Untitled

Núria Rivera Brugués

2023-11-29

```
# BiocManager::install() # Actualiza paquetes instalados
if (!require(BiocManager)) install.packages("BiocManager")
installifnot <- function(pkg) {</pre>
    if (!require(pkg, character.only = T)) {
        BiocManager::install(pkg)
   }
installifnot("oligo")
installifnot("limma")
installifnot("Biobase")
installifnot("arrayQualityMetrics")
installifnot("genefilter")
installifnot("annotate")
installifnot("xtable")
installifnot("gplots")
installifnot("GOstats")
installifnot("mogene10sttranscriptcluster.db")
library("BiocManager")
library("affy")
library("oligo")
library("pd.ht.hg.u133.plus.pm")
library("hgu133plus2.db")
library("ggplot2")
library("hgu133plus2.db")
library(org.Hs.eg.db)
library("mogene10sttranscriptcluster.db")
library(knitr)
library(rmdformats)
library("AnnotationDbi")
```

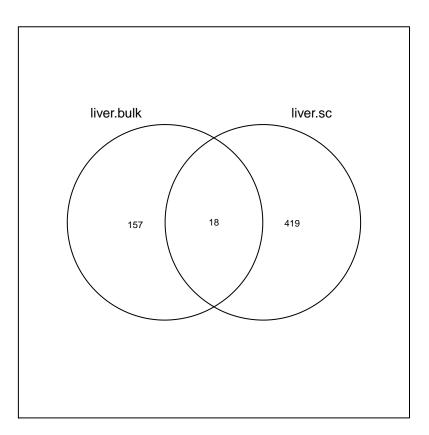
```
library("org.Mm.eg.db")
library(genekitr)
markers kidney.sc <- read.csv2("/home/nrb/Escritorio/perAnna/20 sc kidney/markers.csv")
kidney.sc <- markers_kidney.sc$X</pre>
# write.csv(kidney.sc, '/home/nrb/Escritorio/perAnna/Vennplot/kidney.sc.csv')
markers liver.sc <- read.csv2("/home/nrb/Escritorio/perAnna/18 hepato/markers41.csv")
liver.sc <- markers liver.sc$X</pre>
# write.csv(liver.sc, '/home/nrb/Escritorio/perAnna/Vennplot/liver.sc.csv')
DEG_bulk_kidney <- read.delim("/home/nrb/Documentos/R/kidney/DEG.txt")</pre>
kidney.bulk <- DEG_bulk_kidney$Gene</pre>
# write.csv(kidney.bulk,
# '/home/nrb/Escritorio/perAnna/Vennplot/kidney.bulk.csv')
DEG_bulk_liver <- read.delim("/home/nrb/Documentos/R/liver/DEG_liver.txt")</pre>
liver.bulk <- DEG_bulk_liver$Gene</pre>
# write.csv(liver.bulk, '/home/nrb/Escritorio/perAnna/Vennplot/liver.bulk.csv')
library(biomaRt)
human <- useEnsembl("ensembl", dataset = "hsapiens_gene_ensembl")</pre>
liver.sc <- getBM(attributes = c("hgnc_symbol", "ensembl_gene_id"), filters = "hgnc_symbol",</pre>
    values = liver.sc, mart = human)
write.csv2(liver.sc, "/home/nrb/Escritorio/perAnna/Vennplot/liver.sc2.csv")
kidney.bulk <- getBM(attributes = c("hgnc_symbol", "ensembl_gene_id"), filters = "ensembl_gene_id",
    values = kidney.bulk, mart = human)
write.csv2(kidney.bulk, "/home/nrb/Escritorio/perAnna/Vennplot/kidney.bulk2.csv")
kidney.sc <- mapIds(org.Mm.eg.db, keys = kidney.sc, column = "ENSEMBL", keytype = "SYMBOL")
kidney.sc <- as.data.frame(kidney.sc)</pre>
kidney.sc$symbol <- rownames(kidney.sc)</pre>
kidney.sc$ensembl <- kidney.sc$kidney.sc</pre>
kidney.sc <- kidney.sc[, -1]</pre>
write.csv2(kidney.sc, "/home/nrb/Escritorio/perAnna/Vennplot/kidney.sc2.csv")
liver.bulk <- mapIds(org.Mm.eg.db, keys = liver.bulk, column = "SYMBOL", keytype = "ENSEMBL")</pre>
liver.bulk <- as.data.frame(liver.bulk)</pre>
liver.bulk$ensembl <- rownames(liver.bulk)</pre>
liver.bulk$symbol <- liver.bulk$liver.bulk</pre>
liver.bulk <- liver.bulk[, -1]</pre>
write.csv2(liver.bulk, "/home/nrb/Escritorio/perAnna/Vennplot/liver.bulk2.csv")
library(VennDetail)
library(VennDiagram)
library(venn)
```

