⁶A. Wyniki wskazują, że metylowane adenozyne zmieniają lokalną stabilność konformacyjną oraz wywołują destabilizację w podwyższonych temperaturach. Przyszłe badania integrujące różne dane doświadczalne dotyczące pozycji zmodyfikowanych reszt w sekwencjach RNA, wraz z danymi strukturalnymi dla RNA zawierających takie modyfikacje, będą niezbędne do pełnego zrozumienia, w jaki sposób modyfikacje RNA wpływają na strukturę RNA. W tym celu zainicjowano współpracę między zespołem MODOMICS a twórcami bazy danych Sci-ModoM, której celem jest wyodrębnienie zmodyfikowanych sekwencji transkryptów z zestawów danych sekwencjonowania o wysokiej przepustowości i ustalenie ich preferencji strukturalnych.