**1. Wyjaśnij czy sensownym biologicznie jest aby próbki po normalizacji miały zbliżone średnie?**

Normalizacja jest stosowana aby wziąć pod uwagę różną ilość materiału oraz różną ekstrakcje RNA, odwrotnej transkrypcji, znakowania czy fotodetekcji. Lokalizacja i skala dla różnych mikro macierzy powinna być na ogół taka sama. Lokalizacja i skala box plotu nie definiuje poziomu ekspresji genów, więc jest sensowne aby miały zbliżone średnie.

**2. Zadanie2 Przypadek do analizy z NCBI GEO DataSets https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS810**

* Czego dotyczyło badanie?

Badanie dotyczyło ekspresji genów u 22 pacjentów w różnym stadium Alzheimera.

* Gdzie badano ekspresje genów? (jaki materiał/tkanka)

Mózg - hipokamp

* Ile było wszystkich próbek i na ile grup były one podzielone (Experiment design and value distribution oraz Sample Subsets )

Było 31 próbek podzielonych na 3 grupy – contol, incipient AD, moderate AD, sever AD

* Według jakich kryteriów podzielono próbki na grupy? Opisz wykorzystane parametry. (Experiment design and value distribution oraz Sample Subsets )

Podzielono ze względu na zaawansowanie choroby oraz grupę kontrolną na podstawie wyników MiniMental Status Examination (MMSE) i neurofibrillary tangle (NFT). Oraz formalne kryterium – braak (determinowanie zaawansowania alzhaimera na podstawie autopsji.

* Ile jest genów różnicujących? (Find genes, Find genes that are up/down for this condition(s))

931

* Jak wygląda przebieg ekspresji dla genów SPARC, VSNL1 oraz COL5A2 w kolejnych grupach? (Expression Profiles lub Find genes) Za co odpowiadają te geny? Czy obserwujesz jakieś tendencje zmiany poziomu ich ekspresji w kolejnych grupach? Poszukaj w źródłach zewnętrznych informacji na temat ich związku z chorobą Alzheimera.

**Sparc** – nieznacznie rośnie dla incipient i moderate ze znacznym wzrostem dla severe.

Ten gen koduje bogate w cysteinę białko macierzy pozakomórkowej, które wspomaga kalcyfikację kolagenu w kościach, a także bierze udział w syntezie macierzy pozakomórkowej i zmianach kształtu komórek. Produkt genu związany jest zarówno z supresją nowotworową, jak i z przerzutami – przez wpływ na kształt komórek, co może wspierać inwazję komórek nowotworowych. Zidentyfikowano trzy warianty transkryptów kodujące różne izoformy tego białka.

SPARC jest silnie ekspresjonowany w mózgu w chorobie Alzheimera, zlokalizowanym ze złogami białka Aβ i bierze udział w zapaleniu mózgu oraz naprawie tkanek. Hevin może być obniżony w stanie chorobowym. Potrzebne są dalsze badania nad rolami białek SPARC i Hevin oraz ich szlakami sygnałowymi w neurozapaleniu związanym z Alzheimerem. Normalizacja ekspresji SPARC/Hevin, np. przez hamowanie nadekspresji SPARC, może stanowić nową strategię terapeutyczną w modulacji przebiegu choroby Alzheimera.

VSNL1 – w grupie incipiet i moderate nieznacznie spadła w porównaniu do grupy kontrolnej; znaczny spadek w severe.

Ten gen należy do podrodziny białek neuronalnych czujników wapnia visinin/recoverin. Kodowane białko jest silnie wyrażane w komórkach ziarnistych móżdżku, gdzie wiąże się z błonami w sposób zależny od wapnia i moduluje szlaki sygnalizacyjne ośrodkowego układu nerwowego, regulując aktywność cyklazy adenylanowej. Zaobserwowano warianty alternatywnego splicingu, ale ich pełna długość nie została określona.