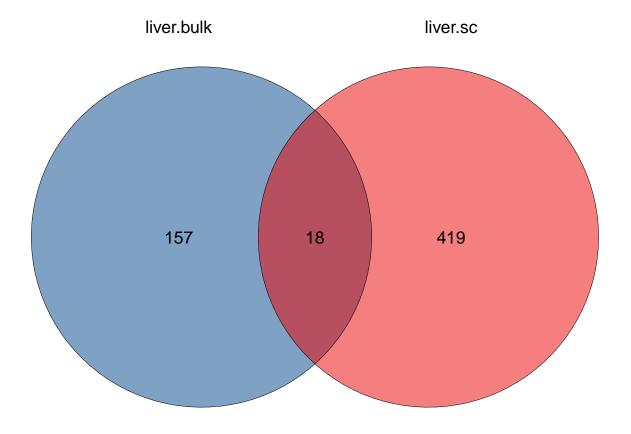
Venn Diagrams

humà symbol

Comparació fetge (bulk-sc)

```
# liver human
ven <- venn(list(liver.bulk = liver.bulk.human, liver.sc = liver.sc.human))</pre>
```





Llistat en comparar bulk-Fetge vs scsn-Fetge

```
library(biomaRt)
mart <- useMart(biomart = "ENSEMBL_MART_ENSEMBL", dataset = "hsapiens_gene_ensembl")
listAttributes(mart)</pre>
```

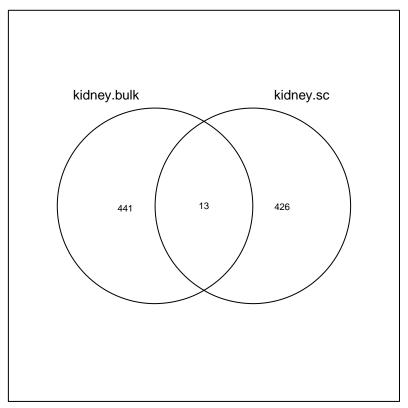
	name	description
1	ensembl_gene_id	Gene stable ID
2	ensembl_gene_id_version	Gene stable ID version
3	ensembl_transcript_id	Transcript stable ID
4	<pre>ensembl_transcript_id_version</pre>	Transcript stable ID version
5	ensembl_peptide_id	Protein stable ID
6	<pre>ensembl_peptide_id_version</pre>	Protein stable ID version
7	ensembl_exon_id	Exon stable ID
8	description	Gene description
9	chromosome_name	Chromosome/scaffold name
10	${\tt start_position}$	Gene start (bp)
11	end_position	Gene end (bp)
12	strand	Strand
13	band	Karyotype band
14	transcript_start	Transcript start (bp)
15	transcript_end	Transcript end (bp)
16	transcription_start_site	Transcription start site (TSS)
17	${\tt transcript_length}$	Transcript length (including UTRs and CDS)
18	${\tt transcript_tsl}$	Transcript support level (TSL)
19	transcript_gencode_basic	GENCODE basic annotation

```
20
                                                             APPRIS annotation
               transcript_appris
21
                                                             Ensembl Canonical
         transcript_is_canonical
                                        RefSeq match transcript (MANE Select)
22
          transcript_mane_select
23 transcript_mane_plus_clinical RefSeq match transcript (MANE Plus Clinical)
24
              external_gene_name
                                                                     Gene name
25
            external_gene_source
                                                           Source of gene name
           page
1 feature_page
2 feature_page
3 feature_page
4 feature_page
5 feature_page
6 feature_page
7 feature_page
8 feature_page
9 feature_page
10 feature_page
11 feature_page
12 feature_page
13 feature_page
14 feature_page
15 feature_page
16 feature_page
17 feature_page
18 feature_page
19 feature_page
20 feature_page
21 feature_page
22 feature_page
23 feature_page
24 feature_page
25 feature_page
 [ reached 'max' / getOption("max.print") -- omitted 3144 rows ]
# Query the database for annotation information
fetge <- getBM(attributes = c("entrezgene_id", "hgnc_symbol", "ensembl_gene_id",</pre>
    "description"), filters = "hgnc_symbol", values = gens_fetge_human, mart = mart)
print(kable(fetge))
```

