

# Towards a reproducible interactome: semantic-based detection of redundancies to unify protein-protein interaction databases

Marc Melkonian, Camille Juigné, Olivier Dameron,  
Gwenaël Rabut\* and Emmanuelle Becker\*

**SUPPLEMENTARY DATA (TABLES AND FIGURES)**

Table S1: Origin of the implicit redundancies in human. The database with the less precise PSI-MI term (the ancestor) is indicated in the first column. The database with the more precise PSI-MI term (the descendant) is indicated in the second column. Number of implicit redundancies in column 3, Percentage of implicit redundancy in column 4.

| <b>Database with<br/>the ancestor PSI-MI term</b> | <b>Database with<br/>the descendant PSI-MI term</b> | <b>Implicit redundancies<br/>Counts</b> | <b>(%)</b> |
|---|---|---|------------|
| BioGRID   | IntAct  | 90368                                   | (84.95%)   |
| BioGRID   | BioGRID   | 9030                                    | (8.49%)    |
| HPRD  | IntAct  | 5309                                    | (4.99%)    |
| BioGRID   | DIP   | 1049                                    | (0.99%)    |
| IntAct  | IntAct  | 440                                     | (0.41%)    |
| IntAct  | BioGRID   | 73                                      | (0.07%)    |
| DIP   | BioGRID   | 43                                      | (0.04%)    |
| DIP   | IntAct  | 31                                      | (0.03%)    |
| DIP   | DIP   | 18                                      | (0.02%)    |
| IntAct  | DIP   | 12                                      | (0.01%)    |
| Intra-database redundancy                         |   | 9488                                    | (8.92%)    |
| Inter-database redundancy                         |   | 96885                                   | (91.08%)   |

Table S2: Origin of the implicit redundancies in yeast. The database with the less precise PSI-MI term (the ancestor) is indicated in the first column. The database with the more precise PSI-MI term (the descendant) is indicated in the second column. Number of implicit redundancies in column 3, Percentage of implicit redundancy in column 4.

| <b>Database with<br/>the ancestor PSI-MI term</b> | <b>Database with<br/>the descendant PSI-MI term</b> | <b>Implicit redundancies<br/>Counts</b> | <b>Implicit redundancies<br/>(%)</b> |
|---|---|---|--------------------------------------|
| BioGRID   | IntAct  | 33426                                   | (71.52%)                             |
| BioGRID   | DIP   | 10091                                   | (21.59%)                             |
| BioGRID   | BioGRID   | 2075                                    | (4.44%)                              |
| DIP   | IntAct  | 426                                     | (0.91%)                              |
| DIP   | BioGRID   | 383                                     | (0.82%)                              |
| IntAct  | IntAct  | 244                                     | (0.52%)                              |
| IntAct  | BioGRID   | 81                                      | (0.17%)                              |
| IntAct  | DIP   | 9                                       | (0.02%)                              |
| DIP   | DIP   | 4                                       | (0.01%)                              |
| Intra-database redundancy                         |   | 2323                                    | (4.97%)                              |
| Inter-database redundancy                         |   | 44416                                   | (95.03%)                             |

Table S3: Couple of (ancestor, child) IDMs producing implicit redundancies in human. Column 1: PSI-MI identifier of the ancestor IDM. Column 2: Name of the ancestor IDM. Column 3: PSI-MI identifier of the child IDM. Column 4: Name of the child IDM. Column 5: Number of such implicit redundancies.