VSNL1 wykazuje istotną współekspresję z genami w szlakach sygnalizacji wapniowej, choroby Alzheimera, długotrwałego wzmocnienia i osłabienia synaptycznego oraz transportu receptorów AMPA. Związek z chorobą Alzheimera częściowo wynika z korelacji z ekspresją APP (amyloidowego białka prekursorowego). Wyniki te wskazują na związek VSNL1 z mechanizmami molekularnymi choroby, w tym ze szlakami zaangażowanymi w patologię synaptyczną. Potrzebne są dalsze badania, aby ustalić, czy APP zwiększa ekspresję VSNL1, VSNL1 zwiększa ekspresję APP, czy oba są regulowane przez wspólne czynniki patogenne.

**COL5A2**- dosyć liniowy spadek w stosnku do stopnia zaawansowania chorby, z dwoma outlierami w severe.

Ten gen koduje łańcuch alfa jednego z rzadkich kolagenów włóknistych. Kolagen typu V, występujący w tkankach zawierających kolagen typu I, reguluje formowanie włókien złożonych z obu typów. Produkt genu jest blisko spokrewniony z kolagenem typu XI, a łańcuchy kolagenowe typów V i XI mogą tworzyć pojedynczy typ kolagenu o specyficznych dla tkanki kombinacjach. Mutacje tego genu są związane z zespołem Ehlersa-Danlosa, typami I i II.

Analiza genów kandydujących do choroby Alzheimera sugeruje, że patogeneza tej choroby może być związana z zaburzeniami transdukcji sygnałów (AGTR1 i PTAFR), obniżeniem zdolności transportu białek (**COL5A2**, COL4A1), upośledzeniem naprawy aksonów (CNR1) oraz zaburzeniami homeostazy wapnia wewnątrzkomórkowego (CACNB2, CACNA1E). Jednak ich potencjalna rola w patologii choroby Alzheimera wymaga dalszej weryfikacji w eksperymentach laboratoryjnych, ponieważ zostały one zidentyfikowane jedynie metodami obliczeniowymi przy użyciu ADA.

* Czym są housekeeping genes? Jaką pełnią rolę w eksperymencie mikromacierzowym? Wybierz trzy przykładowe geny tej kategorii i sprawdź ich ekspresje w kolejnych próbkach.

**Housekeeping genes** to geny niezbędne do podstawowych funkcji komórki, stale podlegające ekspresji we wszystkich typach komórek. W eksperymentach mikromacierzowych pełnią rolę **genów referencyjnych** do normalizacji wyników, co pozwala na dokładne porównanie ekspresji genów między próbkami.

Np. PGK1 lub GAPDH. Ich ekspresja jest równa na wszystkich etapach choroby.

**3. Wyznaczono poziom ekspresji genów dla 6 kolejnych chwil czasowych. Wyznacz współczynnik korelacji pomiędzy poziomem ekspresji genu 1 oraz genu 2. Zinterpretuj uzyskany wynik.**

-0.9730072055183202

Otrzymaliśmy ujemną korelację. Wartość -0.9 sugeruje, że gdy poziom ekspresji genu 1 wzrasta, poziom ekspresji genu 2 zwykle maleje, i odwrotnie. To wskazuje na wyraźny związek odwrotny między tymi dwoma genami. Wartość bliska -1 wskazuje, że związek jest bardzo silny. W praktyce oznacza to, że zmiany w ekspresji jednego genu są ściśle powiązane z przeciwnymi zmianami w ekspresji drugiego genu.

Może to sugerować, że jeden z genów może działać jako supresor drugiego w kontekście badanej funkcji biologicznej lub patologii lub te dwa geny regulują się nawzajem.

**4. Wejdź na STRING: functional protein association networks (string-db.org) a następnie przeanalizuj zestaw potencjalnych genów markerowych dla prognozy raka piersi. Wybierz Multiple proteins, a jako organizm Homo sapiens.**

* Jakie 3 procesy biologiczne mają najmniejszy FDR (false discovery rate) w rozważanej grupie genów?

[GO:0035239](http://amigo.geneontology.org/amigo/term/GO:0035239) Tube morphogenesis

[GO:1902893](http://amigo.geneontology.org/amigo/term/GO:1902893) Regulation of miRNA transcription

[GO:0050678](http://amigo.geneontology.org/amigo/term/GO:0050678) Regulation of epithelial cell proliferation

* Która funkcja molekularna ma najmniejszy FDR (false discovery rate) w rozważanej grupie genów?