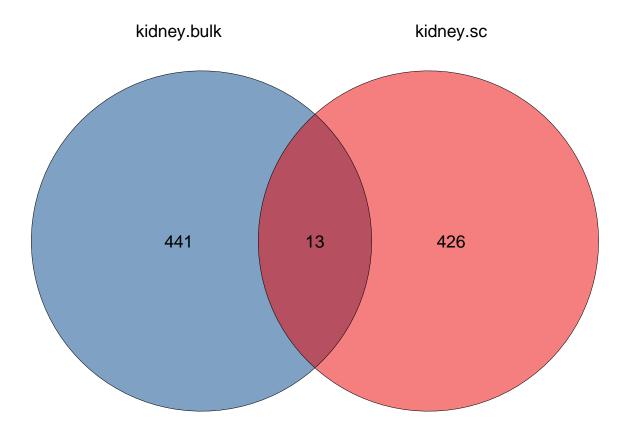
```
| entrezgene_id|hgnc_symbol |ensembl_gene_id |description
    -----|:-----|:-----|
                           |ENSG00000163586 |fatty acid binding protein 1 [Source: HGNC Symbol; Acc: HGNC
          2168 | FABP1
          2644 | GCHFR
                           |ENSG00000137880 |GTP cyclohydrolase I feedback regulator [Source: HGNC Symb
          4508|MT-ATP6
                           |ENSG00000198899 |mitochondrially encoded ATP synthase membrane subunit 6 [
          4509|MT-ATP8
                           |ENSG00000228253 |mitochondrially encoded ATP synthase membrane subunit 8 [
                           |ENSG00000198804 |mitochondrially encoded cytochrome c oxidase I [Source: HG.
          4512|MT-C01
                           |ENSG00000198712 |mitochondrially encoded cytochrome c oxidase II [Source:H
          4513 MT-C02
                           |ENSG00000198938 |mitochondrially encoded cytochrome c oxidase III [Source:
          4514 | MT-CO3
                           |ENSG00000198727 |mitochondrially encoded cytochrome b [Source: HGNC Symbol;
          4519 | MT-CYB
                           |ENSG00000198888 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
          4535 | MT-ND1
                           |ENSG00000198763 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
          4536 | MT-ND2
                           |ENSG00000198840 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
          4537 | MT-ND3
```

Comparació ronyó (bulk-sc)

```
ven <- venn(list(kidney.bulk = kidney.bulk.human, kidney.sc = kidney.sc.human))</pre>
```



```
vendetail <- venndetail(list(kidney.bulk = kidney.bulk.human, kidney.sc = kidney.sc.human))</pre>
res <- result(vendetail)</pre>
df.gens <- res[res$Subset == "Shared", ]</pre>
(gens_ronyo_human <- as.vector(df.gens$Detail))</pre>
 [1] "SYP"
                                        "OGT"
                 "RSRP1"
                             "PROCA1"
                                                    "NKTR"
                                                                "MT-ND6"
 [7] "MT-ND5"
                 "MT-ND4L" "MT-ND2"
                                        "MT-ND1"
                                                    "MT-CYB"
                                                                "CATSPER2"
[13] "ANLN"
library(genekitr)
plotVenn(list(kidney.bulk = kidney.bulk.human, kidney.sc = kidney.sc.human), use_venn = TRUE,
    color = (ggsci::pal_lancet())(2), alpha_degree = 0.5, main_text_size = 4.5, border_thick = 0)
```

