

W ostatnich latach odkryto wiele funkcjonalnych RNA, w tym rybozymy katalityczne, ryboprzełączniki wiążące ligandy oraz długie niekodujące RNA, które odgrywają kluczowe role w różnych aspektach regulacji ekspresji genów. Postępy w badaniach nad RNA zmieniają współczesną biologię i medycynę, czego przykładem jest opracowanie szczepionek mRNA i systemów edycji genomu prowadzonych przez RNA. Funkcjonalne RNA często przyjmują złożone struktury definiujące ich funkcję, np. poprzez wiązanie z białkami lub małymi cząsteczkami. Podobnie jak RNA, DNA może tworzyć różne struktury niekanoniczne, które są zaangażowane w utrzymanie stabilności genetycznej i powiązane z chorobami człowieka, takimi jak dystrofia mięśniowa czy choroba Alzheimer'a.

Trójwymiarowa struktura kwasów nukleinowych jest zorganizowana hierarchicznie i może być postrzegana jako zestaw bloków konstrukcyjnych, tzw. modułów 3D. Powtarzalne moduły, które zachowują swoje cechy strukturalne w różnych cząsteczkach, nazywane są motywami strukturalnymi i często pełnią rolę jednostek funkcjonalnych. Na przykład rybosomalne RNA wykorzystują motywy typu A-minor (płaskie trójki zasad) do rozpoznawania odpowiednich cząsteczek transportującego RNA. Centra katalityczne wielu rybozymów służą jako miejsca wiązania jonów magnezu niezbędnych do ich aktywacji. Telomerowe DNA o strukturze G-kwadrupleksów (płaskie tetrady guanozynowe) odgrywają ważną rolę w procesach nowotworowych, przez co stanowią przedmiot aktywnych badań nad nowymi lekami. Identyfikacja i charakterystyka tych powtarzających się motywów dostarczają istotnych informacji o mechanizmach leżących u podstaw zwijania i funkcjonowania ustrukturyzowanych kwasów nukleinowych.

Nasza wiedza o strukturalnych motywach RNA i DNA pozostaje jednak niekompletna. **Nie istnieje zunifikowane źródło ani narzędzie informatyczne, które zapewniłoby kompleksowy przegląd nawet najczęściej występujących i najlepiej opisanych motywów.** W efekcie, przykładowo motywy typu A-minor i kink-turn, oba odkryte w 2001 roku, wciąż nie mają swoich artykułów na Wikipedii. **W konsekwencji kluczowe odkrycia są czasami niedoceniane lub niespójnie cytowane, co powoduje, że istotne fakty mogą pozostać przez naukowców niezauważone.** Na przykład badacze, którzy w 2015 roku wykazali, że zaburzenie motywu kink-turn w spliceosomalnym RNA prowadzi do zaburzeń neurorozwojowych, nie użyli terminu „kink-turn”. Co więcej, mimo że wiadomo, iż wiele motywów formuje się tylko w obecności określonych jonów lub ligandów, informacje te nie są konsekwentnie stosowane do oceny niejednoznacznych miejsc wiązania małych cząsteczek w eksperymentalnych strukturach 3D. **Ta luka prawdopodobnie wynika z ograniczeń istniejących metod obliczeniowych, które opierają się wyłącznie na adnotacjach oddziaływań par zasad.** Niedawno opracowaliśmy nowy algorytm ARTEM, pierwszy umożliwiający identyfikację motywów strukturalnych kwasów nukleinowych oraz miejsc wiązania o dowolnym rozmiarze, sieci oddziaływań i kolejności nukleotydów.

Głównym celem proponowanego projektu jest opracowanie ARTEMANN, narzędzia bioinformatycznego do identyfikacji powtarzających się motywów strukturalnych w doświadczalnie wyznaczonych trójwymiarowych strukturach kwasów nukleinowych, oraz jego zastosowanie do adnotacji motywów i miejsc wiązania jonów metali, ligandów i białek w znanych strukturach, co pozwoli znaleźć dotychczas ukryte podobieństwa strukturalne i funkcjonalne oraz ich potencjalne implikacje ewolucyjne. Dzięki nowym danym stworzymy ulepszoną klasyfikację trójwymiarowych struktur kwasów nukleinowych i ich modułów, wzbogacając opis ich charakterystycznych cech i wariacji.

Nowa metoda będzie oparta na naszym algorytmie superpozycji ARTEM, który potrafi identyfikować podobieństwa między dowolnymi strukturami kwasów nukleinowych bez uprzedniej wiedzy o dopasowaniu nukleotyd-nukleotyd. ARTEMANN zostanie wyposażony w konfigurację parametrów specyficznych dla motywów, dostosowane do bogatego zestawu obecnie znanych powtarzających się modułów 3D i miejsc wiązania, które zostaną zebrane w ręcznie opracowanym katalogu.

Oczekiwane wyniki tego projektu będą dwojakie. Opracowany katalog motywów, wraz z rozwiniętą metodologią i jej implementacją, mają potencjał zainteresować tysiące badaczy zajmujących się biologią kwasów nukleinowych, ponieważ zapewni łatwo dostępne adnotacje, których obecnie brakuje w badaniach strukturalnych RNA i DNA. Ponadto kompleksowa analiza ustrukturyzowanych RNA i DNA z wykorzystaniem proponowanej metody wyjaśni nowe cechy strukturalne i funkcjonalne tych cząsteczek, co istotnie powiększy naszą wiedzę na temat ich roli w kluczowych procesach komórkowych.