Search of mRNA targets with multiMiR

J. Andrés Castillo Rivas

Table of contents

1 multiMiR

1.1 Installation

```
#if (!require("BiocManager", quietly = TRUE))
#    install.packages("BiocManager")

#BiocManager::install("multiMiR")
#browseVignettes("multiMiR")

#install.packages("rentrez")
```

2 Documentation

```
library("multiMiR")
```

```
Welcome to multiMiR.

multiMiR database URL has been set to the
default value: http://multimir.org/

Database Version: 2.4.0  Updated: 2024-08-28
```

```
library("kableExtra")
library("dplyr")
```

```
Adjuntando el paquete: 'dplyr'

The following object is masked from 'package:kableExtra':

    group_rows

The following object is masked from 'package:multiMiR':

    select
```

```
The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
library("ggplot2") # Para visualización (opcional)
library("clusterProfiler") # Para análisis funcional (opcional)
```

```
clusterProfiler v4.12.6 Learn more at https://yulab-smu.top/contribution-knowledge-mining/

Please cite:

G Yu. Thirteen years of clusterProfiler. The Innovation. 2024,
5(6):100722

Adjuntando el paquete: 'clusterProfiler'

The following object is masked from 'package:multiMiR':

    select

The following object is masked from 'package:stats':

    filter
```

```r
library("org.Hs.eg.db") # Base de datos de genes humanos
```

```
Cargando paquete requerido: AnnotationDbi

Cargando paquete requerido: stats4

Cargando paquete requerido: BiocGenerics
```

```
Adjuntando el paquete: 'BiocGenerics'

The following objects are masked from 'package:dplyr':

    combine, intersect, setdiff, union

The following objects are masked from 'package:stats':

    IQR, mad, sd, var, xtabs

The following objects are masked from 'package:base':

    anyDuplicated, aperm, append, as.data.frame, basename, cbind,
    colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
    get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,
    match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,
    Position, rank, rbind, Reduce, rownames, sapply, setdiff, table,
    tapply, union, unique, unsplit, which.max, which.min

Cargando paquete requerido: Biobase

Welcome to Bioconductor

    Vignettes contain introductory material; view with
    'browseVignettes()'. To cite Bioconductor, see
    'citation("Biobase")', and for packages 'citation("pkgname")'.

Cargando paquete requerido: IRanges

Cargando paquete requerido: S4Vectors

Adjuntando el paquete: 'S4Vectors'

The following object is masked from 'package:clusterProfiler':

    rename

The following objects are masked from 'package:dplyr':

    first, rename
```

```
The following object is masked from 'package:utils':

    findMatches

The following objects are masked from 'package:base':

    expand.grid, I, unname

Adjuntando el paquete: 'IRanges'

The following object is masked from 'package:clusterProfiler':

    slice

The following objects are masked from 'package:dplyr':

    collapse, desc, slice

Adjuntando el paquete: 'AnnotationDbi'

The following object is masked from 'package:clusterProfiler':

    select

The following object is masked from 'package:dplyr':

    select
```

```
library("rentrez")
```

3 Library

The multiMiR database is now versioned. By default multiMiR will use the most recent version each time multiMiR is loaded. However it is now possible to switch between database versions and get information about the multiMiR database versions. `multimir_dbInfoVersions()` returns a dataframe with the available versions.

```
db.ver = multimir_dbInfoVersions()
db.ver
```

```
  VERSION    UPDATED                        RDA      DBNAME
1   2.4.0 2024-08-28 multimir_cutoffs_2.4.rda multimir2_4
2   2.3.0 2020-04-15 multimir_cutoffs_2.3.rda multimir2_3
3   2.2.0 2017-08-08 multimir_cutoffs_2.2.rda multimir2_2
4   2.1.0 2016-12-22 multimir_cutoffs_2.1.rda multimir2_1
5   2.0.0 2015-05-01     multimir_cutoffs.rda    multimir
                  SCHEMA PUBLIC                TABLES
1 multiMiR_DB_schema.sql      1 multiMiR_dbTables.txt
2 multiMiR_DB_schema.sql      1 multiMiR_dbTables.txt
3 multiMiR_DB_schema.sql      1 multiMiR_dbTables.txt
4 multiMiR_DB_schema.sql      1 multiMiR_dbTables.txt
5 multiMiR_DB_schema.sql      1 multiMiR_dbTables.txt
```

To switch between versions we can use `multimir_switchDBVersion()`.

```
multimir_switchDBVersion(db_version = "2.4.0")
```

```
Now using database version: 2.4.0
```

```
# curr_vers  <- vers_table[1, "VERSION"]  # current version
# multimir_switchDBVersion(db_version = curr_vers)
```

The remaining functions will query the selected version until the package is reloaded or until we switch to another version.

Information from each external database is stored in a table in the multiMiR database. To see a list of the tables, we can use the `multimir_dbTables()` function.

```
db.tables = multimir_dbTables()
db.tables
```

```
 [1] "diana_microt" "elmmo"        "map_counts"   "map_metadata" "microcosm"
 [6] "mir2disease"  "miranda"      "mirdb"        "mirecords"    "mirna"
[11] "mirtarbase"   "pharmaco_mir" "phenomir"     "pictar"       "pita"
[16] "tarbase"      "target"       "targetscan"
```

The function `multimir_dbInfo()` will display information about the external miRNA and miRNA-target databases in multiMiR, including version, release date, link to download the data, and the corresponding table in multiMiR.

```
db.info = multimir_dbInfo()
db.info <- as.data.frame(db.info)

print(db.info)
```

```
       map_name                    source_name source_version  source_date
1  diana_microt                  DIANA-microT              5   Sept, 2013
2         elmmo                         EIMMo              5    Jan, 2011
3     microcosm                     MicroCosm              5   Sept, 2009
4   mir2disease                   miR2Disease                Mar 14, 2011
5       miranda                       miRanda                   Aug, 2010
6         mirdb                         miRDB              6   June, 2019
7     mirecords                     miRecords              4 Apr 27, 2013
8    mirtarbase                    miRTarBase              9    Sept 2021
9  pharmaco_mir Pharmaco-miR (Verified Sets)
10     phenomir                      PhenomiR              2 Feb 15, 2011
11       pictar                        PicTar              2 Dec 21, 2012
12         pita                          PITA              6 Aug 31, 2008
13      tarbase                       TarBase              9         2023
14   targetscan                    TargetScan              8    Sept 2021
                                                                      source_url
1  http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=microT_CDS/index
2                        http://www.mirz.unibas.ch/miRNAtargetPredictionBulk.php
3      http://www.ebi.ac.uk/enright-srv/microcosm/cgi-bin/targets/v5/download.pl
4                                                  http://www.mir2disease.org
5                              http://www.microrna.org/microrna/getDownloads.do
6                                                        http://mirdb.org
7                                    http://mirecords.biolead.org/download.php
8                               http://mirtarbase.mbc.nctu.edu.tw/php/download.php
9                               http://www.pharmaco-mir.org/home/download_VERSE_db
10                                    http://mips.helmholtz-muenchen.de/phenomir/
11                                               http://dorina.mdc-berlin.de
12                         http://genie.weizmann.ac.il/pubs/mir07/mir07_data.html
13                                        https://dianalab.e-ce.uth.gr/tarbasev9
14        https://www.targetscan.org/cgi-bin/targetscan/data_download.vert80.cgi
```

```
#db.info |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

3.1 Predicted tables

```
predicted_tables()
```

```
[1] "diana_microt" "elmmo"        "microcosm"    "miranda"      "mirdb"
[6] "pictar"       "pita"         "targetscan"
```

3.2 Validated Tables

```
validated_tables()
```

```
[1] "mirecords"  "mirtarbase" "tarbase"
```

3.3 Diseasedrug Tables

```
diseasedrug_tables()
```

```
[1] "mir2disease"  "pharmaco_mir" "phenomir"
```

To see how many records are in these 14 external databases we refer to the `multimir_dbCount` function.

```
db.count = multimir_dbCount()
db.count
```

```
       map_name human_count mouse_count rat_count total_count
1  diana_microt     7664602     3747171         0    11411773
2         elmmo     3959112     1449133    547191     5955436
3     microcosm      762987      534735    353378     1651100
4   mir2disease        2875           0         0        2875
5       miranda     5429955     2379881    247368     8057204
```

```
6         mirdb     1990425     1091263     199250     3280938
7     mirecords        2425         449        171        3045
8    mirtarbase      957034      116689       1384     1075107
9  pharmaco_mir         308           5          0         313
10     phenomir       15138         491          0       15629
11       pictar      404066      302236          0      706302
12         pita     7710936     5163153          0    12874089
13      tarbase     1290272      473266       1031     1764713
14   targetscan    13964425    10387912          0    24352337
```

```
print(db.count)
```

```
       map_name human_count mouse_count rat_count total_count
1  diana_microt     7664602     3747171         0    11411773
2         elmmo     3959112     1449133    547191     5955436
3     microcosm      762987      534735    353378     1651100
4   mir2disease        2875           0         0        2875
5       miranda     5429955     2379881    247368     8057204
6         mirdb     1990425     1091263    199250     3280938
7     mirecords        2425         449       171        3045
8    mirtarbase      957034      116689      1384     1075107
9  pharmaco_mir         308           5         0         313
10     phenomir       15138         491         0       15629
11       pictar      404066      302236         0      706302
12         pita     7710936     5163153         0    12874089
13      tarbase     1290272      473266      1031     1764713
14   targetscan    13964425    10387912         0    24352337
```

```
#db.count |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

4 Objective

1.Bioinformatic prediction of miR-520-5p and miR-655-3p mRNA targets

Publically available databases will be used to search for mir-520d-5p and mir-655-3p predicted mRNA targets. Most relevant mRNA search will be selected according to the highest score prediction from several algorithms. We intend to construct a list (35-40 mRNA targets) that will

be further categorized according to their putative role on metabolism (by bibliographic analysis) to finally summarize 10-15 miRNA-mRNA targets with potential interest in metabolism. Gene ontology /annotation and enrichment analysis of mRNA targets will be performed using public data base (genecodis)

5 Methodology

To create the validated and predicted table, it was necessary to examine specific parameters available in the tables to ensure reliable results. These parameters were analyzed for each scenario (Validated and Predicted) since they vary depending on whether the goal is to validate or predict targets.

5.1 Parameters

5.1.1 Validated

In this case, there is no need to specify a cutoff, as this table will contain only validated targets. Therefore, you should filter based on the columns of the table that you will be obtaining:

5.1.1.1 Table Information

The columns represent various features or metadata associated with the interactions between microRNAs (miRNAs) and their potential targets.

1. **database**: Indicates the source database of the information (e.g., *mirtarbase*, *targetscan*, *miRDB*). This is useful if you want to focus on specific sources or compare results across databases.
2. **mature_mirna_acc**: The accession identifier for the mature version of the miRNA in the database (e.g., *MIMAT0003331*). This unique identifier is useful for unambiguous miRNA identification.
3. **mature_mirna_id**: The name of the mature miRNA in standard format (e.g., *hsa-miR-655-3p*). You can use this to filter specific miRNA interactions.
4. **target_symbol**: The gene symbol of the miRNA target (e.g., *TGFBR2*). Useful for focusing on results for a particular gene.
5. **target_entrez**: The unique gene identifier in the *Entrez Gene* database (e.g., *7048*). This is important for bioinformatic analyses requiring unique gene identifiers.