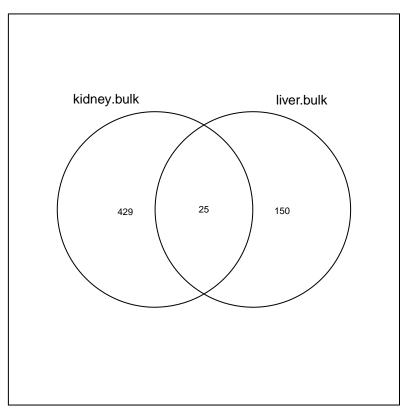


Llistat en comparar bulk-Ronyó vs scsn-Ronyó

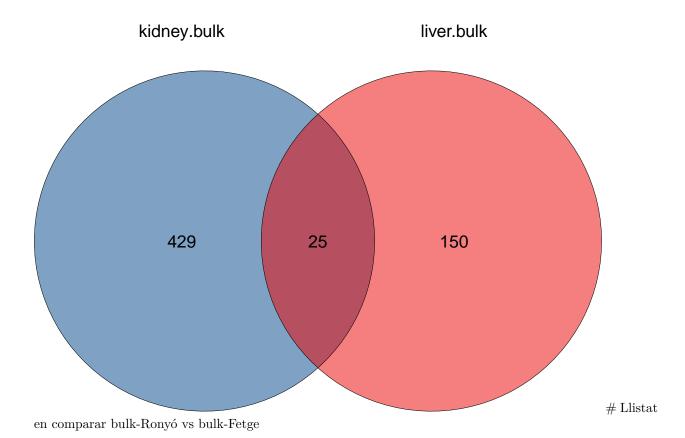
```
| entrezgene_id|hgnc_symbol |ensembl_gene_id |description
          54443 | ANLN
                             |ENSG00000011426 |anillin, actin binding protein [Source: HGNC Symbol; Acc: HG
                             |ENSG00000166762 |cation channel sperm associated 2 [Source: HGNC Symbol; Acc
        117155 | CATSPER2
           4519|MT-CYB
                             |ENSG00000198727 |mitochondrially encoded cytochrome b [Source: HGNC Symbol;
                             |ENSG00000198888 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
           4535 | MT-ND1
          4536 | MT-ND2
                             |ENSG00000198763 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
          4539|MT-ND4L
                             |ENSG00000212907 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
                             |ENSG00000198786 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
           4540 | MT-ND5
                             |ENSG00000198695 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
          4541 | MT-ND6
          4820 | NKTR
                             |ENSG00000114857 | natural killer cell triggering receptor [Source: HGNC Symb
                             |ENSG00000147162 | O-linked N-acetylglucosamine (GlcNAc) transferase [Source
          8473 | OGT
        147011 | PROCA1
                             |ENSG00000167525 |protein interacting with cyclin A1 [Source: HGNC Symbol; Ac
                             |ENSG00000117616 | arginine and serine rich protein 1 [Source: HGNC Symbol; Ac
          57035 | RSRP1
                             |ENSG00000102003 |synaptophysin [Source: HGNC Symbol; Acc: HGNC: 11506]
          6855|SYP
```

Comparació bulks (ronyo-fetge)

```
ven <- venn(list(kidney.bulk = kidney.bulk.human, liver.bulk = liver.bulk.human))</pre>
```



```
vendetail <- venndetail(list(kidney.bulk = kidney.bulk.human, liver.bulk = liver.bulk.human))</pre>
res <- result(vendetail)</pre>
df.gens <- res[res$Subset == "Shared", ]</pre>
(gens_bulks_human <- as.vector(df.gens$Detail))</pre>
 [1] "TTLL3"
                 "SYNGAP1"
                            "MT-ND6"
                                        "MT-ND5"
                                                    "MT-ND4L"
                                                                "MT-ND4"
 [7] "MT-ND3"
                 "MT-ND2"
                            "MT-ND1"
                                        "MT-CYB"
                                                    "MT-C03"
                                                                "MT-C02"
[13] "MT-C01"
                 "MT-ATP8"
                            "MT-ATP6"
                                        "MLXIPL"
                                                    "MAPK8IP3" "LENG8"
[19] "KCNC3"
                 "DOCK3"
                             "COL27A1"
                                        "CDH24"
                                                    "CCNL2"
                                                                "ASIC3"
[25] "ABCC10"
library(genekitr)
plotVenn(list(kidney.bulk = kidney.bulk.human, liver.bulk = liver.bulk.human), use_venn = TRUE,
    color = (ggsci::pal_lancet())(2), alpha_degree = 0.5, main_text_size = 4.5, border_thick = 0)
```

