

Mgr inż. Monika Kwiatkowska

Genomiczna charakterystyka długich niekodujących RNA w genomie danio pręgowanego

Genomy kręgowców kodują tysiące długich niekodujących RNA (lncRNA) — transkryptów o długości przekraczającej 200 nukleotydów, które mają ograniczony potencjał kodowania białek. Pomimo rosnącej liczby lncRNA zaangażowanych w kluczowe procesy biologiczne, ponad 97% z nich wciąż nie ma przypisanej funkcji. Modele zwierzęce mogą odegrać istotną rolę w poznaniu biologicznych funkcji lncRNA, jednak powodzenie tych badań w dużej mierze zależy od jakości adnotacji genomu. Chociaż danio pręgowane (*Danio rerio*) stał się obiecującym modelem kręgowców do badania biologii lncRNA, jakość adnotacji jego genomu wciąż pozostaje znacznie gorsza w porównaniu z adnotacjami genów u ludzi i myszy, co istotnie ogranicza jego zastosowanie. Zadaniem tego projektu było opracowanie kompleksowego i precyzyjnego katalogu genów lncRNA w genomie danio pręgowanego. W tym celu zoptymalizowano protokół CapTrap-CLS. Metoda CapTrap-seq, która służy do przygotowywania bibliotek pełnej długości, wykazuje przewagę w specyficznym wzbogacaniu o transkrypty z kompletnymi końcami 5' i 3', jednocześnie skutecznie redukując ilość rybosomalnego RNA (rRNA). Zastosowanie CapTrap-seq u danio pręgowanego znacząco poprawiło adnotację miejsc rozpoczęcia transkrypcji (TSS) dla biologicznie istotnych genów. Ponadto, aby zwiększyć detekcję słabo reprezentowanych lncRNA, CapTrap-seq połączono z metodą Capture Long-read Sequencing (CLS) — ukierunkowaną techniką sekwencjonowania RNA, która łączy wychwytywanie RNA z sekwencjonowaniem długich odczytów za pomocą platformy ONT (Oxford Nanopore Technologies). To podejście skutecznie wzbogaciło lncRNA, znacznie zwiększając ich reprezentację w bibliotekach sekwencyjnych, co w efekcie doprowadziło do rozszerzenia referencyjnej adnotacji lncRNA dla genomu danio pręgowanego. Dodatkowo, ukierunkowane sekwencjonowanie RNA z regionów pozycyjnie zachowanych pomiędzy człowiekiem, myszą a danio pręgowanym ujawniło nowe izoformy transkryptów dla potencjalnie funkcjonalnych lncRNA oraz istotnie poprawiło detekcję nowych genów w przestrzeniach międzygenowych. Wynik ten podkreśla znaczenie pozycyjnej zachowawczości jako efektywnej strategii w odkrywaniu nowych, potencjalnie funkcjonalnych cząsteczek lncRNA. Zoptymalizowana adnotacja genomu danio pręgowanego, stworzona w ramach tego projektu doktorskiego, stanowi solidną podstawę do wzmocnienia biologicznej wartości tego gatunku jako modelu zwierzęcego do badań nad funkcją lncRNA.