

Resources and interaction types in OmniPath

Marton Olbei

10/12/2020

In this tutorial we show you how to query interactions from one of the many resources included in OmniPath, customize the data by interaction types, and further quality controls.

We'll start by importing libraries, first `OmniPathR`, and `dplyr` for data wrangling.

```
library(OmniPathR)
library(dplyr)
```

OmniPath contains many resources we can choose to collate our desired network from. To browse the list available resources call the `get_interaction_resources` function.

```
get_interaction_resources() %>% tibble()
```

```
## # A tibble: 135 x 1
##   .
##   <chr>
## 1 ABS
## 2 ACSN
## 3 ACSN_SignalLink3
## 4 Adhesome
## 5 AlzPathway
## 6 ARACNe-GTex_DoRothEA
## 7 ARN
## 8 Baccin2019
## 9 BEL-Large-Corpus_ProtMapper
## 10 BioGRID
## # ... with 125 more rows
```

OmniPath can serve multiple kinds of interactions, based on the quality of the interactors or the interactions themselves:

- `post_translational` i.e. physical interactions of proteins, protein-protein interactions (or PPIs)
- `transcriptional` i.e. gene regulatory interactions
- `post_transcriptional` i.e. miRNA-mRNA interactions
- `mirna_transcriptional` i.e. transcriptional regulation of miRNA genes

In the following code blocks we are going to query all of them, and show the URLs these queries generate, through which the data is also accessible, through a browser.

First, let's take a look at PPI interactions.

URL: https://omnipathdb.org/interactions?genesymbols=yes&datasets=omnipath,pathwayextra,kinaseextra,ligreextra&organisms=9606&fields=sources,references,curation_effort&license=academic

By default, the query returns data from the `omnipath` dataset, which means literature curated activity flow (directed, signed interactions in most cases, curation effort).

```
interactions_PPI <- import_post_translational_interactions(  
  organism = 9606  
)  
interactions_PPI %>% tibble()
```

```
## # A tibble: 75,524 x 16  
##   source target source_genesymb~ target_genesymb~ is_directed is_stimulation  
##   <chr> <chr> <chr> <chr> <int> <int>  
## 1 P0DP24 P48995 CALM2 TRPC1 1 0  
## 2 Q03135 P48995 CAV1 TRPC1 1 1  
## 3 P14416 P48995 DRD2 TRPC1 1 1  
## 4 Q02790 P48995 FKBP4 TRPC1 1 0  
## 5 Q99750 P48995 MDFI TRPC1 1 0  
## 6 Q14571 P48995 ITPR2 TRPC1 1 1  
## 7 P29966 P48995 MARCKS TRPC1 1 0  
## 8 P48995 Q13255 TRPC1 GRM1 1 0  
## 9 Q13255 P48995 GRM1 TRPC1 1 1  
## 10 Q13586 P48995 STIM1 TRPC1 1 1  
## # ... with 75,514 more rows, and 10 more variables: is_inhibition <int>,  
## # consensus_direction <int>, consensus_stimulation <int>,  
## # consensus_inhibition <int>, dip_url <chr>, sources <chr>, references <chr>,  
## # curation_effort <int>, n_references <int>, n_resources <int>
```

We can use these properties to further specify our queries, e.g.:

```
interactions_curation_effort <- import_post_translational_interactions(  
  organism = 9606  
) %>% filter(curation_effort > 7)  
interactions_curation_effort %>% tibble()
```

```
## # A tibble: 5,566 x 16  
##   source target source_genesymb~ target_genesymb~ is_directed is_stimulation  
##   <chr> <chr> <chr> <chr> <int> <int>  
## 1 Q03135 P48995 CAV1 TRPC1 1 1  
## 2 Q14571 P48995 ITPR2 TRPC1 1 1  
## 3 Q13586 P48995 STIM1 TRPC1 1 1  
## 4 P48995 Q13507 TRPC1 TRPC3 1 1  
## 5 Q13507 P48995 TRPC3 TRPC1 1 1  
## 6 P48995 Q9UBN4 TRPC1 TRPC4 1 1  
## 7 Q9UBN4 P48995 TRPC4 TRPC1 1 1  
## 8 P48995 Q9UL62 TRPC1 TRPC5 1 1  
## 9 Q9UL62 P48995 TRPC5 TRPC1 1 1  
## 10 P48995 Q13563 TRPC1 PKD2 1 1  
## # ... with 5,556 more rows, and 10 more variables: is_inhibition <int>,  
## # consensus_direction <int>, consensus_stimulation <int>,  
## # consensus_inhibition <int>, dip_url <chr>, sources <chr>, references <chr>,  
## # curation_effort <int>, n_references <int>, n_resources <int>
```

The `curation_effort` value we filtered our query on shows the unique database - citation pairs, i.e. how many times was an interaction described in a paper and mentioned in a database.