[GO:0001228](http://amigo.geneontology.org/amigo/term/GO:0001228) DNA-binding transcription activator activity, RNA polymerase II-specifc

[GO:0003690](http://amigo.geneontology.org/amigo/term/GO:0003690) Double-stranded DNA binding

[GO:0000978](http://amigo.geneontology.org/amigo/term/GO:0000978) RNA polymerase II cis-regulatory region sequence-specific DNA binding

* Która ścieżka KEGG ma najmniejszy FDR (false discovery rate) w rozważanej grupie genów?

[hsa05206](https://www.kegg.jp/kegg-bin/show_pathway?hsa05206) MicroRNAs in cancer

[hsa04010](https://www.kegg.jp/kegg-bin/show_pathway?hsa04010) MAPK signaling pathway

[hsa05219](https://www.kegg.jp/kegg-bin/show_pathway?hsa05219) Bladder cancer

**5. Poniżej zamieszczono dane odnośnie ekspresji genów w komórkach drożdży podczas procesu oddychania – fermentacji alkoholowej.**

**Wyróżniamy dwa główne etapy tego procesu: - rozkład glukozy do kwasu pirogronowego, - przemianę kwasu pirogronowego do alkoholu.**

**Każdy z etapów kontrolowany jest przez 2 klasy genów odpowiedzialnych za te procesy. Dane pochodzą z 7 chwil czasowych (kolejne kolumny). Skomentuj i porównaj poniższe wyniki w kontekście powyższych informacji.**

**W jaki sposób podzieliłbyś rozważane chwile czasowe?**

Można wyraźnie zauważyć że ekspresja genów jest sprężona, w momencie końca ekspresji jednego zaczyna się drugi. Jest to związane z etapowością fermentacji alkoholowej, jedna grupa genów odpowiada ze rozkład glukozy do kwasu pirogronowego a druga za przemianę kwasu pirogronowego do alkoholu.

**6. Wejdź na stronę https://singlecell.broadinstitute.org/single\_cell/. Znajdują się tam dane z eksperymentów scRNA-seq.** Wyszukaj eksperyment: Single-cell transcriptomics of the spinal cord of a severe SMA mouse

* Czego dotyczył eksperyment? (Summary)

Spinal muscular atrophy (SMA) jest chorobą neuronów ruchowych spowodowaną utratą białka SMN (survival of motor neuron). Coraz więcej dowodów wskazuje jednak, że komórki nieneuronalne odgrywają istotną rolę w patogenezie SMA. Dotychczasowe badania nie pozwalały na zrozumienie zmian w transkryptomie na poziomie pojedynczych komórek rdzenia kręgowego w SMA, co ograniczało pełne poznanie roli konkretnych typów komórek. W tym badaniu przeprowadzono sekwencjonowanie RNA na poziomie pojedynczych komórek rdzenia kręgowego u myszy z ciężką postacią SMA, identyfikując dziesięć typów komórek oraz ich geny różnicowo eksprymowane.

* Ile komórek i genów rozważano? (Summary) - Ile typów komórek rozważano (Explore)

22155 komórek 18772 genów

* Sprawdź ekspresję genów: hemoglobiny (HBB) oraz SPARC wśród rozważanych typów komórek. W której grupie występuje największa ich ekspresja? (Explore, search genes)

HBB (Hbb-bs, Hbb-bt) najwyższa ekspresja (99-100%) występuje w grupie Blood; w innych szczątkowa

SPARC (Sparc, Sparcl1) ma bardziej zróżnicowaną ekspresje w grupach.

Sparc:

* Duża: Vasculature, Schawann\_cell, Microglia (100,100, 97 %)
* Średnia: Astrocyte, Ependymal, OPC, COP (98, 90, 87, 82%)
* Reszta 40-50%

Sparcl1:

* Duża: Astrocyte (100%)
* Średnia: Ependymal, Vasculature (93, 88%)
* Reszta 40-60%

A network of cells with different colored circles

Description automatically generated with medium confidence