6. **target_ensembl**: The unique gene identifier in the *Ensembl* database (e.g., *ENSG00000163513*). Similar to the previous field, this is helpful for integration with tools and analyses using Ensembl.
7. **experiment**: Describes the experimental methods used to validate the interaction (e.g., *Luciferase reporter assay/Western blot/qRT-PCR*). This field is crucial if you want to prioritize results with strong experimental evidence.
8. **support_type**: Indicates the level of support for the interaction, such as "Functional MTI" (functional miRNA-target interaction). If you are only interested in functional or strongly supported interactions, this field is essential for filtering.
9. **pubmed_id**: The identifier of the article in PubMed where the interaction was reported (e.g., *23690952*). Useful for exploring the original studies to obtain more details.
10. **DB.link**: A link to the entry in the original database where the interaction can be verified. This is helpful for directly consulting the source.
11. **type**: Specifies whether the interaction is "validated" or "predicted." Validated interactions are usually more reliable as they are backed by experimental evidence. This is one of the most important fields for filtering reliable interactions.

5.1.1.2 Which Columns Are Useful for Filtering More Accurate Results?

To obtain more specific and reliable results, the following columns can be particularly useful:

- **type**: Filter by "validated" if you are only interested in experimentally confirmed interactions.
- **experiment**: Prioritize interactions with high-confidence experimental methods, such as *Luciferase reporter assay* or *Western blot*.
- **support_type**: Use "Functional MTI" to focus on functional interactions supported by experimental evidence.
- **pubmed_id**: Examine the references to validate the reliability of the reported data.
- **database**: Select databases that are most relevant or recognized for your research.

5.1.2 Predicted

In the function `get_multimir` you need to adjust the following paramaters: #### **1. predicted.cutoff.type = " "** ##### `predicted.cutoff.type = "p"` - **Default Value (`predicted.cutoff = NULL`)**:
- Automatically selects the **top 20% of predictions** if percentage-based cutoff is used

(`predicted.cutoff.type = "p"`). This behavior is generally suitable for human studies because:
- It provides a balanced set of high-quality predictions.
- The filter is designed to reduce noise (i.e., less reliable interactions).
- The resulting set remains manageable, even in large-scale analyses.
- **Recommended if you prioritize quality over quantity** and do not have a specific criterion to customize the cutoff.

- **Customizing the Percentage (`predicted.cutoff = X`)**:
  - If you want a more restricted (stricter) set, you can reduce this percentage (e.g., selecting the top 10%, `predicted.cutoff = 10`).

5.1.2.0.1 `predicted.cutoff.type = "n"`

Absolute cutoff is useful if you have a fixed dataset size in mind or if you want to maximize the scope of the analysis without focusing on relative proportions.

- **Default Value (`predicted.cutoff = NULL`)**:
  - Selects the **top 300,000 predictions** or all available records if there are fewer than 300,000.
  - This approach can be helpful for comprehensive studies but is less precise if you aim to prioritize high-quality predictions.
- **Customizing the Number (`predicted.cutoff = X`)**:
  - For example, setting `predicted.cutoff = 50,000` will select the top 50,000 predictions, providing a balance between result size and relevance.
  - This is useful if you have computational limitations or know the exact number of records you need to process.

5.1.2.1 predicted.site = "conserved"

For humans, the default settings are reasonable in most cases because:

- **Conserved Predictions**: Conserved target sites (`predicted.site = "conserved"`) are generally more reliable for humans, as they are supported by evolutionary criteria and are more likely to be functional.
- **Top 20% as Default**: Provides high-quality predictions by filtering interactions with the highest scores.

- **Biological Significance**: Conserved sites are often associated with critical functions in organisms. In the case of miRNAs, these sites are more likely to represent real and functional interactions between the miRNA and the target gene. For example, if a binding site for a miRNA in a human gene is also present in mice and rats, it is likely to have an essential biological role.
- **Higher Likelihood of Experimental Validation**: Since conserved sites hold greater biological relevance, they are more likely to have been experimentally validated, increasing confidence in the prediction.
- **Reduction of Noise in Predictions**: Prediction databases such as **TargetScan**, **miRanda**, and **PITA** generate numerous potential interactions, but not all are biologically significant.

- Filtering for conserved sites removes less reliable predictions (based solely on sequence matches) and prioritizes those more likely to be functional.
- **Reliability of the Database Using `predicted.site = "conserved"`**

1. **Focus on More Robust Predictions**:
    - By setting `predicted.site = "conserved"`, you obtain results that meet conservation criteria in at least one of the three databases (**TargetScan**, **miRanda**, or **PITA**). This means the selected target sites have additional support in terms of evolutionary relevance.
2. **Limitations**:
    - **Lower Coverage**: Some miRNAs or genes may not have conserved target sites in these databases, potentially reducing the number of predictions.
    - **Context Dependency**: Not all functional interactions are conserved. Species- or tissue-specific interactions might be excluded.

- **When to Use Conserved Sites**

1. **Recommended**:

- If you are looking for miRNA-gene interactions with a high likelihood of being functional and relevant.
- For comparative studies across species or general approaches in evolutionary biology.

2. **Avoid It**:

- If you are working with species-specific miRNAs or genes with emerging, less-studied roles.

- If you want to maximize the coverage of predictions and are willing to analyze additional results (potentially with more noise).

5.2 Search

5.3 miRNA targets

5.3.1 hsa-miR-520d-5p

5.3.1.1 1. Validated Datasets

```
# The default is to search validated interactions in human
val_hsa_miR_520d_5p <- get_multimir(org = "hsa",
                                    mirna = 'hsa-miR-520d-5p',
                                    table = "validated",
                                    add.link = TRUE,
                                    use.tibble = TRUE,
                                    summary = TRUE)
```

```
Searching mirecords ...

Warning: `as_data_frame()` was deprecated in tibble 2.0.0.
i Please use `as_tibble()` (with slightly different semantics) to convert to a
  tibble, or `as.data.frame()` to convert to a data frame.
i The deprecated feature was likely used in the multiMiR package.
  Please report the issue at <https://github.com/KechrisLab/multiMiR/issues>.

Searching mirtarbase ...
Searching tarbase ...

Some of the links to external databases may be broken due to outdated identifiers in these da
```

```
# Check which types of associations were returned
table(val_hsa_miR_520d_5p@data$type)
```

```
validated
      471
```

```r
dim(val_hsa_miR_520d_5p@data)
```

```
[1] 471  11
```

```r
# Detailed information of the validated miRNA-target interaction
head(val_hsa_miR_520d_5p@data)
```

```
# A tibble: 6 x 11
  database   mature_mirna_acc mature_mirna_id target_symbol target_entrez
  <chr>      <chr>            <chr>           <chr>         <chr>
1 mirtarbase MIMAT0002855     hsa-miR-520d-5p NCOA3         8202
2 mirtarbase MIMAT0002855     hsa-miR-520d-5p NCOA3         8202
3 mirtarbase MIMAT0002855     hsa-miR-520d-5p NCOA3         8202
4 mirtarbase MIMAT0002855     hsa-miR-520d-5p NCOA3         8202
5 mirtarbase MIMAT0002855     hsa-miR-520d-5p CD164         8763
6 mirtarbase MIMAT0002855     hsa-miR-520d-5p CD164         8763
# i 6 more variables: target_ensembl <chr>, experiment <chr>,
#   support_type <chr>, pubmed_id <chr>, DB.link <chr>, type <chr>
```

```r
# Filter with the updated databases, experiment, support_type y type
filter_val_hsa_miR_520d_5p <-val_hsa_miR_520d_5p@data[
  val_hsa_miR_520d_5p@data[["database"]] %in% c("mirtarbase", "tarbase") & # Filtrar por b
  grepl("Luciferase|Western blot|qRT-PCR", val_hsa_miR_520d_5p@data[["experiment"]], ignor
  val_hsa_miR_520d_5p@data[["support_type"]]  %in% c("Functional MTI", "Functional MTI (We
  val_hsa_miR_520d_5p@data[["type"]] == "validated", # Filtro por tipo validado
]

print(filter_val_hsa_miR_520d_5p)
```

```
# A tibble: 1 x 11
  database   mature_mirna_acc mature_mirna_id target_symbol target_entrez
  <chr>      <chr>            <chr>           <chr>         <chr>
1 mirtarbase MIMAT0002855     hsa-miR-520d-5p PPIB          5479
# i 6 more variables: target_ensembl <chr>, experiment <chr>,
#   support_type <chr>, pubmed_id <chr>, DB.link <chr>, type <chr>
```

```r
# Table with kable
#filter_val_hsa_miR_520d_5p |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

5.3.1.2 2. Predicted

```r
# Conserved Sites
# The default is to search validated interactions in human
pre_hsa_miR_520d_5p <- get_multimir(org = "hsa",
                                    mirna = 'hsa-miR-520d-5p',
                                    table = "predicted",
                                    predicted.cutoff = 10,
                                    predicted.cutoff.type = "p",
                                    predicted.site = "conserved",
                                    add.link = TRUE,
                                    use.tibble = TRUE,
                                    summary = TRUE)
```

```
Searching diana_microt ...
Searching elmmo ...
Searching microcosm ...
Searching miranda ...
Searching mirdb ...
Searching pictar ...
Searching pita ...
Searching targetscan ...

Some of the links to external databases may be broken due to outdated identifiers in these da
```

```r
# Check which types of associations were returned
table(pre_hsa_miR_520d_5p@data$type)
```

```
predicted
     7954
```

```r
dim(pre_hsa_miR_520d_5p@data)
```

```
[1] 7954    9
```

```
  # Detailed information of the validated miRNA-target interaction
  head(pre_hsa_miR_520d_5p@data)
```

```
# A tibble: 6 x 9
  database     mature_mirna_acc mature_mirna_id target_symbol target_entrez
  <chr>        <chr>            <chr>           <chr>         <chr>
1 diana_microt MIMAT0002855     hsa-miR-520d-5p TEAD1         7003
2 diana_microt MIMAT0002855     hsa-miR-520d-5p ELAVL2        1993
3 diana_microt MIMAT0002855     hsa-miR-520d-5p ELAVL2        1993
4 diana_microt MIMAT0002855     hsa-miR-520d-5p CPEB3         22849
5 diana_microt MIMAT0002855     hsa-miR-520d-5p CPEB3         22849
6 diana_microt MIMAT0002855     hsa-miR-520d-5p ATAD2B        54454
# i 4 more variables: target_ensembl <chr>, score <chr>, DB.link <chr>,
#   type <chr>
```

```
  # Filter with the updated databases (targetscan)
  filter_pre_hsa_miR_520d_5p <-  pre_hsa_miR_520d_5p@data[pre_hsa_miR_520d_5p@data[["databas
  #as.data.frame(pre_hsa_miR_520d_5p@data)
  dim(filter_pre_hsa_miR_520d_5p)
```

```
[1] 575   9
```

```
  print(filter_pre_hsa_miR_520d_5p)
```

```
# A tibble: 575 x 9
   database mature_mirna_acc mature_mirna_id target_symbol target_entrez
   <chr>    <chr>            <chr>           <chr>         <chr>
 1 mirdb    MIMAT0002855     hsa-miR-520d-5p PELI2         57161
 2 mirdb    MIMAT0002855     hsa-miR-520d-5p PUM2          23369
 3 mirdb    MIMAT0002855     hsa-miR-520d-5p ELAVL2        1993
 4 mirdb    MIMAT0002855     hsa-miR-520d-5p MTM1          4534
 5 mirdb    MIMAT0002855     hsa-miR-520d-5p STRBP         55342
 6 mirdb    MIMAT0002855     hsa-miR-520d-5p PTPRE         5791
 7 mirdb    MIMAT0002855     hsa-miR-520d-5p ELAVL2        1993
 8 mirdb    MIMAT0002855     hsa-miR-520d-5p CPEB2         132864
 9 mirdb    MIMAT0002855     hsa-miR-520d-5p MAPK1         5594
```