| Redundant ancestor |                                       | Redundant child |  | Counts |
|--------------------|---------------------------------------|-----------------|--|--------|
| PSI-MI code        | PSI-MI term                           | PSI-MI code     | PSI-MI term                                |        |
| MI:0004            | Affinity chromatography technology    | MI:0007         | Anti tag coimmunoprecipitation             | 25333  |
| MI:0018            | Two hybrid                            | MI:0397         | Two hybrid array                           | 16113  |
| MI:0018            | Two hybrid                            | MI:1112         | Two hybrid prey pooling approach           | 11238  |
| MI:0018            | Two hybrid                            | MI:1356         | Validated two hybrid                       | 11123  |
| MI:0018            | Two hybrid                            | MI:0398         | Two hybrid pooling approach                | 10713  |
| MI:0004            | Affinity chromatography technology    | MI:0096         | Pull down                                  | 9896   |
| MI:0004            | Affinity chromatography technology    | MI:0006         | Anti bait coimmunoprecipitation            | 6617   |
| MI:0004            | Affinity chromatography technology    | MI:0676         | Tandem affinity purification               | 5968   |
| MI:1313            | Proximity labelling technology        | MI:1314         | Proximity-dependent biotin identification  | 5339   |
| MI:0401            | Biochemical                           | MI:0004         | Affinity chromatography technology         | 935    |
| MI:0415            | Enzymatic study                       | MI:0424         | Protein kinase assay                       | 736    |
| MI:0004            | Affinity chromatography technology    | MI:0019         | Coimmunoprecipitation                      | 734    |
| MI:0401            | Biochemical                           | MI:0096         | Pull down                                  | 467    |
| MI:0090            | Protein complementation assay         | MI:0112         | Ubiquitin reconstruction                   | 192    |
| MI:0004            | Affinity chromatography technology    | MI:0729         | Luminescence based mammalian [...]         | 111    |
| MI:0401            | Biochemical                           | MI:0006         | Anti bait coimmunoprecipitation            | 110    |
| MI:0401            | Biochemical                           | MI:0007         | Anti tag coimmunoprecipitation             | 77     |
| MI:0020            | Transmission electron microscopy      | MI:0410         | Electron tomography                        | 74     |
| MI:0401            | Biochemical                           | MI:0415         | Enzymatic study                            | 60     |
| MI:0401            | Biochemical                           | MI:0071         | Molecular sieving                          | 45     |
| MI:0428            | Imaging technique                     | MI:0663         | Confocal microscopy                        | 37     |
| MI:0019            | Coimmunoprecipitation                 | MI:0006         | Anti bait coimmunoprecipitation            | 34     |
| MI:0401            | Biochemical                           | MI:0019         | Coimmunoprecipitation                      | 30     |
| MI:0019            | Coimmunoprecipitation                 | MI:0007         | Anti tag coimmunoprecipitation             | 27     |
| MI:0428            | Imaging technique                     | MI:0416         | Fluorescence microscopy                    | 27     |
| MI:0415            | Enzymatic study                       | MI:0435         | Protease assay                             | 26     |
| MI:0091            | Chromatography technology             | MI:0004         | Affinity chromatography technology         | 20     |
| MI:0018            | Two hybrid                            | MI:0728         | Gal4 vp16 complementation                  | 18     |
| MI:0401            | Biochemical                           | MI:0676         | Tandem affinity purification               | 17     |
| MI:0004            | Affinity chromatography technology    | MI:0402         | Chromatin immunoprecipitation assay        | 16     |
| MI:0401            | Biochemical                           | MI:0029         | Cosedimentation through density gradient   | 16     |
| MI:0401            | Biochemical                           | MI:0047         | Far western blotting                       | 16     |
| MI:0415            | Enzymatic study                       | MI:0406         | Deacetylase assay                          | 13     |
| MI:0090            | Protein complementation assay         | MI:0018         | Two hybrid                                 | 10     |
| MI:0401            | Biochemical                           | MI:0030         | Cross-linking study                        | 10     |
| MI:0401            | Biochemical                           | MI:0091         | Chromatography technology                  | 10     |
| MI:0415            | Enzymatic study                       | MI:0434         | Phosphatase assay                          | 10     |
| MI:0415            | Enzymatic study                       | MI:0515         | Methyltransferase assay                    | 10     |
| MI:0067            | Light scattering                      | MI:0826         | X ray scattering                           | 9      |
| MI:0018            | Two hybrid                            | MI:0437         | Protein three hybrid                       | 8      |
| MI:0415            | Enzymatic study                       | MI:0997         | Ubiquitinase assay                         | 8      |
| MI:0892            | Solid phase assay                     | MI:0411         | Enzyme linked immunosorbent assay          | 8      |
| MI:0090            | Protein complementation assay         | MI:0231         | Mammalian protein protein interaction trap | 7      |
| MI:0091            | Chromatography technology             | MI:0096         | Pull down                                  | 7      |
| MI:0091            | Chromatography technology             | MI:0006         | Anti bait coimmunoprecipitation            | 6      |
| MI:0091            | Chromatography technology             | MI:0007         | Anti tag coimmunoprecipitation             | 6      |
| MI:0401            | Biochemical                           | MI:0226         | Ion exchange chromatography                | 6      |
| MI:0004            | Affinity chromatography technology    | MI:0858         | Immunodepleted coimmunoprecipitation       | 5      |
| MI:0232            | Transcriptional complementation assay | MI:0018         | Two hybrid                                 | 4      |
| MI:0401            | Biochemical                           | MI:0027         | Cosedimentation                            | 4      |
| MI:0415            | Enzymatic study                       | MI:0516         | Methyltransferase radiometric assay        | 4      |
| MI:0415            | Enzymatic study                       | MI:0889         | Acetylase assay                            | 4      |
| MI:0051            | Fluorescence technology               | MI:0055         | Fluorescent resonance energy transfer      | 3      |
| MI:0067            | Light scattering                      | MI:0038         | Dynamic light scattering                   | 3      |
| MI:0091            | Chromatography technology             | MI:0071         | Molecular sieving                          | 3      |
| MI:0400            | Affinity technology                   | MI:0007         | Anti tag coimmunoprecipitation             | 3      |