We can include interactions without explicit literature references as well, by including the extra datasets `pathwayextra`, `kinaseextra`, or `ligreextra`.

To get just one of these extra sets, one can call the specific function for it:

URL https://omnipathdb.org/interactions?genesymbols=yes&datasets=pathwayextra&organisms=9606&fields=sources,references,curation_effort&license=academic

```
interactions_pathwayextra <- import_pathwayextra_interactions(  
  organism = 9606  
)  
interactions_pathwayextra %>% tibble()
```

```
## # A tibble: 41,817 x 16  
##   source target source_genesymb~ target_genesymb~ is_directed is_stimulation  
##   <chr> <chr> <chr> <chr> <int> <int>  
## 1 P48995 Q13255 TRPC1 GRM1 1 0  
## 2 Q13255 P48995 GRM1 TRPC1 1 1  
## 3 P20591 Q9Y210 MX1 TRPC6 1 1  
## 4 O60500 Q9Y210 NPFS1 TRPC6 1 1  
## 5 Q13976 Q9Y210 PRKG1 TRPC6 1 0  
## 6 Q9NP85 Q9Y210 NPFS2 TRPC6 1 1  
## 7 P17612 Q8NER1 PRKACA TRPV1 1 1  
## 8 P12931 Q8NER1 SRC TRPV1 1 1  
## 9 Q96J02 Q9HBA0 ITCH TRPV4 1 1  
## 10 Q9UEF7 Q9NQA5 KL TRPV5 1 1  
## # ... with 41,807 more rows, and 10 more variables: is_inhibition <int>,  
## # consensus_direction <int>, consensus_stimulation <int>,  
## # consensus_inhibition <int>, dip_url <chr>, sources <chr>, references <chr>,  
## # curation_effort <int>, n_references <int>, n_resources <int>
```

To get all PPI interactions call `import_all_interactions`. By default only directed interactions are included, but we can include the `directed = no` flag to get everything.

URL: <https://omnipathdb.org/interactions?genesymbols=yes&fields=sources,references&datasets=omnipath,pathwayextra,kinaseextra,ligreextra&directed=no>

```
all_interactions <- import_all_interactions(  
  organism = 9606,  
  directed = 'no'  
)  
all_interactions %>% tibble()
```

```
## # A tibble: 176,864 x 17  
##   source target source_genesymb~ target_genesymb~ is_directed is_stimulation  
##   <chr> <chr> <chr> <chr> <int> <int>  
## 1 P0DP24 P48995 CALM2 TRPC1 1 0  
## 2 Q03135 P48995 CAV1 TRPC1 1 1  
## 3 P14416 P48995 DRD2 TRPC1 1 1  
## 4 Q02790 P48995 FKBP4 TRPC1 1 0  
## 5 P48995 Q86YM7 TRPC1 HOMER1 0 0  
## 6 Q99750 P48995 MDFI TRPC1 1 0
```

```
## 7 Q14571 P48995 ITPR2          TRPC1          1          1
## 8 P48995 Q14573 TRPC1          ITPR3          0          0
## 9 P29966 P48995 MARCKS        TRPC1          1          0
## 10 P48995 Q13255 TRPC1        GRM1           1          0
## # ... with 176,854 more rows, and 11 more variables: is_inhibition <int>,
## #   consensus_direction <int>, consensus_stimulation <int>,
## #   consensus_inhibition <int>, dip_url <chr>, sources <chr>, references <chr>,
## #   curation_effort <int>, dorothea_level <chr>, n_references <int>,
## #   n_resources <int>
```

The other interaction types have their own built-in functions as well. This query accesses interactions from DoRothEA, from confidence levels A to D, from highest to lowest. It is set to pull out A and B by default, but naturally we can extend it.