```
10 mirdb    MIMAT0002855     hsa-miR-520d-5p ARPP19        10776
# i 565 more rows
# i 4 more variables: target_ensembl <chr>, score <chr>, DB.link <chr>,
#   type <chr>
```

```
# Table with kable
#filter_pre_hsa_miR_520d_5p |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

5.3.1.3 3. Filter and Combine

```
combined_targets_miR_520d_5p <- filter_val_hsa_miR_520d_5p %>%
  mutate(type = "validated") %>%
  bind_rows(
    filter_pre_hsa_miR_520d_5p %>% mutate(type = "predicted") # Agregar etiquetas de orige
  ) %>%
  distinct()
```

5.3.1.4 4. Resumir targets por miRNA

```
summary_targets_miR_520d_5p <- combined_targets_miR_520d_5p %>%
  group_by(mature_mirna_id, type) %>%
  summarise(
    num_targets = n(),
    top_targets = paste0(unique(target_symbol)[1:20], collapse = ", ")
  )
```

```
`summarise()` has grouped output by 'mature_mirna_id'. You can override using
the `.groups` argument.
```

```
print(summary_targets_miR_520d_5p)
```

```
# A tibble: 2 x 4
# Groups:   mature_mirna_id [1]
  mature_mirna_id type      num_targets top_targets
  <chr>           <chr>           <int> <chr>
1 hsa-miR-520d-5p predicted         575 PELI2, PUM2, ELAVL2, MTM1, STRBP, PTPRE~
2 hsa-miR-520d-5p validated           1 PPIB, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

```
#summary_targets_miR_520d_5p |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

5.3.1.5 5: Análisis funcional

5.3.1.5.1 Validated

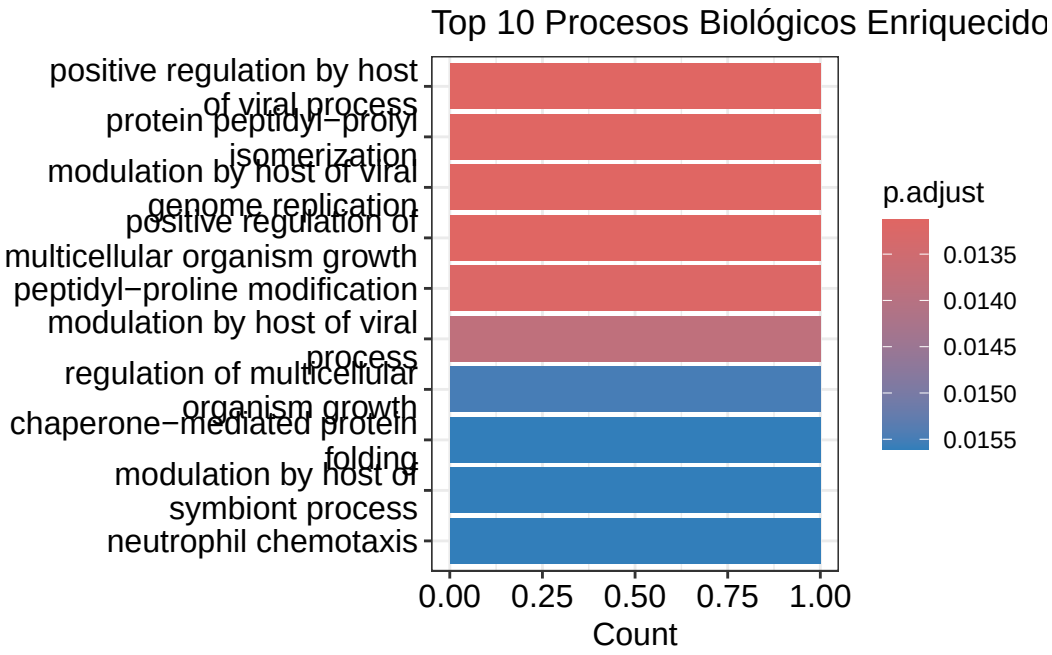
```
# Gene Ontology (GO) para todos los genes combinados
gene_symbols_miR_520d_5p_validated <- filter_val_hsa_miR_520d_5p$target_symbol %>% unique(
go_results_miR_520d_5p_validated <- enrichGO(
  gene = gene_symbols_miR_520d_5p_validated,
  OrgDb = org.Hs.eg.db,
  keyType = "SYMBOL",
  ont = "BP",  # BP Biological Processes
  pAdjustMethod = "BH",
  qvalueCutoff = 0.05, # Relacionada con el control de la tasa de falsos descubrimientos (
  pvalueCutoff = 0.05
)

head(go_results_miR_520d_5p_validated@result,20)
```

```
                   ID                                                 Description
GO:0044794 GO:0044794            positive regulation by host of viral process
GO:0000413 GO:0000413                  protein peptidyl-prolyl isomerization
GO:0044827 GO:0044827          modulation by host of viral genome replication
GO:0040018 GO:0040018    positive regulation of multicellular organism growth
GO:0018208 GO:0018208                           peptidyl-proline modification
GO:0044788 GO:0044788                     modulation by host of viral process
GO:0040014 GO:0040014             regulation of multicellular organism growth
GO:0061077 GO:0061077                      chaperone-mediated protein folding
GO:0051851 GO:0051851                 modulation by host of symbiont process
GO:0030593 GO:0030593                                  neutrophil chemotaxis
GO:0051702 GO:0051702 biological process involved in interaction with symbiont
GO:0071621 GO:0071621                                granulocyte chemotaxis
GO:0019079 GO:0019079                             viral genome replication
GO:1990266 GO:1990266                                 neutrophil migration
GO:0035264 GO:0035264                          multicellular organism growth
GO:0097530 GO:0097530                                granulocyte migration
GO:0048639 GO:0048639            positive regulation of developmental growth
```

```
GO:0060348 GO:0060348                                      bone development
GO:0050821 GO:0050821                                  protein stabilization
GO:0006457 GO:0006457                                        protein folding
           GeneRatio   BgRatio  RichFactor FoldEnrichment    zScore      pvalue
GO:0044794       1/1  20/18888 0.050000000      944.40000 30.714817 0.001058873
GO:0000413       1/1  25/18888 0.040000000      755.52000 27.468527 0.001323592
GO:0044827       1/1  26/18888 0.038461538      726.46154 26.934393 0.001376535
GO:0040018       1/1  31/18888 0.032258065      609.29032 24.663542 0.001641254
GO:0018208       1/1  39/18888 0.025641026      484.30769 21.984260 0.002064803
GO:0044788       1/1  49/18888 0.020408163      385.46939 19.607891 0.002594240
GO:0040014       1/1  64/18888 0.015625000      295.12500 17.150073 0.003388395
GO:0061077       1/1  75/18888 0.013333333      251.84000 15.837929 0.003970775
GO:0051851       1/1  86/18888 0.011627907      219.62791 14.786071 0.004553155
GO:0030593       1/1 107/18888 0.009345794      176.52336 13.248523 0.005664972
GO:0051702       1/1 119/18888 0.008403361      158.72269 12.558769 0.006300296
GO:0071621       1/1 128/18888 0.007812500      147.56250 12.106300 0.006776789
GO:0019079       1/1 129/18888 0.007751938      146.41860 12.058964 0.006829733
GO:1990266       1/1 129/18888 0.007751938      146.41860 12.058964 0.006829733
GO:0035264       1/1 147/18888 0.006802721      128.48980 11.291138 0.007782719
GO:0097530       1/1 154/18888 0.006493506      122.64935 11.029476 0.008153325
GO:0048639       1/1 169/18888 0.005917160      111.76331 10.524415 0.008947480
GO:0060348       1/1 215/18888 0.004651163       87.85116  9.319397 0.011382889
GO:0050821       1/1 218/18888 0.004587156       86.64220  9.254307 0.011541720
GO:0006457       1/1 225/18888 0.004444444       83.94667  9.107506 0.011912325
             p.adjust qvalue geneID Count
GO:0044794 0.01313003     NA   PPIB     1
GO:0000413 0.01313003     NA   PPIB     1
GO:0044827 0.01313003     NA   PPIB     1
GO:0040018 0.01313003     NA   PPIB     1
GO:0018208 0.01321474     NA   PPIB     1
GO:0044788 0.01383595     NA   PPIB     1
GO:0040014 0.01548980     NA   PPIB     1
GO:0061077 0.01561082     NA   PPIB     1
GO:0051851 0.01561082     NA   PPIB     1
GO:0030593 0.01561082     NA   PPIB     1
GO:0051702 0.01561082     NA   PPIB     1
GO:0071621 0.01561082     NA   PPIB     1
GO:0019079 0.01561082     NA   PPIB     1
GO:1990266 0.01561082     NA   PPIB     1
GO:0035264 0.01630665     NA   PPIB     1
GO:0097530 0.01630665     NA   PPIB     1
GO:0048639 0.01684232     NA   PPIB     1
GO:0060348 0.01840514     NA   PPIB     1
```

```
GO:0050821 0.01840514     NA   PPIB     1
GO:0006457 0.01840514     NA   PPIB     1
```

```
#go_results_miR_520d_5p_validated@result |>
#  kable(format = "html") |>
#  kable_styling("striped")


# Visualización del análisis funcional
barplot(go_results_miR_520d_5p_validated, showCategory = 10, title = "Top 10 Procesos Biol
```



5.3.1.5.2 Predicted

```
# Gene Ontology (GO) para todos los genes combinados
gene_symbols_miR_520d_5p_predicted <- filter_pre_hsa_miR_520d_5p$target_symbol %>% unique(
go_results_miR_520d_5p_predicted <- enrichGO(
  gene = gene_symbols_miR_520d_5p_predicted,
  OrgDb = org.Hs.eg.db,
  keyType = "SYMBOL",
  ont = "BP",  # Biological Processes
  pAdjustMethod = "BH",
  qvalueCutoff = 0.05, # Relacionada con el control de la tasa de falsos descubrimientos (
```