End of Table S3:

| Redundant ancestor |                                    | Redundant child |   | Counts |
|--------------------|------------------------------------|-----------------|---|--------|
| PSI-MI code        | PSI-MI term                        | PSI-MI code     | PSI-MI term                                       |        |
| MI:0401            | Biochemical                        | MI:0434         | Phosphatase assay                                 | 3      |
| MI:0415            | Enzymatic study                    | MI:0512         | Zymography  | 3      |
| MI:0415            | Enzymatic study                    | MI:0870         | Demethylase assay                                 | 3      |
| MI:0892            | Solid phase assay                  | MI:0047         | Far western blotting                              | 3      |
| MI:0018            | Two hybrid                         | MI:0727         | Lexa b52 complementation                          | 2      |
| MI:0090            | Protein complementation assay      | MI:0809         | Bimolecular fluorescence complementation          | 2      |
| MI:0400            | Affinity technology                | MI:0004         | Affinity chromatography technology                | 2      |
| MI:0400            | Affinity technology                | MI:0096         | Pull down   | 2      |
| MI:0401            | Biochemical                        | MI:0028         | Cosedimentation in solution                       | 2      |
| MI:0401            | Biochemical                        | MI:0404         | Comigration in non denaturing gel electrophoresis | 2      |
| MI:0401            | Biochemical                        | MI:0411         | Enzyme linked immunosorbent assay                 | 2      |
| MI:0515            | Methyltransferase assay            | MI:0516         | Methyltransferase radiometric assay               | 2      |
| MI:0004            | Affinity chromatography technology | MI:0963         | Interactome parallel affinity capture             | 1      |
| MI:0018            | Two hybrid                         | MI:0399         | Two hybrid fragment pooling approach              | 1      |
| MI:0051            | Fluorescence technology            | MI:0054         | Fluorescence-activated cell sorting               | 1      |
| MI:0051            | Fluorescence technology            | MI:0420         | Kinase homogeneous time resolved fluorescence     | 1      |
| MI:0400            | Affinity technology                | MI:0006         | Anti bait coimmunoprecipitation                   | 1      |
| MI:0400            | Affinity technology                | MI:0405         | Competition binding                               | 1      |
| MI:0400            | Affinity technology                | MI:0411         | Enzyme linked immunosorbent assay                 | 1      |
| MI:0400            | Affinity technology                | MI:0440         | Saturation binding                                | 1      |
| MI:0401            | Biochemical                        | MI:0049         | Filter binding                                    | 1      |
| MI:0401            | Biochemical                        | MI:0405         | Competition binding                               | 1      |
| MI:0401            | Biochemical                        | MI:0424         | Protein kinase assay                              | 1      |
| MI:0401            | Biochemical                        | MI:0440         | Saturation binding                                | 1      |
| MI:0415            | Enzymatic study                    | MI:0419         | Gtpase assay                                      | 1      |
| MI:0415            | Enzymatic study                    | MI:0420         | Kinase homogeneous time resolved fluorescence     | 1      |
| MI:0415            | Enzymatic study                    | MI:0423         | In-gel kinase assay                               | 1      |
| MI:0415            | Enzymatic study                    | MI:1010         | Neddylase assay                                   | 1      |
| MI:0417            | Footprinting                       | MI:0605         | Enzymatic footprinting                            | 1      |
| MI:0424            | Protein kinase assay               | MI:0420         | Kinase homogeneous time resolved fluorescence     | 1      |
| MI:0990            | Cleavage assay                     | MI:0435         | Protease assay                                    | 1      |