URL: https://omnipathdb.org/interactions?genesymbols=yes&fields=sources,references&datasets=dorothea,tf_target&dorothea_levels=A,B,C,D

```
interactions_regulatory <- import_transcriptional_interactions(
  organism = 9606,
  dorothea_levels = c("A","B", "C", "D")
)
interactions_regulatory %>% tibble()
```

```
## # A tibble: 341,504 x 17
##   source target source_genesymb~ target_genesymb~ is_directed is_stimulation
##   <chr> <chr> <chr> <chr> <int> <int>
## 1 Q9H2P0 P33527 ADNP ABCC1 1 0
## 2 Q9H2P0 O95255 ADNP ABCC6 1 0
## 3 Q9H2P0 Q8WTS1 ADNP ABHD5 1 0
## 4 Q9H2P0 Q9ULW3 ADNP ABT1 1 0
## 5 Q9H2P0 Q9BR61 ADNP ACBD6 1 0
## 6 Q9H2P0 Q6ZNF0 ADNP ACP7 1 0
## 7 Q9H2P0 Q9H324 ADNP ADAMTS10 1 0
## 8 Q9H2P0 Q8WXS8 ADNP ADAMTS14 1 0
## 9 Q9H2P0 P51828 ADNP ADCY7 1 0
## 10 Q9H2P0 Q9Y653 ADNP ADGRG1 1 0
## # ... with 341,494 more rows, and 11 more variables: is_inhibition <int>,
## #   consensus_direction <int>, consensus_stimulation <int>,
## #   consensus_inhibition <int>, dip_url <lgl>, sources <chr>, references <chr>,
## #   curation_effort <int>, dorothea_level <chr>, n_references <int>,
## #   n_resources <int>
```

To access `post_transcriptional` and `mirna_transcriptional` interactions, we can utilise their respective functions, or call the corresponding URLs:

- miRNA - mRNA https://omnipathdb.org/interactions?genesymbols=yes&datasets=mirnatarget&organisms=9606&fields=sources,references,curation_effort&license=academic
- TF - miRNA: https://omnipathdb.org/interactions?genesymbols=yes&datasets=tf_mirna&organisms=9606&fields=sources,references,curation_effort&license=academic

```

interactions_post_transcriptional <- import_mirnatarget_interactions(
  organism = 9606
)
interactions_mirna_transcriptional <- import_tf_mirna_interactions(
  organism = 9606
)

```

```
interactions_post_transcriptional %>% tibble()
```

```

## # A tibble: 8,278 x 16
##   source target source_genesymb~ target_genesymb~ is_directed is_stimulation
##   <chr> <chr> <chr> <chr> <int> <int>
## 1 MIMAT~ P01116 hsa-let-7a KRAS 1 0
## 2 MIMAT~ P52926 hsa-let-7a HMGA2 1 0
## 3 MIMAT~ P10415 hsa-let-7a BCL2 1 0
## 4 MIMAT~ P01106 hsa-let-7a MYC 1 0
## 5 MIMAT~ P30304 hsa-let-7a CDC25A 1 0
## 6 MIMAT~ Q00534 hsa-let-7a CDK6 1 0
## 7 MIMAT~ P35240 hsa-let-7a NF2 1 0
## 8 MIMAT~ Q96PU4 hsa-let-7a UHRF2 1 0
## 9 MIMAT~ Q9UHF5 hsa-let-7a IL17B 1 0
## 10 MIMAT~ P49427 hsa-let-7b CDC34 1 0
## # ... with 8,268 more rows, and 10 more variables: is_inhibition <int>,
## # consensus_direction <int>, consensus_stimulation <int>,
## # consensus_inhibition <int>, dip_url <lg1>, sources <chr>, references <chr>,
## # curation_effort <int>, n_references <int>, n_resources <int>

```

```
interactions_mirna_transcriptional %>% tibble()
```

```

## # A tibble: 4,979 x 16
##   source target source_genesymb~ target_genesymb~ is_directed is_stimulation
##   <chr> <chr> <chr> <chr> <int> <int>
## 1 Q9UKV8 MIMAT~ AG02 hsa-miR-155-5p 1 0
## 2 Q9UKV8 MIMAT~ AG02 hsa-miR-155* 1 0
## 3 P35869 MIMAT~ AHR hsa-miR-106b* 1 1
## 4 P35869 MIMAT~ AHR hsa-miR-106b-5p 1 1
## 5 P35869 MIMAT~ AHR hsa-miR-132-5p 1 1
## 6 P35869 MIMAT~ AHR hsa-miR-132 1 1
## 7 P35869 MIMAT~ AHR hsa-miR-212-5p 1 1
## 8 P35869 MIMAT~ AHR hsa-miR-212-3p 1 1
## 9 P35869 MIMAT~ AHR hsa-miR-25 1 1
## 10 P35869 MIMAT~ AHR hsa-miR-25* 1 1
## # ... with 4,969 more rows, and 10 more variables: is_inhibition <int>,
## # consensus_direction <int>, consensus_stimulation <int>,
## # consensus_inhibition <int>, dip_url <lg1>, sources <chr>, references <chr>,
## # curation_effort <int>, n_references <int>, n_resources <int>

```

In this tutorial we learned:

- The various interaction types in OmniPath
- The differences between the encoded interaction types
- How to access and query these interaction types