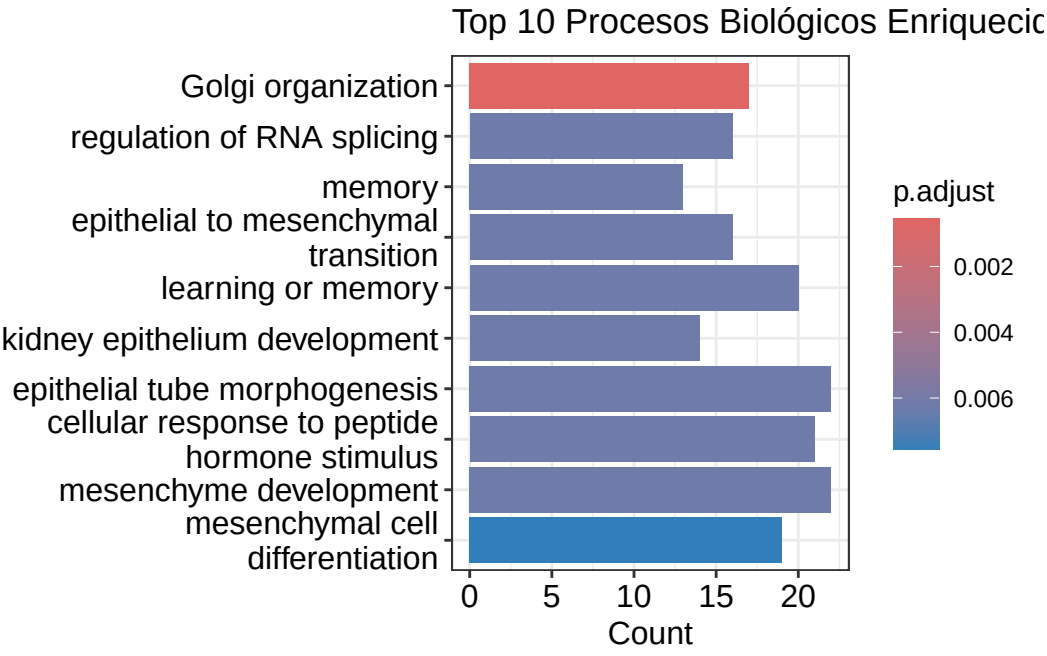
```
  pvalueCutoff = 0.05
)

head(go_results_miR_520d_5p_predicted@result,20 )
```

```
                  ID                                   Description GeneRatio
GO:0007030 GO:0007030                             Golgi organization    17/439
GO:0043484 GO:0043484                  regulation of RNA splicing    16/439
GO:0007613 GO:0007613                                      memory    13/439
GO:0001837 GO:0001837          epithelial to mesenchymal transition    16/439
GO:0007611 GO:0007611                           learning or memory    20/439
GO:0072073 GO:0072073                kidney epithelium development    14/439
GO:0060562 GO:0060562                epithelial tube morphogenesis    22/439
GO:0071375 GO:0071375 cellular response to peptide hormone stimulus    21/439
GO:0060485 GO:0060485                         mesenchyme development    22/439
GO:0048762 GO:0048762              mesenchymal cell differentiation    19/439
GO:0050803 GO:0050803   regulation of synapse structure or activity    18/439
GO:0048568 GO:0048568                  embryonic organ development    26/439
GO:1901653 GO:1901653                 cellular response to peptide    23/439
GO:0032869 GO:0032869         cellular response to insulin stimulus    16/439
GO:0072001 GO:0072001                     renal system development    21/439
GO:0048638 GO:0048638            regulation of developmental growth    21/439
GO:0021543 GO:0021543                          pallium development    15/439
GO:0048639 GO:0048639   positive regulation of developmental growth    14/439
GO:0050807 GO:0050807            regulation of synapse organization    17/439
GO:0003184 GO:0003184                 pulmonary valve morphogenesis     5/439
            BgRatio RichFactor FoldEnrichment   zScore       pvalue
GO:0007030 156/18888 0.10897436       4.688628 7.136124 1.401228e-07
GO:0043484 186/18888 0.08602151       3.701080 5.710541 7.534365e-06
GO:0007613 126/18888 0.10317460       4.439093 5.974712 7.632717e-06
GO:0001837 187/18888 0.08556150       3.681288 5.684068 8.066780e-06
GO:0007611 278/18888 0.07194245       3.095328 5.429109 8.657087e-06
GO:0072073 152/18888 0.09210526       3.962834 5.657415 1.283945e-05
GO:0060562 335/18888 0.06567164       2.825526 5.200328 1.319290e-05
GO:0071375 311/18888 0.06752412       2.905229 5.225970 1.364153e-05
GO:0060485 336/18888 0.06547619       2.817117 5.184232 1.381931e-05
GO:0048762 269/18888 0.07063197       3.038945 5.195540 1.875100e-05
GO:0050803 249/18888 0.07228916       3.110245 5.170691 2.276126e-05
GO:0048568 455/18888 0.05714286       2.458575 4.858069 2.557950e-05
GO:1901653 376/18888 0.06117021       2.631852 4.930322 2.615897e-05
GO:0032869 206/18888 0.07766990       3.341752 5.213029 2.702248e-05
GO:0072001 329/18888 0.06382979       2.746280 4.929035 3.149411e-05
```

```
GO:0048638 330/18888 0.06363636       2.737958 4.913128 3.292573e-05
GO:0021543 191/18888 0.07853403       3.378931 5.097297 4.221070e-05
GO:0048639 169/18888 0.08284024       3.564206 5.165142 4.222932e-05
GO:0050807 243/18888 0.06995885       3.009983 4.864540 5.768813e-05
GO:0003184  19/18888 0.26315789      11.322383 6.943999 5.888725e-05
              p.adjust       qvalue
GO:0007030 0.000563854 0.0004945596
GO:0043484 0.006178765 0.0054194308
GO:0007613 0.006178765 0.0054194308
GO:0001837 0.006178765 0.0054194308
GO:0007611 0.006178765 0.0054194308
GO:0072073 0.006178765 0.0054194308
GO:0060562 0.006178765 0.0054194308
GO:0071375 0.006178765 0.0054194308
GO:0060485 0.006178765 0.0054194308
GO:0048762 0.007545403 0.0066181164
GO:0050803 0.007767033 0.0068125098
GO:0048568 0.007767033 0.0068125098
GO:1901653 0.007767033 0.0068125098
GO:0032869 0.007767033 0.0068125098
GO:0072001 0.008280821 0.0072631558
GO:0048638 0.008280821 0.0072631558
GO:0021543 0.009440599 0.0082804038
GO:0048639 0.009440599 0.0082804038
GO:0050807 0.011848115 0.0103920502
GO:0003184 0.011848115 0.0103920502

GO:0007030                        MAPK1/RAB33B/TMED7/GOLGA6B/PDCD10/GOLGA6A/GOLGA6D/GOLGA8M/VA
GO:0043484                                                RBM39/SRSF10/MBNL1/RBM12/CLK4/RB
GO:0007613                                                                          SYT4/S
GO:0001837                                                      DLG5/USF3/HEY2/HMGA2/SMAD3/
GO:0007611                              MAPK1/SYT4/SCN2A/MEF2C/CNR1/TAFA2/PPP1R1B/AMFR/KCNK2/
GO:0072073                                                                  MTSS1/EPHA7/MEI
GO:0060562                        DLG5/MTSS1/EPHA7/FZD3/MEF2C/ARL13B/SIX4/IRX2/CTHRC1/SMAD3/FGI
GO:0071375                           PTPRE/CPEB2/MAPK1/PPP3CA/NR4A2/ATP2B1/PHIP/APPL2/PRKCB/FBI
GO:0060485                        MAPK1/ADAMTS5/DLG5/USF3/MEF2C/SIX4/HEY2/HMGA2/SMAD3/EPHA3/JAG
GO:0048762                                    MAPK1/DLG5/USF3/MEF2C/HEY2/HMGA2/SMAD3/EPHA3
GO:0050803                                  PTPN13/DLG5/LRRTM3/EPHA7/SLITRK4/AGRN/MEF2C/SIX4/NI
GO:0048568 MAPK1/TEAD1/E2F8/SP3/FZD3/MEF2C/ARL13B/PRKRA/SIX4/FBN1/EGLN1/CTHRC1/NCOA1/TSHZ1/HI
GO:1901653                PTPRE/CPEB2/MAPK1/PPP3CA/NR4A2/ATP2B1/PHIP/APPL2/PRKCB/FBN1/SRD5A1/Z
GO:0032869                                                      PTPRE/CPEB2/MAPK1/ATP2B1/PHI
GO:0072001                              DLG5/MTSS1/EPHA7/PPP3CA/MEF2C/SIX4/FBN1/IRX2/SMAD3/JAGI
GO:0048638                                      PUM2/MTM1/SYT4/EPHA7/JARID2/AGRN/MEF2C/SIX4/WWC
```

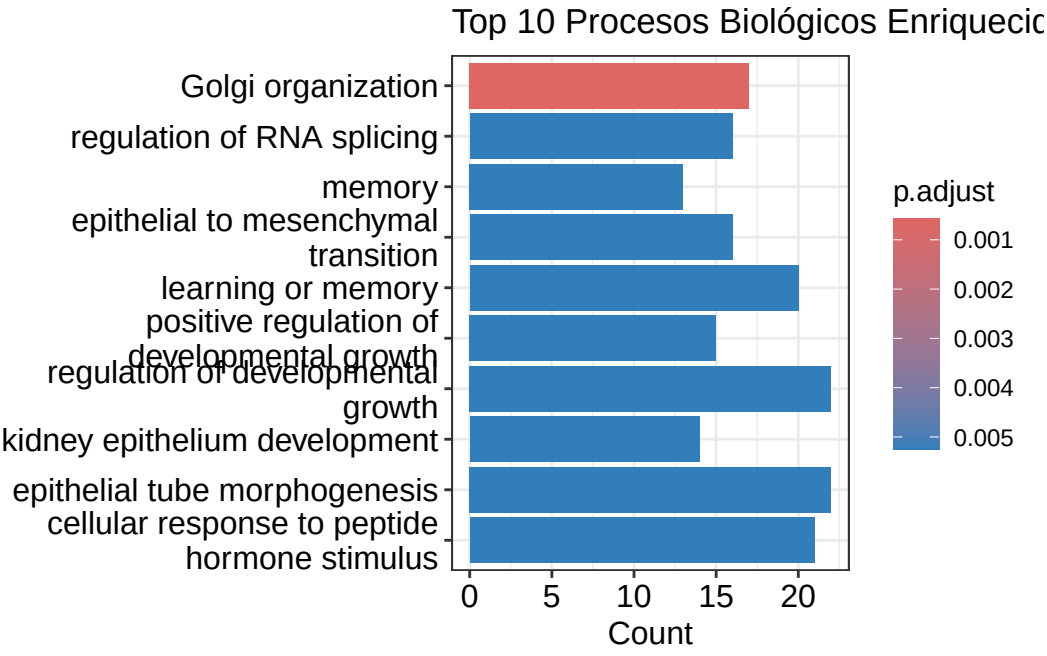
```
GO:0021543                                                                CDK6/ZMIZ1/DLX1/CEP120/S
GO:0048639                                                                                PUM2/MTM1/S
GO:0050807                                                  PTPN13/DLG5/LRRTM3/EPHA7/SLITRK4/AGRN/MEF2C
GO:0003184
           Count
GO:0007030    17
GO:0043484    16
GO:0007613    13
GO:0001837    16
GO:0007611    20
GO:0072073    14
GO:0060562    22
GO:0071375    21
GO:0060485    22
GO:0048762    19
GO:0050803    18
GO:0048568    26
GO:1901653    23
GO:0032869    16
GO:0072001    21
GO:0048638    21
GO:0021543    15
GO:0048639    14
GO:0050807    17
GO:0003184     5
```

```
#go_results_miR_520d_5p_predicted@result |>
#  kable(format = "html") |>
#  kable_styling("striped")


# Visualización del análisis funcional
barplot(go_results_miR_520d_5p_predicted, showCategory = 10, title = "Top 10 Procesos Biol
```



5.3.1.5.3 Combined

```
# Gene Ontology (GO) para todos los genes combinados
gene_symbols_miR_520d_5p_combined <- combined_targets_miR_520d_5p$target_symbol %>% unique
go_results_miR_520d_5p_combined <- enrichGO(
  gene = gene_symbols_miR_520d_5p_combined,
  OrgDb = org.Hs.eg.db,
  keyType = "SYMBOL",
  ont = "BP",  # Biological Processes
  pAdjustMethod = "BH",
  qvalueCutoff = 0.05, # Relacionada con el control de la tasa de falsos descubrimientos (
  pvalueCutoff = 0.05
)

# Visualización del análisis funcional
barplot(go_results_miR_520d_5p_combined, showCategory = 10, title = "Top 10 Procesos Bioló
```



5.3.1.6 6: Exportar resultados

```
write.csv(filter_val_hsa_miR_520d_5p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRN
write.csv(filter_pre_hsa_miR_520d_5p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRN
write.csv(combined_targets_miR_520d_5p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/mi
write.csv(summary_targets_miR_520d_5p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/miR
```

