

Tissue specificity index analysis - tau

Deisiany

2024-01-13

Load libraries

read in data

For Tissue specificity index analysis - tau - we will use the 133 genes which had average expression across all tissues equal or greater than 2, and was used in our PCA analysis.

```
# Load full genes
data <- read_csv("Tau_133_genes_with_expression_across_tissues_are_greater_than_2_SoyARCs.csv")

## Rows: 133 Columns: 20
## -- Column specification -----
## Delimiter: ","
## chr (5): Transcript ID, heatmap_label, Family, Class, Clade
## dbl (15): AM, OF, IAM, IBM, RootTip, Cotyledon, Hypocotyl, SAM6D, SAM17D, SA...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
data <- data %>% select(-c(Class, Clade, expr_median))
```

Calculate the maximal component value

$X^{(i)} = X_i / \max(X_i)$ for $1 \leq i \leq n$

```
# lets start by assigning our data to df
max_component_df <- data

for (Xi in 1:nrow(max_component_df)) {
  max_component_df[Xi, 4:17] <- max_component_df[Xi, 4:17] / max(max_component_df[Xi, 4:17], na.rm = TRUE)
}

max_component_df
```

```
## # A tibble: 133 x 17
##   `Transcript ID` heatmap_label Family    AM    OF    IAM    IBM RootTip
##   <chr>           <chr>      <chr>  <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Glyma.01G002100.1 Glyma.01G002100|G~ ARF    0.816 0.890 0.797 0.945 0.486
## 2 Glyma.01G019400.1 Glyma.01G019400|G~ IAA    0.382 0.718 0      0.140 0.349
## 3 Glyma.01G019400.2 Glyma.01G019400|G~ IAA    0.175 0      0.266 0.339 0.0983
## 4 Glyma.01G019400.3 Glyma.01G019400|G~ IAA    0.254 0.404 0.234 0.301 0.182
## 5 Glyma.01G098000.3 Glyma.01G098000|G~ IAA    0.396 0.298 0.311 0.559 0.190
## 6 Glyma.01G103500.1 Glyma.01G103500|G~ ARF    0.694 0.214 0.420 0.839 0.146
```

```
## 7 Glyma.02G000500.1 Glyma.02G000500|G~ IAA 0.252 0.315 0.137 0.114 0.271
## 8 Glyma.02G065300.1 Glyma.02G065300|G~ AFB/T~ 0.522 0.424 0.506 0.772 0.251
## 9 Glyma.02G142500.3 Glyma.02G142500|G~ IAA 0.0277 1 0.138 0.0538 0.0816
## 10 Glyma.02G142600.1 Glyma.02G142600|G~ IAA 0.0541 1 0.243 0.266 0.238
## # i 123 more rows
## # i 9 more variables: Cotyledon <dbl>, Hypocotyl <dbl>, SAM6D <dbl>,
## # SAM17D <dbl>, SAM38D <dbl>, Callus <dbl>, Leaf <dbl>, Root <dbl>,
## # Nodule <dbl>

# Calculate Tau (Tissue Specificity Index) for each gene
tau_df <- max_component_df

tau_df$tau <- apply(tau_df[, 4:17], 1, #one here indicates the function will be applied to each row
                    function(row) sum(1 - row) / 13)

tau_df$tau <- numeric(nrow(tau_df)) # Create an empty column

for (i in 1:nrow(tau_df)) {
  tau_df$tau[i] <- sum(1 - tau_df[i, 4:17]) / 13
}

(tau_df_values <- tau_df %>% select(`Transcript ID`, heatmap_label, tau))

## # A tibble: 133 x 3
##   `Transcript ID` heatmap_label tau
##   <chr>          <chr>          <dbl>
## 1 Glyma.01G002100.1 Glyma.01G002100|GmARF7/19_F.1 0.376
## 2 Glyma.01G019400.1 Glyma.01G019400|GmIAA8-9-B.1 0.615
## 3 Glyma.01G019400.2 Glyma.01G019400|GmIAA8-9-B.2 0.796
## 4 Glyma.01G019400.3 Glyma.01G019400|GmIAA8-9-B.3 0.800
## 5 Glyma.01G098000.3 Glyma.01G098000|GmIAA8-9-D.3 0.643
## 6 Glyma.01G103500.1 Glyma.01G103500|GmARF9_A.1 0.711
## 7 Glyma.02G000500.1 Glyma.02G000500|GmIAA1-4-D.1 0.831
## 8 Glyma.02G065300.1 Glyma.02G065300|GmAFB2/3_C.1 0.619
## 9 Glyma.02G142500.3 Glyma.02G142500|GmIAA16-F.3 0.944
## 10 Glyma.02G142600.1 Glyma.02G142600|GmIAA1-4-H.1 0.895
## # i 123 more rows
```