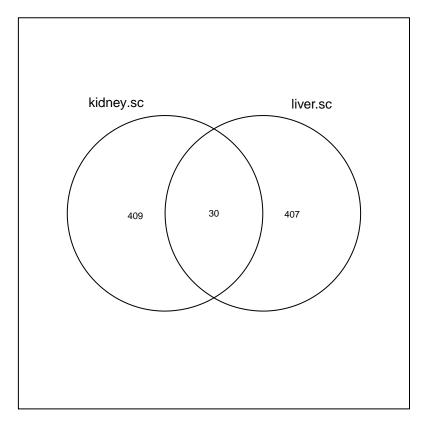


```
| entrezgene_id|hgnc_symbol |ensembl_gene_id |description
      -----:|:-----:
                            |ENSG00000124574 | ATP binding cassette subfamily C member 10 [Source: HGNC S
         89845 | ABCC10
          9311|ASIC3
                            |ENSG00000213199 |acid sensing ion channel subunit 3 [Source: HGNC Symbol; Ac
                            [ENSG00000221978 | cyclin L2 [Source: HGNC Symbol; Acc: HGNC: 20570]
         81669 | CCNL2
         64403 | CDH24
                            |ENSG00000139880 |cadherin 24 [Source: HGNC Symbol; Acc: HGNC: 14265]
         85301 | COL27A1
                            |ENSG00000196739 |collagen type XXVII alpha 1 chain [Source:HGNC Symbol;Acc
          1795 | DOCK3
                            |ENSG00000088538 |dedicator of cytokinesis 3 [Source: HGNC Symbol; Acc: HGNC: 2
                            |ENSG00000131398 |potassium voltage-gated channel subfamily C member 3 [Sou
          3748 | KCNC3
        114823 | LENG8
                            |ENSG00000276458 |leukocyte receptor cluster member 8 [Source: HGNC Symbol; A
                            |ENSG00000274305 |leukocyte receptor cluster member 8 [Source: HGNC Symbol; A
        114823 | LENG8
                            |ENSG00000276681 |leukocyte receptor cluster member 8 [Source: HGNC Symbol; A
        114823 | LENG8
                            |ENSG00000167615 |leukocyte receptor cluster member 8 [Source: HGNC Symbol; A
        114823 | LENG8
                            |ENSG00000138834 |mitogen-activated protein kinase 8 interacting protein 3
         23162|MAPK8IP3
         51085|MLXIPL
                            |ENSG00000009950 |MLX interacting protein like [Source: HGNC Symbol; Acc: HGNC
                            |ENSG00000198899 |mitochondrially encoded ATP synthase membrane subunit 6 [
          4508 | MT-ATP6
                            |ENSG00000228253 |mitochondrially encoded ATP synthase membrane subunit 8 [
          4509 | MT-ATP8
                            |ENSG00000198804 |mitochondrially encoded cytochrome c oxidase I [Source: HG
          4512 | MT-CO1
                            |ENSG00000198712 |mitochondrially encoded cytochrome c oxidase II [Source:H
          4513 | MT-CO2
                            |ENSG00000198938 |mitochondrially encoded cytochrome c oxidase III [Source:
          4514 | MT-CO3
          4519 | MT-CYB
                            |ENSG00000198727 |mitochondrially encoded cytochrome b [Source: HGNC Symbol;
```

```
4535 | MT-ND1
                  |ENSG00000198888 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
                  |ENSG00000198763 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
4536 | MT-ND2
                  |ENSG00000198840 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
4537 | MT-ND3
                  |ENSG00000198886 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
4538 | MT-ND4
                  |ENSG00000212907 |mitochondrially encoded NADH:ubiquinone oxidoreductase co
4539 | MT-ND4L
                  |ENSG00000198786 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
4540 | MT-ND5
                  |ENSG00000198695 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
4541 MT-ND6
                  |ENSG00000227460 |synaptic Ras GTPase activating protein 1 [Source: HGNC Sym
8831 SYNGAP1
8831|SYNGAP1
                  |ENSG00000197283 |synaptic Ras GTPase activating protein 1 [Source: HGNC Sym
                  |ENSG00000214021 |tubulin tyrosine ligase like 3 [Source: HGNC Symbol; Acc: HG
26140|TTLL3
```

Comparació scs (ronyo-fetge)

```
ven <- venn(list(kidney.sc = kidney.sc.human, liver.sc = liver.sc.human))</pre>
```