5.3.2 hsa-miR-655-3p

5.3.2.1 1. Validated

```
# The default is to search validated interactions in human
val_hsa_miR_655_3p <- get_multimir(org = "hsa",
                                   mirna = 'hsa-miR-655-3p',
                                   table = "validated",
                                   add.link = TRUE,
                                   use.tibble = TRUE,
                                   summary = TRUE)
```

```
Searching mirecords ...
Searching mirtarbase ...
```

```
Searching tarbase ...

Some of the links to external databases may be broken due to outdated identifiers in these da
```

```r
# Check which types of associations were returned
table(val_hsa_miR_655_3p@data$type)
```

```
validated
      319
```

```r
# Detailed information of the validated miRNA-target interaction
head(val_hsa_miR_655_3p@data)
```

```
# A tibble: 6 x 11
  database   mature_mirna_acc mature_mirna_id target_symbol target_entrez
  <chr>      <chr>            <chr>           <chr>         <chr>
1 mirtarbase MIMAT0003331     hsa-miR-655-3p  TGFBR2        7048
2 mirtarbase MIMAT0003331     hsa-miR-655-3p  TGFBR2        7048
3 mirtarbase MIMAT0003331     hsa-miR-655-3p  CAPRIN2       65981
4 mirtarbase MIMAT0003331     hsa-miR-655-3p  CAPRIN2       65981
5 mirtarbase MIMAT0003331     hsa-miR-655-3p  CAPRIN2       65981
6 mirtarbase MIMAT0003331     hsa-miR-655-3p  CAPRIN2       65981
# i 6 more variables: target_ensembl <chr>, experiment <chr>,
#   support_type <chr>, pubmed_id <chr>, DB.link <chr>, type <chr>
```

```r
# Filter with the updated databases, experiment, support_type y type
filter_val_hsa_miR_655_3p <- val_hsa_miR_655_3p@data[
  val_hsa_miR_655_3p@data[["database"]] %in% c("mirtarbase", "tarbase") & # Filtrar por ba
  grepl("Luciferase|Western blot|qRT-PCR", val_hsa_miR_655_3p@data[["experiment"]], ignore
  val_hsa_miR_655_3p@data[["support_type"]] %in% c("Functional MTI", "Functional MTI (Weak
  val_hsa_miR_655_3p@data[["type"]] == "validated", # Filtro por tipo validado
]

print(filter_val_hsa_miR_655_3p)
```

```
# A tibble: 6 x 11
  database   mature_mirna_acc mature_mirna_id target_symbol target_entrez
  <chr>      <chr>            <chr>           <chr>         <chr>
```

```
1 mirtarbase MIMAT0003331     hsa-miR-655-3p  TGFBR2        7048
2 mirtarbase MIMAT0003331     hsa-miR-655-3p  ZEB1          6935
3 mirtarbase MIMAT0003331     hsa-miR-655-3p  ADAM10        102
4 mirtarbase MIMAT0003331     hsa-miR-655-3p  PRRX1         5396
5 mirtarbase MIMAT0003331     hsa-miR-655-3p  MAGI2         9863
6 mirtarbase MIMAT0003331     hsa-miR-655-3p  PTTG1         9232
# i 6 more variables: target_ensembl <chr>, experiment <chr>,
#   support_type <chr>, pubmed_id <chr>, DB.link <chr>, type <chr>
```

```r
# Table with kable

#filter_val_hsa_miR_655_3p |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

5.3.2.2 2. Predicted

```r
# Conserved Sites
# The default is to search validated interactions in human
pre_hsa_miR_655_3p <- get_multimir(org = "hsa",
                                   mirna = 'hsa-miR-655-3p',
                                   table = "predicted",
                                   predicted.cutoff = 10,
                                   predicted.cutoff.type = "p",
                                   predicted.site = "conserved",
                                   add.link = TRUE,
                                   use.tibble = TRUE,
                                   summary = TRUE)
```

```
Searching diana_microt ...
Searching elmmo ...
Searching microcosm ...
Searching miranda ...
Searching mirdb ...
Searching pictar ...
Searching pita ...
Searching targetscan ...

Some of the links to external databases may be broken due to outdated identifiers in these da
```

```r
# Check which types of associations were returned
table(pre_hsa_miR_655_3p@data$type)
```

```
predicted
     4453
```

```r
dim(pre_hsa_miR_655_3p@data)
```

```
[1] 4453    9
```

```r
# Detailed information of the validated miRNA-target interaction
head(pre_hsa_miR_655_3p@data)
```

```
# A tibble: 6 x 9
  database     mature_mirna_acc mature_mirna_id target_symbol target_entrez
  <chr>        <chr>            <chr>           <chr>         <chr>
1 diana_microt MIMAT0003331     hsa-miR-655-3p  INO80D        54891
2 diana_microt MIMAT0003331     hsa-miR-655-3p  INO80D        54891
3 diana_microt MIMAT0003331     hsa-miR-655-3p  CD47          961
4 diana_microt MIMAT0003331     hsa-miR-655-3p  POU2F1        5451
5 diana_microt MIMAT0003331     hsa-miR-655-3p  ZNF521        25925
6 diana_microt MIMAT0003331     hsa-miR-655-3p  CLCF1         23529
# i 4 more variables: target_ensembl <chr>, score <chr>, DB.link <chr>,
#   type <chr>
```

```r
# Filter with the updated databases (targetscan)
filter_pre_hsa_miR_655_3p <- pre_hsa_miR_655_3p@data[pre_hsa_miR_655_3p@data[["database"]]

dim(filter_pre_hsa_miR_655_3p)
```

```
[1] 34  9
```

```r
print(filter_pre_hsa_miR_655_3p)
```

```
# A tibble: 34 x 9
   database   mature_mirna_acc mature_mirna_id target_symbol target_entrez
   <chr>      <chr>            <chr>           <chr>         <chr>
 1 targetscan MIMAT0003331     hsa-miR-655-3p  ADM           133
 2 targetscan MIMAT0003331     hsa-miR-655-3p  GTF2F2        2963
 3 targetscan MIMAT0003331     hsa-miR-655-3p  C2orf76       130355
 4 targetscan MIMAT0003331     hsa-miR-655-3p  TAF13         6884
 5 targetscan MIMAT0003331     hsa-miR-655-3p  TSPAN7        7102
 6 targetscan MIMAT0003331     hsa-miR-655-3p  RPL38         6169
 7 targetscan MIMAT0003331     hsa-miR-655-3p  BRAT1         221927
 8 targetscan MIMAT0003331     hsa-miR-655-3p  RBM24         221662
 9 targetscan MIMAT0003331     hsa-miR-655-3p  PRRX2         51450
10 targetscan MIMAT0003331     hsa-miR-655-3p  RGS1          5996
# i 24 more rows
# i 4 more variables: target_ensembl <chr>, score <chr>, DB.link <chr>,
#   type <chr>
```

```
# Table with kable
#filter_pre_hsa_miR_655_3p |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

5.3.2.3 3. Filter and Combine

```
combined_targets_miR_655_3p <- filter_val_hsa_miR_655_3p %>%
  mutate(type = "validated") %>%
  bind_rows(
    filter_pre_hsa_miR_655_3p %>% mutate(type = "predicted") # Agregar etiquetas de origen
  ) %>%
  distinct()
```

5.3.2.4 4. Resumir targets por miRNA

```
summary_targets_miR_655_3p <- combined_targets_miR_655_3p %>%
  group_by(mature_mirna_id, type) %>%
  summarise(
    num_targets = n(),
    top_targets = paste0(unique(target_symbol)[1:20], collapse = ", ")
  )
```

```
`summarise()` has grouped output by 'mature_mirna_id'. You can override using
the `.groups` argument.
```

```
print(summary_targets_miR_655_3p)
```

```
# A tibble: 2 x 4
# Groups:   mature_mirna_id [1]
  mature_mirna_id type      num_targets top_targets
  <chr>           <chr>           <int> <chr>
1 hsa-miR-655-3p  predicted          34 ADM, GTF2F2, C2orf76, TAF13, TSPAN7, RP~
2 hsa-miR-655-3p  validated           6 TGFBR2, ZEB1, ADAM10, PRRX1, MAGI2, PTT~
```

```
# Table with kable
#summary_targets_miR_655_3p |>
#  kable(format = "html") |>
#  kable_styling("striped")
```

5.3.2.5 5: Análisis funcional

5.3.2.5.1 Validated

```
# Gene Ontology (GO) para todos los genes combinados
gene_symbols_miR_655_3p_validated <- filter_val_hsa_miR_655_3p$target_symbol %>% unique()
go_results_miR_655_3p_validated <- enrichGO(
  gene = gene_symbols_miR_655_3p_validated,
  OrgDb = org.Hs.eg.db,
  keyType = "SYMBOL",
  ont = "BP",  # Biological Processes
  pAdjustMethod = "BH",
  qvalueCutoff = 0.05, # Relacionada con el control de la tasa de falsos descubrimientos (
  pvalueCutoff = 0.05
)

head(go_results_miR_655_3p_validated@result,20)
```