Table S4: Couple of (ancestor, child) IDMs producing implicit redundancies in yeast. Column 1: PSI-MI identifier of the ancestor IDM. Column 2: Name of the ancestor IDM. Column 3: PSI-MI identifier of the child IDM. Column 4: Name of the child IDM. Column 5: Number of such implicit redundancies.

| Redundant ancestor |                                    | Redundant child |   | Counts |
|--------------------|------------------------------------|-----------------|---|--------|
| PSI-MI code        | PSI-MI term                        | PSI-MI code     | PSI-MI term                                       |        |
| MI:0004            | Affinity chromatography technology | MI:0676         | Tandem affinity purification                      | 21353  |
| MI:0004            | Affinity chromatography technology | MI:0007         | Anti tag coimmunoprecipitation                    | 11383  |
| MI:0004            | Affinity chromatography technology | MI:0096         | Pull down   | 3322   |
| MI:0090            | Protein complementation assay      | MI:0112         | Ubiquitin reconstruction                          | 2726   |
| MI:0090            | Protein complementation assay      | MI:0111         | Dihydrofolate reductase reconstruction            | 2645   |
| MI:0018            | Two hybrid                         | MI:0398         | Two hybrid pooling approach                       | 1302   |
| MI:0018            | Two hybrid                         | MI:0397         | Two hybrid array                                  | 858    |
| MI:0004            | Affinity chromatography technology | MI:0019         | Coimmunoprecipitation                             | 728    |
| MI:0401            | Biochemical                        | MI:0004         | Affinity chromatography technology                | 687    |
| MI:0004            | Affinity chromatography technology | MI:0006         | Anti bait coimmunoprecipitation                   | 377    |
| MI:0401            | Biochemical                        | MI:0096         | Pull down   | 250    |
| MI:0415            | Enzymatic study                    | MI:0424         | Protein kinase assay                              | 215    |
| MI:0019            | Coimmunoprecipitation              | MI:0007         | Anti tag coimmunoprecipitation                    | 169    |
| MI:0090            | Protein complementation assay      | MI:0018         | Two hybrid  | 112    |
| MI:0401            | Biochemical                        | MI:0007         | Anti tag coimmunoprecipitation                    | 83     |
| MI:0401            | Biochemical                        | MI:0071         | Molecular sieving                                 | 79     |
| MI:0401            | Biochemical                        | MI:0019         | Coimmunoprecipitation                             | 69     |
| MI:0401            | Biochemical                        | MI:0006         | Anti bait coimmunoprecipitation                   | 45     |
| MI:0401            | Biochemical                        | MI:0415         | Enzymatic study                                   | 39     |
| MI:0019            | Coimmunoprecipitation              | MI:0006         | Anti bait coimmunoprecipitation                   | 35     |
| MI:0401            | Biochemical                        | MI:0676         | Tandem affinity purification                      | 33     |
| MI:0020            | Transmission electron microscopy   | MI:0410         | Electron tomography                               | 22     |
| MI:0091            | Chromatography technology          | MI:0004         | Affinity chromatography technology                | 19     |
| MI:0401            | Biochemical                        | MI:0091         | Chromatography technology                         | 16     |
| MI:0401            | Biochemical                        | MI:0029         | Cosedimentation through density gradient          | 15     |
| MI:0091            | Chromatography technology          | MI:0071         | Molecular sieving                                 | 14     |
| MI:0401            | Biochemical                        | MI:0028         | Cosedimentation in solution                       | 12     |
| MI:0401            | Biochemical                        | MI:0030         | Cross-linking study                               | 11     |
| MI:0400            | Affinity technology                | MI:0004         | Affinity chromatography technology                | 10     |
| MI:0401            | Biochemical                        | MI:0226         | Ion exchange chromatography                       | 9      |