"CYP2E1"

```
vendetail <- venndetail(list(kidney.sc = kidney.sc.human, liver.sc = liver.sc.human))</pre>
res <- result(vendetail)</pre>
df.gens <- res[res$Subset == "Shared", ]</pre>
(gens.sc_human <- as.vector(df.gens$Detail))</pre>
 [1] "ZEB2"
                "UPP2"
                           "SRC"
                                      "SMC2"
                                                 "SLC27A2" "SDC4"
                                                                       "PTPRD"
 [8] "PRKCB"
                "PDE4D"
                           "PAG1"
                                      "MT1M"
                                                 "MT-ND5"
                                                            "MT-ND4L"
                                                                       "MT-ND2"
                                                 "GRAMD1B" "GPX3"
[15] "MT-ND1"
                "MT-CYB"
                           "MGST1"
                                      "MAGI2"
                                                                       "GNG2"
```

"FGG"

[22] "FUT2"

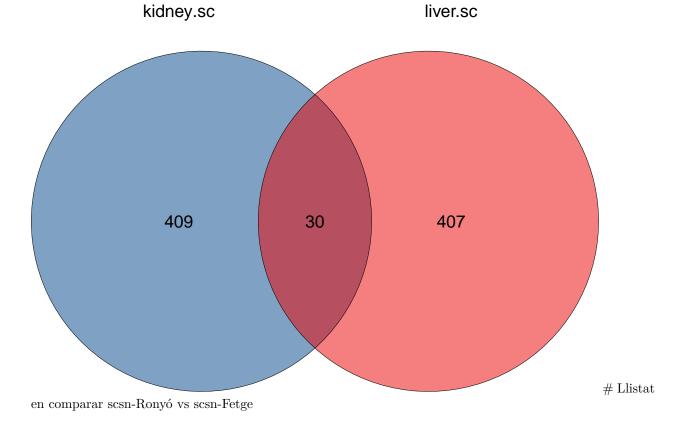
"CD320"

"BNC2"

"ANTXR2"

"CD74"

```
library(genekitr)
plotVenn(list(kidney.sc = kidney.sc.human, liver.sc = liver.sc.human), use_venn = TRUE,
    color = (ggsci::pal_lancet())(2), alpha_degree = 0.5, main_text_size = 4.5, border_thick = 0)
```



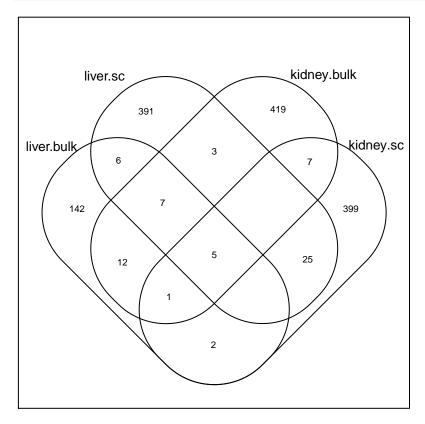
```
scs <- getBM(attributes = c("entrezgene_id", "hgnc_symbol", "ensembl_gene_id", "description"),
    filters = "hgnc_symbol", values = gens.sc_human, mart = mart)
print(kable(scs))</pre>
```

```
| entrezgene id|hgnc symbol |ensembl gene id |description
                             |ENSG00000023839 | ATP binding cassette subfamily C member 2 [Source: HGNC Sys
           1244 | ABCC2
            125 | ADH1B
                             |ENSG00000196616 |alcohol dehydrogenase 1B (class I), beta polypeptide [Sou
                             |ENSG00000163297 | ANTXR cell adhesion molecule 2 [Source: HGNC Symbol; Acc: HG
         118429 | ANTXR2
                             |ENSG00000173068 |basonuclin 2 [Source: HGNC Symbol; Acc: HGNC: 30988]
          54796 | BNC2
                             |ENSG00000167775 | CD320 molecule [Source: HGNC Symbol; Acc: HGNC: 16692]
          51293 | CD320
            972 | CD74
                             |ENSG00000019582 |CD74 molecule [Source:HGNC Symbol;Acc:HGNC:1697]
                             |ENSG00000130649 | cytochrome P450 family 2 subfamily E member 1 [Source: HGN
           1571 | CYP2E1
           2266|FGG
                             |ENSG00000171557 |fibrinogen gamma chain [Source: HGNC Symbol; Acc: HGNC: 3694]
                             |ENSG00000176920 |fucosyltransferase 2 [Source: HGNC Symbol; Acc: HGNC: 4013]
           2524 | FUT2
                             |ENSG00000186469 | G protein subunit gamma 2 [Source: HGNC Symbol; Acc: HGNC: 44
          54331 | GNG2
                             |ENSG00000211445 |glutathione peroxidase 3 [Source: HGNC Symbol; Acc: HGNC: 455
           2878 | GPX3
                             |ENSG00000023171 | GRAM domain containing 1B [Source: HGNC Symbol; Acc: HGNC: 29
          57476 | GRAMD1B
                             |ENSG00000187391 |membrane associated guanylate kinase, WW and PDZ domain c
           9863|MAGI2
           4257 | MGST1
                             |ENSG00000008394 |microsomal glutathione S-transferase 1 [Source: HGNC Symbo
```