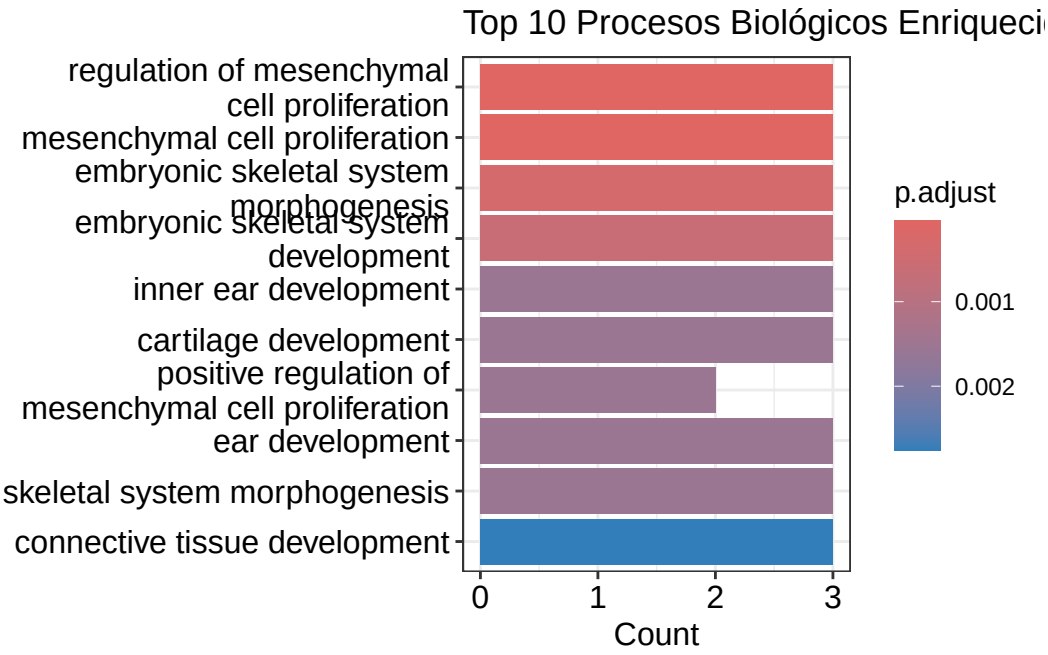
```
                   ID
GO:0010464 GO:0010464
GO:0010463 GO:0010463
```

```
GO:0048704 GO:0048704
GO:0048706 GO:0048706
GO:0048839 GO:0048839
GO:0051216 GO:0051216
GO:0002053 GO:0002053
GO:0043583 GO:0043583
GO:0048705 GO:0048705
GO:0061448 GO:0061448
GO:0048562 GO:0048562
GO:0048701 GO:0048701
GO:0090092 GO:0090092
GO:0032924 GO:0032924
GO:0090102 GO:0090102
GO:0141091 GO:0141091
GO:0007178 GO:0007178
GO:1904888 GO:1904888
GO:0048568 GO:0048568
GO:0048844 GO:0048844
                                                                                  Descri
GO:0010464                                                regulation of mesenchymal cell prolifera
GO:0010463                                                             mesenchymal cell prolifera
GO:0048704                                                  embryonic skeletal system morphoge
GO:0048706                                                    embryonic skeletal system develo
GO:0048839                                                                 inner ear develo
GO:0051216                                                                 cartilage develo
GO:0002053                                        positive regulation of mesenchymal cell prolifera
GO:0043583                                                                         ear develo
GO:0048705                                                               skeletal system morphoge
GO:0061448                                                               connective tissue develo
GO:0048562                                                               embryonic organ morphoge
GO:0048701                                                          embryonic cranial skeleton morphoge
GO:0090092 regulation of transmembrane receptor protein serine/threonine kinase signaling pa
GO:0032924                                                                activin receptor signaling pa
GO:0090102                                                                        cochlea develo
GO:0141091                      transforming growth factor beta receptor superfamily signaling pa
GO:0007178                    transmembrane receptor protein serine/threonine kinase signaling pa
GO:1904888                                                         cranial skeletal system develo
GO:0048568                                                                 embryonic organ develo
GO:0048844                                                                        artery morphoge
           GeneRatio   BgRatio  RichFactor FoldEnrichment    zScore
GO:0010464       3/6  33/18888 0.090909091      286.18182 29.227970
GO:0010463       3/6  45/18888 0.066666667      209.86667 25.005382
GO:0048704       3/6  94/18888 0.031914894      100.46809 17.233409
```

```
GO:0048706       3/6 128/18888 0.023437500       73.78125 14.727913
GO:0048839       3/6 196/18888 0.015306122       48.18367 11.836547
GO:0051216       3/6 209/18888 0.014354067       45.18660 11.450384
GO:0002053       2/6  26/18888 0.076923077      242.15385 21.934115
GO:0043583       3/6 223/18888 0.013452915       42.34978 11.072475
GO:0048705       3/6 229/18888 0.013100437       41.24017 10.921104
GO:0061448       3/6 288/18888 0.010416667       32.79167  9.691393
GO:0048562       3/6 297/18888 0.010101010       31.79798  9.536350
GO:0048701       2/6  45/18888 0.044444444      139.91111 16.630348
GO:0090092       3/6 317/18888 0.009463722       29.79180  9.215394
GO:0032924       2/6  50/18888 0.040000000      125.92000 15.766406
GO:0090102       2/6  50/18888 0.040000000      125.92000 15.766406
GO:0141091       3/6 388/18888 0.007731959       24.34021  8.280714
GO:0007178       3/6 417/18888 0.007194245       22.64748  7.968257
GO:1904888       2/6  72/18888 0.027777778       87.44444 13.100045
GO:0048568       3/6 455/18888 0.006593407       20.75604  7.603977
GO:0048844       2/6  81/18888 0.024691358       77.72840 12.335932
                 pvalue     p.adjust       qvalue            geneID Count
GO:0010464 9.683080e-08 3.950697e-05 1.386209e-05 TGFBR2/ZEB1/PRRX1     3
GO:0010463 2.514779e-07 5.130149e-05 1.800052e-05 TGFBR2/ZEB1/PRRX1     3
GO:0048704 2.361693e-06 3.211903e-04 1.126983e-04 TGFBR2/ZEB1/PRRX1     3
GO:0048706 5.990226e-06 6.110031e-04 2.143870e-04 TGFBR2/ZEB1/PRRX1     3
GO:0048839 2.150866e-05 1.552440e-03 5.447158e-04 ZEB1/ADAM10/PRRX1     3
GO:0051216 2.606295e-05 1.552440e-03 5.447158e-04 TGFBR2/ZEB1/PRRX1     3
GO:0002053 2.723852e-05 1.552440e-03 5.447158e-04      TGFBR2/PRRX1     2
GO:0043583 3.163472e-05 1.552440e-03 5.447158e-04 ZEB1/ADAM10/PRRX1     3
GO:0048705 3.424500e-05 1.552440e-03 5.447158e-04 TGFBR2/ZEB1/PRRX1     3
GO:0061448 6.782057e-05 2.756692e-03 9.672605e-04 TGFBR2/ZEB1/PRRX1     3
GO:0048562 7.432259e-05 2.756692e-03 9.672605e-04 TGFBR2/ZEB1/PRRX1     3
GO:0048701 8.275024e-05 2.783121e-03 9.765338e-04      TGFBR2/PRRX1     2
GO:0090092 9.021146e-05 2.783121e-03 9.765338e-04 TGFBR2/ZEB1/MAGI2     3
GO:0032924 1.023206e-04 2.783121e-03 9.765338e-04      TGFBR2/MAGI2     2
GO:0090102 1.023206e-04 2.783121e-03 9.765338e-04       ZEB1/ADAM10     2
GO:0141091 1.642920e-04 4.189445e-03 1.469981e-03 TGFBR2/ZEB1/MAGI2     3
GO:0007178 2.033491e-04 4.824194e-03 1.692700e-03 TGFBR2/ZEB1/MAGI2     3
GO:1904888 2.128321e-04 4.824194e-03 1.692700e-03      TGFBR2/PRRX1     2
GO:0048568 2.631082e-04 5.496656e-03 1.928651e-03 TGFBR2/ZEB1/PRRX1     3
GO:0048844 2.694439e-04 5.496656e-03 1.928651e-03      TGFBR2/PRRX1     2
```

```
#go_results_miR_655_3p_validated@result |>
#  kable(format = "html") |>
```

```
#  kable_styling("striped")


# Visualización del análisis funcional
barplot(go_results_miR_655_3p_validated, showCategory = 10, title = "Top 10 Procesos Bioló
```



5.3.2.5.2 Predicted

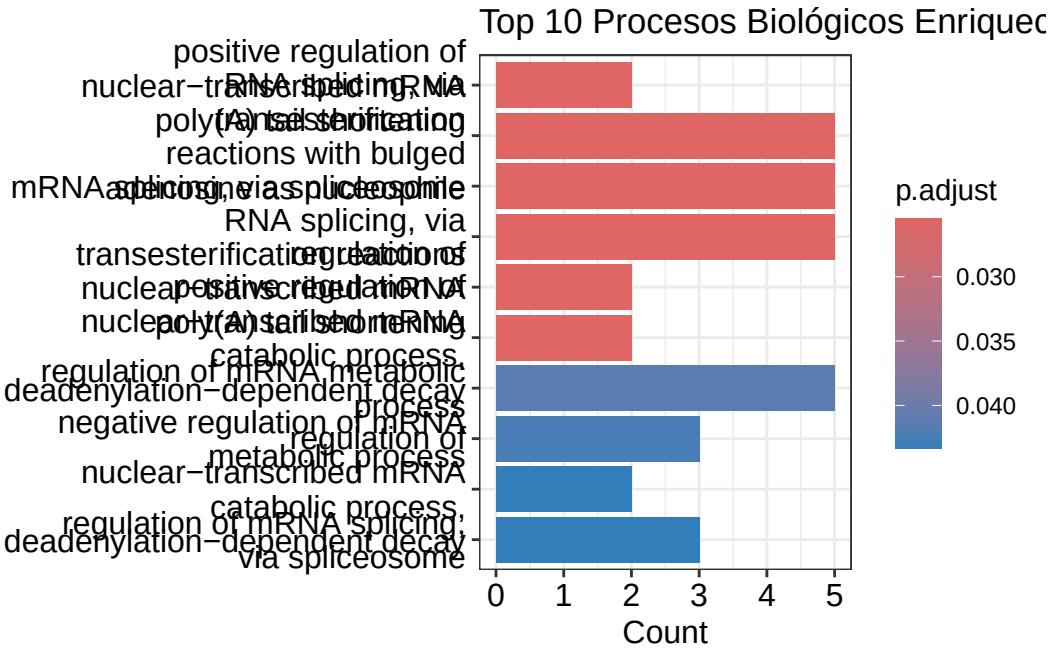
```
# Gene Ontology (GO) para todos los genes combinados
gene_symbols_miR_655_3p_predicted <- filter_pre_hsa_miR_655_3p$target_symbol %>% unique()
go_results_miR_655_3p_predicted <- enrichGO(
  gene = gene_symbols_miR_655_3p_predicted,
  OrgDb = org.Hs.eg.db,
  keyType = "SYMBOL",
  ont = "BP",  # Biological Processes
  pAdjustMethod = "BH",
  qvalueCutoff = 0.05, # Relacionada con el control de la tasa de falsos descubrimientos (
  pvalueCutoff = 0.05
)

head(go_results_miR_655_3p_predicted@result,20)
```

```
                   ID
GO:0060213 GO:0060213
GO:0000377 GO:0000377
GO:0000398 GO:0000398
GO:0000375 GO:0000375
GO:0060211 GO:0060211
GO:1900153 GO:1900153
GO:1903311 GO:1903311
GO:1903312 GO:1903312
GO:1900151 GO:1900151
GO:0048024 GO:0048024
GO:0006417 GO:0006417
GO:1903019 GO:1903019
GO:0008380 GO:0008380
GO:0000289 GO:0000289
GO:0022613 GO:0022613
GO:0050684 GO:0050684
GO:0043484 GO:0043484
GO:0000381 GO:0000381
GO:0061157 GO:0061157
GO:0050779 GO:0050779

GO:0060213                            positive regulation of nuclear-transcribed mRNA poly(A) t
GO:0000377              RNA splicing, via transesterification reactions with bulged adenosine
GO:0000398                                                                  mRNA splicing, v
GO:0000375                                                    RNA splicing, via transesterifica
GO:0060211                                       regulation of nuclear-transcribed mRNA poly(A) t
GO:1900153 positive regulation of nuclear-transcribed mRNA catabolic process, deadenylation-d
GO:1903311                                                              regulation of mRNA met
GO:1903312                                                     negative regulation of mRNA met
GO:1900151          regulation of nuclear-transcribed mRNA catabolic process, deadenylation-d
GO:0048024                                                       regulation of mRNA splicing, v
GO:0006417                                                                    regulation
GO:1903019                                          negative regulation of glycoprotein met
GO:0008380
GO:0000289                                               nuclear-transcribed mRNA poly(A) t
GO:0022613                                                        ribonucleoprotein comp
GO:0050684                                                               regulation of m
GO:0043484                                                                  regulation o
GO:0000381                                       regulation of alternative mRNA splicing, v
GO:0061157                                                                        mRNA o
GO:0050779                                                                         RNA o
           GeneRatio   BgRatio RichFactor FoldEnrichment    zScore       pvalue
```

