

Ćwiczenia 4 – Integracja i analiza danych z różnych eksperymentów

Ćw1. Weryfikacja metylacji rejonów wysp CpG potwierdzonych dwoma różnymi eksperymentami

1. Otwórz Galaxy workbench <http://courses.combio.pl/>
2. Utwórz nową historię
3. Z menu narzędzi wybierz zakładkę Get Data a następnie UCSC Main table browser
4. Wybierz dane dla genomu człowieka **hg19**:
 - a. Group: Regulation track: HAIB Methyl RRBS
 - b. table: HeLa-S3 1
 - c. region: zaznacz position chr21: 33000000-34000000
 - d. output format: BED
5. Kliknij **Get output**
6. Na następnej stronie potwierdź wysłanie pliku do Galaxy
7. Powtórz powyższe czynności aby ściągnąć tabelę:
 - a. Group: Regulation Track: HAIB Methyl450
 - b. table: HeLa S3
 - c. region: zaznacz position chr21: 33000000-34000000
 - d. output format: BED
8. Powtórz powyższe czynności aby ściągnąć tabelę:
 - a. Group: Regulation Track: CpG Islands
 - b. table: cpGIslandsExt
 - c. region: zaznacz position chr21:33000000-34000000
 - d. output format: BED

9. Zmień nazwy tabeli na krótsze:
 - a. dla tabeli z pkt.4 podaj RRBS
 - b. dla tabeli z pkt.7 podaj M450
 - c. dla tabeli z pkt.8 podaj CpG
10. Wybierz narzędzie **bedtools Multiple Intersect**
11. Wprowadź trzy pobrane z UCSC tabele, wybierz opcję **Report empty regions: No**
12. Uruchom narzędzie
13. Z menu narzędzi wybierz zakładkę **Filter and Sort** i wybierz narzędzie **Filter**
14. Jako plik do filtrowania wybierz plik utworzony w poprzednim punkcie
15. Aby pozostawić jedynie regiony zawierające dane z wszystkich trzech tabel jako warunek filtrowania wpisz **c4>2** (wartość w kolumnie 4 ma być większa od 2)
16. Uruchom narzędzie
17. Pytanie: Ile metylowanych rejonów wykrytych dwoma metodami i nachodzących na wyspy CpG udało się zidentyfikować?
18. Wykonaj analogiczne porównanie jak w pkt. 10-16, tym razem porównując wyłącznie dwie tabele z miejscami metylacji (z pkt. 4 i z pkt 7). Pamiętaj aby podczas filtrowania dostosować wpisywany warunek tak, aby wyświetlić tylko te rejony które zawierają dane z obu tabel.
19. Pytanie: Ile metylowanych rejonów wykrytych dwoma niezależnymi metodami udało ci się znaleźć?
20. Porównując wynik uzyskany w pkt. 20 z wynikiem uzyskanym w pkt. 17 wywnioskuj ile jest rejonów potwierdzonych dwoma metodami, które nie nachodzą na wyspy CpG?
21. Zapisz plik z pkt.18 będący odfiltrowanym wynikiem analizy w komórkach HeLa na dysku

Ćw.2 Porównanie potwierdzonych profili metylacji dla dwóch linii komórkowych

1. Powtórz ćw.1 tym razem w pozycji „table” wybierając dane dla hepatocytów. Możesz pominąć ściąganie pozycji wysp CpG oraz porównanie trzech tabeli i przejdź od razu do porównanie dwóch tabeli zawierających dane o metylacji.
2. Zapisz plik będący odfiltrowanym wynikiem analizy hepatocytów na dysku
3. Otwórz zapisany plik dla hepatocytów oraz komórek HeLa w Excelu
4. Przeformatuj pliki do formatu BED (pamiętaj o linii definiującej plik BED - <https://genome.ucsc.edu/FAQ/FAQformat.html#format1.5>) i zapisz zmiany
5. Załaduj utworzone pliki BED do UCSC Genome Browser
6. Porównaj wyniki analizy dla komórek HeLa i hepatocytów
7. Pytanie: Czy istnieją różnice pomiędzy profilami metylacji w hepatocytach i komórkach HeLa?

Ćw.3 Weryfikacja regionów promotorowych

1. Utwórz nową historię
2. Przeprowadź podobną analizę jak w ćwiczeniu 1 oraz 2, jednak ukierunkowana na identyfikację miejsc aktywnych transkrypcyjnie. Do analizy użyj następujących danych z UCSC Genome Browser:
 - a. track: HAIB TFBS (transcription factor binding sites)
 - i. table: HeLa Pol II (na końcu nazwy PkRep1);
 - ii. table: HepG Pol2 (na końcu nazwy PkRep1) (zawierają rejony genomowe na których obserwowana jest aktywna polimeraza RNA II dla komórek HeLa oraz linii komórkowej raka wątroby)
 - b. track: DUKE DnaseI HS
 - i. table: HeLaS3 Pk;
 - ii. table: HepG2 Pk; (zawierają rejony genomowe wrażliwe na trawienie Dnazą I, są to rejony tzw. „otwartej”, aktywnej transkrypcyjnie chromatyny)
3. Najpierw porównaj miejsca występowania pol II z rejonami otwartej chromatyny w ramach linii komórkowej HeLa oraz osobno w ramach linii HepG (analogicznie do ćw1)
4. Następnie porównaj wzorce uzyskane dla poszczególnych linii komórkowych w pkt.3 pomiędzy komórkami HeLa i nowotworem wątroby (analogicznie do ćw.2)
5. Wklej do protokołu koordynaty przykładowego rejonu w którym widoczne są różnice w aktywności transkrypcyjnej chromatyny pomiędzy komórkami.