
Zmienność strukturalna genomów

Opracowanie: Aleksandra Świercz

Zajęcia 1. Zapoznanie się z formatem plików VCF

1. W katalogu `/home/Lab1/` znajdują się 2 pliki `.vcf`. Zawierają one wyznaczone warianty strukturalne krótkie (SNP) dla dwóch próbek rakowych genomu ludzkiego traktowanych różnym stężeniem leku Paclitaxel (PTX): 16PTX i 32PTX. Obie próbki były porównane do próbki kontrolnej (nie traktowanej lekiem) o nazwie A_2780. Mapowanie i dalsza analiza była przeprowadzona na wersji genomu ludzkiego grch37.
2. Połącz ze sobą oba pliki za pomocą narzędzia `vcftools`: `vcf-merge`.
 - Pliki muszą być wcześniej poindexowane i spakowane – postępuj zgodnie z podpowiedzią programu.
 - Ile powstało wspólnych wariantów z obu plików? Co `vcf-tools` zrobił ze wspólnymi kolumnami, np. FILTER?
 - Porównaj warianty powstałe na chromosomie pierwszym w pozycjach: 12073251 i 15993699. Oba miejsca są multialleliczne (występuje więcej niż jeden alternatywny wariant). Czy w taki sam sposób został zapisany wariant dla próbki 16PTX? Zauważ zmienioną kolejność w kolumnie FORMAT i następnych. Czy wpłynęło to na zmianę kolejności pól AF i AD w kolejnych kolumnach?
3. Wykonaj połączenie plików za pomocą drugiego narzędzia: `bcftools merge`. W razie błędu z łączeniem plików, zajrzyj tu jak rozwiązać problem: (AS_FilterStatus należy zmienić na 1, albo każdą linię multiallelic pozmienić)
<https://github.com/broadinstitute/gatk/issues/6857>
4. Zrób adnotację wariantów z połączonego pliku (dowolnym narzędziem) w programie VEP (Ensembl Variant Effect Predictor). Pamiętaj o wybraniu odpowiedniej wersji genomu

ludzkiego! Wyniki analizy zamieść w sprawozdaniu (wykresy które generuje VEP). Zamieść link do wyników analizy w sprawozdaniu (pamiętaj że jest on ważny ok miesiąca, więc nie czekaj z realizacją całego sprawozdania).

5. Wszystkie pliki z ćwiczenia proszę zapisywać w swoim katalogu domowym. W sprawozdaniu zamieść wszystkie polecenia. Skomentuj wyniki. Które narzędzie wydaje się być bardziej przydatne w celu łączenia VCFów?

Linki do manuali narzędzi:

<http://samtools.github.io/bcftools/bcftools.html>

https://vcftools.github.io/man_latest.html

http://vcftools.sourceforge.net/perl_module.html#vcf-merge

Sprawozdanie wyślij na adres: Aleksandra.Swiercz@cs.put.poznan.pl