GO:0060213	2/32	13/18888	0.15384615	90.807692	13.343613	0.0002143868
GO:0000377	5/32	337/18888	0.01483680	8.757418	5.919434	0.0002378612
GO:0000398	5/32	337/18888	0.01483680	8.757418	5.919434	0.0002378612
GO:0000375	5/32	341/18888	0.01466276	8.654692	5.876243	0.0002512008
GO:0060211	2/32	15/18888	0.13333333	78.700000	12.401604	0.0002879873
GO:1900153	2/32	15/18888	0.13333333	78.700000	12.401604	0.0002879873
GO:1903311	5/32	404/18888	0.01237624	7.305074	5.277334	0.0005462988
GO:1903312	3/32	100/18888	0.03000000	17.707500	6.900855	0.0006388390
GO:1900151	2/32	26/18888	0.07692308	45.403846	9.333512	0.0008810777
GO:0048024	3/32	112/18888	0.02678571	15.810268	6.475929	0.0008881381
GO:0006417	5/32	451/18888	0.01108647	6.543792	4.908872	0.0008976565
GO:1903019	2/32	28/18888	0.07142857	42.160714	8.978892	0.0010225968
GO:0008380	5/32	484/18888	0.01033058	6.097624	4.680218	0.0012301626
GO:0000289	2/32	32/18888	0.06250000	36.890625	8.370722	0.0013361599
GO:0022613	5/32	499/18888	0.01002004	5.914329	4.583182	0.0014083556
GO:0050684	3/32	134/18888	0.02238806	13.214552	5.845406	0.0014897551
GO:0043484	3/32	186/18888	0.01612903	9.520161	4.810519	0.0037775812
GO:0000381	2/32	55/18888	0.03636364	21.463636	6.260890	0.0039044545
GO:0061157	3/32	198/18888	0.01515152	8.943182	4.628648	0.0044992049
GO:0050779	3/32	202/18888	0.01485149	8.766089	4.571425	0.0047570607

	p.adjust	qvalue	geneID	Count
GO:0060213	0.02548688	0.02071488	TOB1/AGO2	2
GO:0000377	0.02548688	0.02071488	RBM24/SNRNP40/SRSF5/SRSF7/KHDRBS2	5
GO:0000398	0.02548688	0.02071488	RBM24/SNRNP40/SRSF5/SRSF7/KHDRBS2	5
GO:0000375	0.02548688	0.02071488	RBM24/SNRNP40/SRSF5/SRSF7/KHDRBS2	5
GO:0060211	0.02548688	0.02071488	TOB1/AGO2	2
GO:1900153	0.02548688	0.02071488	TOB1/AGO2	2
GO:1903311	0.04144067	0.03368158	RBM24/SRSF7/KHDRBS2/TOB1/AGO2	5
GO:1903312	0.04240294	0.03446368	RBM24/SRSF7/TOB1	3
GO:1900151	0.04333233	0.03521906	TOB1/AGO2	2
GO:0048024	0.04333233	0.03521906	RBM24/SRSF7/KHDRBS2	3
GO:0006417	0.04333233	0.03521906	RPL38/RBM24/EIF1B/TOB1/AGO2	5
GO:1903019	0.04524991	0.03677760	PTX3/AGO2	2
GO:0008380	0.04944125	0.04018418	RBM24/SNRNP40/SRSF5/SRSF7/KHDRBS2	5
GO:0000289	0.04944125	0.04018418	TOB1/AGO2	2
GO:0022613	0.04944125	0.04018418	RPL38/SRSF5/NOA1/RRP9/AGO2	5
GO:0050684	0.04944125	0.04018418	RBM24/SRSF7/KHDRBS2	3
GO:0043484	0.11114991	0.09033891	RBM24/SRSF7/KHDRBS2	3
GO:0000381	0.11114991	0.09033891	RBM24/KHDRBS2	2
GO:0061157	0.11114991	0.09033891	RBM24/TOB1/AGO2	3
GO:0050779	0.11114991	0.09033891	RBM24/TOB1/AGO2	3

```
#go_results_miR_655_3p_predicted@result |>
#  kable(format = "html") |>
#  kable_styling("striped")

# Visualización del análisis funcional
barplot(go_results_miR_655_3p_predicted, showCategory = 10, title = "Top 10 Procesos Bioló
```



5.3.2.5.3 Combined

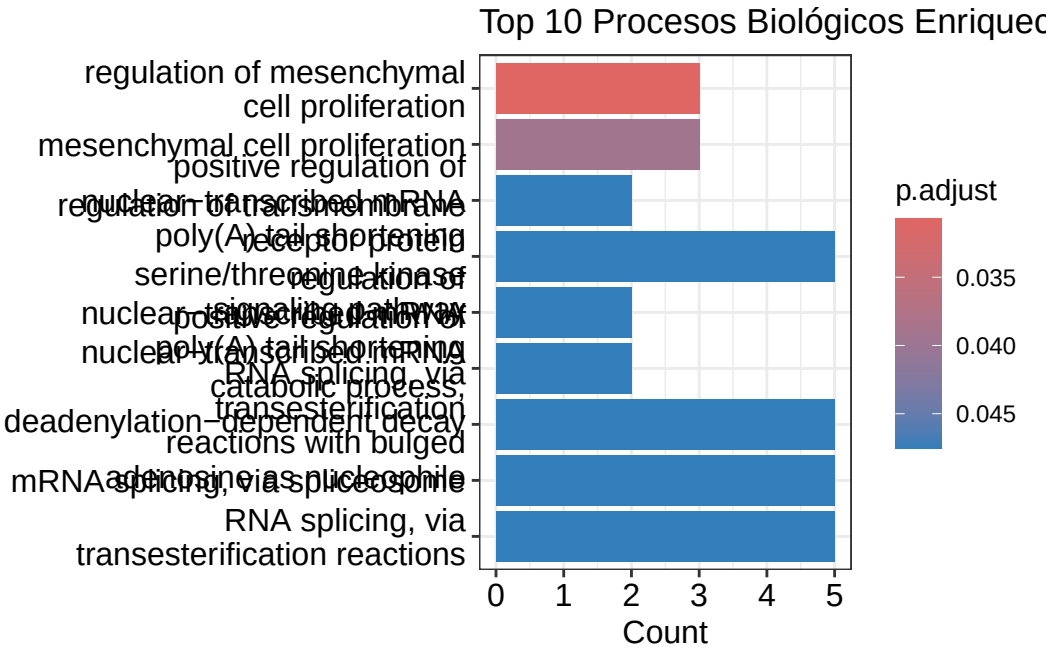
```
# Gene Ontology (GO) para todos los genes combinados
gene_symbols_miR_655_3p_combined <- combined_targets_miR_655_3p$target_symbol %>% unique()
go_results_miR_655_3p_combined <- enrichGO(
  gene = gene_symbols_miR_655_3p_combined,
  OrgDb = org.Hs.eg.db,
  keyType = "SYMBOL",
  ont = "BP",  # Biological Processes
  pAdjustMethod = "BH",
  qvalueCutoff = 0.05, # Relacionada con el control de la tasa de falsos descubrimientos (
  pvalueCutoff = 0.05
)
head(go_results_miR_655_3p_combined@result,20)
```

```
                   ID
GO:0010464 GO:0010464
GO:0010463 GO:0010463
GO:0060213 GO:0060213
GO:0090092 GO:0090092
GO:0060211 GO:0060211
GO:1900153 GO:1900153
GO:0000377 GO:0000377
GO:0000398 GO:0000398
GO:0000375 GO:0000375
GO:0042474 GO:0042474
GO:0048704 GO:0048704
GO:0060395 GO:0060395
GO:0141091 GO:0141091
GO:0043583 GO:0043583
GO:1903312 GO:1903312
GO:1903311 GO:1903311
GO:0002053 GO:0002053
GO:1900151 GO:1900151
GO:0007178 GO:0007178
GO:0048024 GO:0048024

GO:0010464                                                    regulation of mesenchymal cell
GO:0010463                                                                mesenchymal cell
GO:0060213                              positive regulation of nuclear-transcribed mRNA poly(A) t
GO:0090092           regulation of transmembrane receptor protein serine/threonine kinase sig
GO:0060211                                      regulation of nuclear-transcribed mRNA poly(A) t
GO:1900153 positive regulation of nuclear-transcribed mRNA catabolic process, deadenylation-d
GO:0000377              RNA splicing, via transesterification reactions with bulged adenosine
GO:0000398                                                                mRNA splicing, v
GO:0000375                                                  RNA splicing, via transesterifica
GO:0042474                                                                      middle ear
GO:0048704                                                        embryonic skeletal syster
GO:0060395                                                              SMAD protein signa
GO:0141091                              transforming growth factor beta receptor superfamily sig
GO:0043583                                                                              e
GO:1903312                                                    negative regulation of mRNA met
GO:1903311                                                            regulation of mRNA met
GO:0002053                                                positive regulation of mesenchymal cell
GO:1900151          regulation of nuclear-transcribed mRNA catabolic process, deadenylation-d
GO:0007178                             transmembrane receptor protein serine/threonine kinase sig
GO:0048024                                                    regulation of mRNA splicing, v
           GeneRatio   BgRatio RichFactor FoldEnrichment    zScore       pvalue
```

GO:0010464	3/37	33/18888	0.09090909	46.407862	11.566179	3.625480e-05
GO:0010463	3/37	45/18888	0.06666667	34.032432	9.828493	9.277778e-05
GO:0060213	2/37	13/18888	0.15384615	78.536383	12.389360	2.873078e-04
GO:0090092	5/37	317/18888	0.01577287	8.051837	5.609568	3.619361e-04
GO:0060211	2/37	15/18888	0.13333333	68.064865	11.511591	3.858067e-04
GO:1900153	2/37	15/18888	0.13333333	68.064865	11.511591	3.858067e-04
GO:0000377	5/37	337/18888	0.01483680	7.573983	5.394796	4.786548e-04
GO:0000398	5/37	337/18888	0.01483680	7.573983	5.394796	4.786548e-04
GO:0000375	5/37	341/18888	0.01466276	7.485139	5.353956	5.050573e-04
GO:0042474	2/37	21/18888	0.09523810	48.617761	9.672580	7.659213e-04
GO:0048704	3/37	94/18888	0.03191489	16.292122	6.584719	8.204433e-04
GO:0060395	3/37	95/18888	0.03157895	16.120626	6.545588	8.460581e-04
GO:0141091	5/37	388/18888	0.01288660	6.578434	4.918781	9.045213e-04
GO:0043583	4/37	223/18888	0.01793722	9.156708	5.428358	9.202028e-04
GO:1903312	3/37	100/18888	0.03000000	15.314595	6.358489	9.817631e-04
GO:1903311	5/37	404/18888	0.01237624	6.317902	4.786832	1.083398e-03
GO:0002053	2/37	26/18888	0.07692308	39.268191	8.650593	1.178068e-03
GO:1900151	2/37	26/18888	0.07692308	39.268191	8.650593	1.178068e-03
GO:0007178	5/37	417/18888	0.01199041	6.120941	4.684764	1.247143e-03
GO:0048024	3/37	112/18888	0.02678571	13.673745	5.959744	1.361675e-03

	p.adjust	qvalue	geneID	Count
GO:0010464	0.03070782	0.02408082	TGFBR2/ZEB1/PRRX1	3
GO:0010463	0.03929139	0.03081199	TGFBR2/ZEB1/PRRX1	3
GO:0060213	0.04753151	0.03727382	TOB1/AGO2	2
GO:0090092	0.04753151	0.03727382	TGFBR2/ZEB1/MAGI2/ING2/TOB1	5
GO:0060211	0.04753151	0.03727382	TOB1/AGO2	2
GO:1900153	0.04753151	0.03727382	TOB1/AGO2	2
GO:0000377	0.04753151	0.03727382	RBM24/SNRNP40/SRSF5/SRSF7/KHDRBS2	5
GO:0000398	0.04753151	0.03727382	RBM24/SNRNP40/SRSF5/SRSF7/KHDRBS2	5
GO:0000375	0.04753151	0.03727382	RBM24/SNRNP40/SRSF5/SRSF7/KHDRBS2	5
GO:0042474	0.05512798	0.04323091	PRRX1/RPL38	2
GO:0048704	0.05512798	0.04323091	TGFBR2/ZEB1/PRRX1	3
GO:0060395	0.05512798	0.04323091	TGFBR2/MAGI2/TOB1	3
GO:0141091	0.05512798	0.04323091	TGFBR2/ZEB1/MAGI2/ING2/TOB1	5
GO:0043583	0.05512798	0.04323091	ZEB1/ADAM10/PRRX1/RPL38	4
GO:1903312	0.05512798	0.04323091	RBM24/SRSF7/TOB1	3
GO:1903311	0.05512798	0.04323091	RBM24/SRSF7/KHDRBS2/TOB1/AGO2	5
GO:0002053	0.05512798	0.04323091	TGFBR2/PRRX1	2
GO:1900151	0.05512798	0.04323091	TOB1/AGO2	2
GO:0007178	0.05512798	0.04323091	TGFBR2/ZEB1/MAGI2/ING2/TOB1	5
GO:0048024	0.05512798	0.04323091	RBM24/SRSF7/KHDRBS2	3

```
#go_results_miR_655_3p_combined@result |>
#  kable(format = "html") |>
#  kable_styling("striped")

# Visualización del análisis funcional
barplot(go_results_miR_655_3p_combined, showCategory = 10, title = "Top 10 Procesos Biológ
```