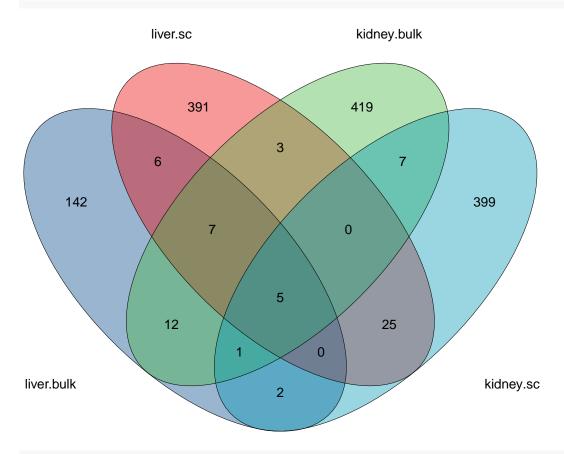
```
4519 | MT-CYB
                    |ENSG00000198727 |mitochondrially encoded cytochrome b [Source: HGNC Symbol;
                    |ENSG00000198888 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
 4535 | MT-ND1
                    |ENSG00000198763 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
 4536 | MT-ND2
                    |ENSG00000212907 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
 4539|MT-ND4L
                    |ENSG00000198786 |mitochondrially encoded NADH:ubiquinone oxidoreductase co.
 4540 | MT-ND5
                    |ENSG00000205364 |metallothionein 1M [Source: HGNC Symbol; Acc: HGNC: 14296]
  4499 | MT1M
                    |ENSG00000076641 |phosphoprotein membrane anchor with glycosphingolipid mic
 55824 | PAG1
                    |ENSG00000113448 |phosphodiesterase 4D [Source: HGNC Symbol; Acc: HGNC: 8783]
 5144 | PDE4D
 5579 | PRKCB
                    |ENSG00000166501 | protein kinase C beta [Source: HGNC Symbol; Acc: HGNC: 9395]
                    |ENSG00000153707 | protein tyrosine phosphatase receptor type D [Source: HGNC
 5789 | PTPRD
 6385|SDC4
                    |ENSG00000124145 |syndecan 4 [Source: HGNC Symbol; Acc: HGNC: 10661]
                    |ENSG00000140284 |solute carrier family 27 member 2 [Source: HGNC Symbol; Acc
 11001|SLC27A2
                    |ENSG00000136824 |structural maintenance of chromosomes 2 [Source: HGNC Symb
 10592|SMC2
                    |ENSG00000291971 |SRC proto-oncogene, non-receptor tyrosine kinase [Source:
 6714|SRC
 6714|SRC
                    |ENSG00000197122 | SRC proto-oncogene, non-receptor tyrosine kinase [Source:
                    |ENSG00000007001 |uridine phosphorylase 2 [Source: HGNC Symbol; Acc: HGNC: 2306
151531 | UPP2
                    |ENSG00000169554 |zinc finger E-box binding homeobox 2 [Source: HGNC Symbol;
  9839 | ZEB2
```

Comparació tot

```
ven <- venn(list(liver.bulk = liver.bulk.human, liver.sc = liver.sc.human, kidney.bulk = kidney.bulk.hum
   kidney.sc = kidney.sc.human))</pre>
```



```
plotVenn(list(liver.bulk = liver.bulk.human, liver.sc = liver.sc.human, kidney.bulk = kidney.bulk.human
    kidney.sc = kidney.sc.human), use_venn = TRUE, color = (ggsci::pal_lancet())(4),
    alpha_degree = 0.4, main_text_size = 3.5, border_thick = 0)
```



plotVenn(list(liver.bulk = liver.bulk.human, liver.sc = liver.sc.human, kidney.bulk = kidney.bulk.human
 kidney.sc = kidney.sc.human), use_venn = FALSE, main_text_size = 12, legend_text_size = 6,
 legend_position = "left")