5.3.2.6 6: Exportar resultados

```
write.csv(filter_val_hsa_miR_655_3p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA
write.csv(filter_pre_hsa_miR_655_3p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA
write.csv(combined_targets_miR_655_3p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/miR
write.csv(summary_targets_miR_655_3p, "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRN
```

5.3.3 Both hsa-miR-655-3p & hsa-miR-520d-5p

5.3.3.1 1. Validated

```
list_miRNAs <- c("hsa-miR-655-3p","hsa-miR-520d-5p")
```

```r
validated_targets_both <- lapply(list_miRNAs, function(miRNA){
  get_multimir(org = "hsa",
               mirna = list_miRNAs,
               table = "validated",
               add.link = TRUE,
               use.tibble = TRUE,
               summary = TRUE
               )@data # Extract the data
}) %>%
  bind_rows() %>%  # Relate
  distinct()  # Combine and eliminate duplicates
```

```
Searching mirecords ...
Searching mirtarbase ...
Searching tarbase ...

Some of the links to external databases may be broken due to outdated identifiers in these da

Searching mirecords ...
Searching mirtarbase ...
Searching tarbase ...

Some of the links to external databases may be broken due to outdated identifiers in these da
```

5.3.3.2 2. Predicted

```r
predicted_targets_both <- lapply(list_miRNAs, function(miRNA){
  get_multimir(org = "hsa",
               mirna = list_miRNAs,
               table = "predicted",
               predicted.cutoff = 10,
               predicted.cutoff.type = "p",
               predicted.site = "conserved",
               add.link = TRUE,
               use.tibble = TRUE,
               summary = TRUE
               )@data
}) %>%
  bind_rows() %>%
  distinct()  # Combine and eliminate duplicates
```

```
Searching diana_microt ...
Searching elmmo ...
Searching microcosm ...
Searching miranda ...
Searching mirdb ...
Searching pictar ...
Searching pita ...
Searching targetscan ...

Some of the links to external databases may be broken due to outdated identifiers in these da

Searching diana_microt ...
Searching elmmo ...
Searching microcosm ...
Searching miranda ...
Searching mirdb ...
Searching pictar ...
Searching pita ...
Searching targetscan ...

Some of the links to external databases may be broken due to outdated identifiers in these da
```

5.3.3.3 3. Filter and combined

```r
combined_targets <- validated_targets_both %>%
  mutate(type = "validated") %>%
  bind_rows(
    predicted_targets_both %>% mutate(type = "predicted") # Agregar etiquetas de origen
  ) %>%
  distinct()
```

5.3.3.4 4. Resumir targets por miRNA

```r
summary_targets <- combined_targets %>%
  group_by(mature_mirna_id, type) %>%
  summarise(
    num_targets = n(),
    top_targets = paste0(unique(target_symbol)[1:10], collapse = ", ")
  )
```

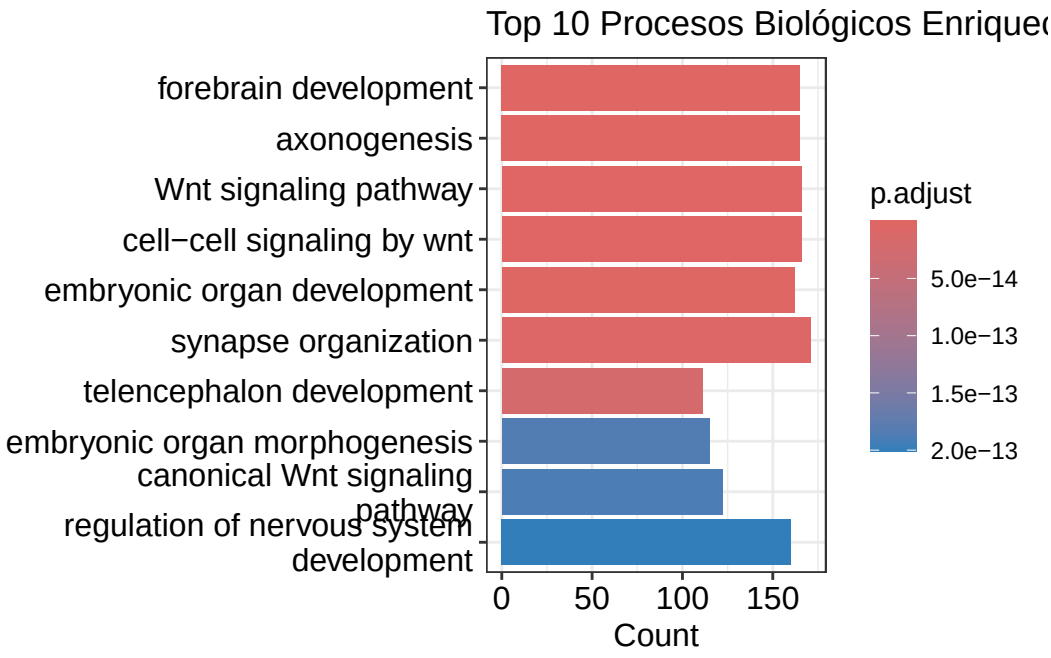
```
`summarise()` has grouped output by 'mature_mirna_id'. You can override using
the `.groups` argument.
```

```
print(summary_targets)
```

```
# A tibble: 4 x 4
# Groups:   mature_mirna_id [2]
  mature_mirna_id type      num_targets top_targets
  <chr>           <chr>           <int> <chr>
1 hsa-miR-520d-5p predicted        5202 TEAD1, ELAVL2, CPEB3, ATAD2B, CPEB2, TA~
2 hsa-miR-520d-5p validated         470 NCOA3, CD164, LAMC1, MYO10, PDCD4, FBLN~
3 hsa-miR-655-3p  predicted        3269 INO80D, CD47, POU2F1, ZNF521, CLCF1, FA~
4 hsa-miR-655-3p  validated         319 TGFBR2, CAPRIN2, TRIM2, ZEB1, HIPK3, JU~
```

5.3.3.5 5: Análisis funcional

```
# Gene Ontology (GO) para todos los genes combinados
gene_symbols <- combined_targets$target_symbol %>% unique()
go_results <- enrichGO(
  gene = gene_symbols,
  OrgDb = org.Hs.eg.db,
  keyType = "SYMBOL",
  ont = "BP",  # Biological Processes
  pAdjustMethod = "BH",
  pvalueCutoff = 0.05
)

# Visualización del análisis funcional
barplot(go_results, showCategory = 10, title = "Top 10 Procesos Biológicos Enriquecidos")
```



5.4 Search in PubMed with entrez

5.4.1 hsa-miR-520d-5p

5.4.1.1 Validated

```
genes_hsa_miR_520d_5p_validated_mirtarbase <- unique(filter_val_hsa_miR_520d_5p$target_sym

# Archivo para guardar los resultados
output_file <- "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_miR_520d_

# Abrir un archivo para escribir
file_conn <- file(output_file, open = "w")

# Bucle para consultar cada gen
for (gene in genes_hsa_miR_520d_5p_validated_mirtarbase) {
  query <- paste0(gene, " AND metabolism", " AND adipose")

  # Buscar en PubMed
  search_results <- entrez_search(db = "pubmed", term = query, retmax = 10)

  # Si hay resultados, recuperar los abstracts
```

```
  if (length(search_results$ids) > 0) {
    abstracts <- entrez_fetch(db = "pubmed", id = search_results$ids, rettype = "abstract"
    writeLines(paste("###---------->", gene, "<----------###", sep = " "), file_conn)
    writeLines(abstracts, file_conn)
    writeLines("\n", file_conn)
  }
}
# Cerrar el archivo
close(file_conn)

cat("Resultados guardados en:", output_file, "\n")
```

```
Resultados guardados en: /home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_mi
```

5.4.1.2 Predicted

```
genes_hsa_miR_520d_5p_predicted <- unique(filter_pre_hsa_miR_520d_5p$target_symbol)

# Archivo para guardar los resultados
output_file <- "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_miR_520d_

# Abrir un archivo para escribir
file_conn <- file(output_file, open = "w")

# Bucle para consultar cada gen
for (gene in genes_hsa_miR_520d_5p_predicted) {
  query <- paste0(gene, " AND metabolism", " AND diabetes",  " AND adipose")

  # Buscar en PubMed
  search_results <- entrez_search(db = "pubmed", term = query, retmax = 10)

  # Si hay resultados, recuperar los abstracts
  if (length(search_results$ids) > 0) {
    abstracts <- entrez_fetch(db = "pubmed", id = search_results$ids, rettype = "abstract"
    writeLines(paste("###---------->", gene, "<----------###", sep = " "), file_conn)
    writeLines(abstracts, file_conn)
    writeLines("\n", file_conn)
  }
}
# Cerrar el archivo
```

```
close(file_conn)

cat("Resultados guardados en:", output_file, "\n")
```

```
Resultados guardados en: /home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_mi
```

5.4.2 hsa-miR-655-3p

5.4.2.1 Validated

```
genes_hsa_miR_655_3p_validated_mirtarbase <- unique(filter_val_hsa_miR_655_3p$target_symbo

# Archivo para guardar los resultados
output_file <- "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_miR_655_3

# Abrir un archivo para escribir
file_conn <- file(output_file, open = "w")

# Bucle para consultar cada gen
for (gene in genes_hsa_miR_655_3p_validated_mirtarbase) {
  query <- paste0(gene, " AND metabolism", " AND adipose")

  # Buscar en PubMed
  search_results <- entrez_search(db = "pubmed", term = query, retmax = 10)

  # Si hay resultados, recuperar los abstracts
  if (length(search_results$ids) > 0) {
    abstracts <- entrez_fetch(db = "pubmed", id = search_results$ids, rettype = "abstract"
    writeLines(paste("###---------->", gene, "<----------###", sep = " "), file_conn)
    writeLines(abstracts, file_conn)
    writeLines("\n", file_conn)
  }
}
# Cerrar el archivo
close(file_conn)

cat("Resultados guardados en:", output_file, "\n")
```

```
Resultados guardados en: /home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_mi
```

5.4.2.2 Predicted

```
genes_hsa_miR_655_3p_predicted <- unique(filter_pre_hsa_miR_655_3p$target_symbol)

# Archivo para guardar los resultados
output_file <- "/home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_miR_655_3

# Abrir un archivo para escribir
file_conn <- file(output_file, open = "w")

# Bucle para consultar cada gen
for (gene in genes_hsa_miR_655_3p_predicted) {
  query <- paste0(gene, " AND metabolism", " AND diabetes", " AND adipose")

  # Buscar en PubMed
  search_results <- entrez_search(db = "pubmed", term = query, retmax = 10)

  # Si hay resultados, recuperar los abstracts
  if (length(search_results$ids) > 0) {
    abstracts <- entrez_fetch(db = "pubmed", id = search_results$ids, rettype = "abstract"
    writeLines(paste("###---------->", gene, "<----------###", sep = " "), file_conn)
    writeLines(abstracts, file_conn)
    writeLines("\n", file_conn)
  }
}
# Cerrar el archivo
close(file_conn)

cat("Resultados guardados en:", output_file, "\n")
```

```
Resultados guardados en: /home/joshoacr13/Documentos/TFM/mirna_analysis/miRNA_targets/hsa_mi
```