| MI:0428            | Imaging technique                  | MI:0416         | Fluorescence microscopy                           | 9      |
| MI:0090            | Protein complementation assay      | MI:0809         | Bimolecular fluorescence complementation          | 8      |
| MI:0091            | Chromatography technology          | MI:0096         | Pull down   | 8      |
| MI:0415            | Enzymatic study                    | MI:0515         | Methyltransferase assay                           | 8      |
| MI:0027            | Cosedimentation                    | MI:0029         | Cosedimentation through density gradient          | 7      |
| MI:0040            | Electron microscopy                | MI:1024         | Scanning electron microscopy                      | 6      |
| MI:0401            | Biochemical                        | MI:0027         | Cosedimentation                                   | 6      |
| MI:0401            | Biochemical                        | MI:0047         | Far western blotting                              | 6      |
| MI:0401            | Biochemical                        | MI:0276         | Blue native page                                  | 6      |
| MI:0018            | Two hybrid                         | MI:0727         | Lexa b52 complementation                          | 4      |
| MI:0091            | Chromatography technology          | MI:0006         | Anti bait coimmunoprecipitation                   | 4      |
| MI:0090            | Protein complementation assay      | MI:1037         | Split renilla luciferase complementation          | 3      |
| MI:0401            | Biochemical                        | MI:0404         | Comigration in non denaturing gel electrophoresis | 3      |
| MI:0018            | Two hybrid                         | MI:0399         | Two hybrid fragment pooling approach              | 2      |
| MI:0084            | Phage display                      | MI:0048         | Filamentous phage display                         | 2      |
| MI:0091            | Chromatography technology          | MI:0007         | Anti tag coimmunoprecipitation                    | 2      |
| MI:0401            | Biochemical                        | MI:0605         | Enzymatic footprinting                            | 2      |
| MI:0415            | Enzymatic study                    | MI:0406         | Deacetylase assay                                 | 2      |
| MI:0004            | Affinity chromatography technology | MI:0402         | Chromatin immunoprecipitation assay               | 1      |
| MI:0051            | Fluorescence technology            | MI:0055         | Fluorescent resonance energy transfer             | 1      |
| MI:0090            | Protein complementation assay      | MI:0397         | Two hybrid array                                  | 1      |
| MI:0400            | Affinity technology                | MI:0006         | Anti bait coimmunoprecipitation                   | 1      |
| MI:0400            | Affinity technology                | MI:0019         | Coimmunoprecipitation                             | 1      |
| MI:0401            | Biochemical                        | MI:0400         | Affinity technology                               | 1      |
| MI:0401            | Biochemical                        | MI:0424         | Protein kinase assay                              | 1      |
| MI:0401            | Biochemical                        | MI:0515         | Methyltransferase assay                           | 1      |

End of Table S4:

| <b>Redundant ancestor</b> |                    | <b>Redundant child</b> |                                    | <b>Counts</b> |
|---------------------------|--------------------|------------------------|------------------------------------|---------------|
| <b>PSI-MI code</b>        | <b>PSI-MI term</b> | <b>PSI-MI code</b>     | <b>PSI-MI term</b>                 |               |
| MI:0401                   | Biochemical        | MI:0807                | Comigration in gel electrophoresis | 1             |
| MI:0415                   | Enzymatic study    | MI:0880                | Atpase assay                       | 1             |
| MI:0415                   | Enzymatic study    | MI:0997                | Ubiquitinase assay                 | 1             |
| MI:0415                   | Enzymatic study    | MI:0998                | Deubiquitinase assay               | 1             |
| MI:0428                   | Imaging technique  | MI:0663                | Confocal microscopy                | 1             |

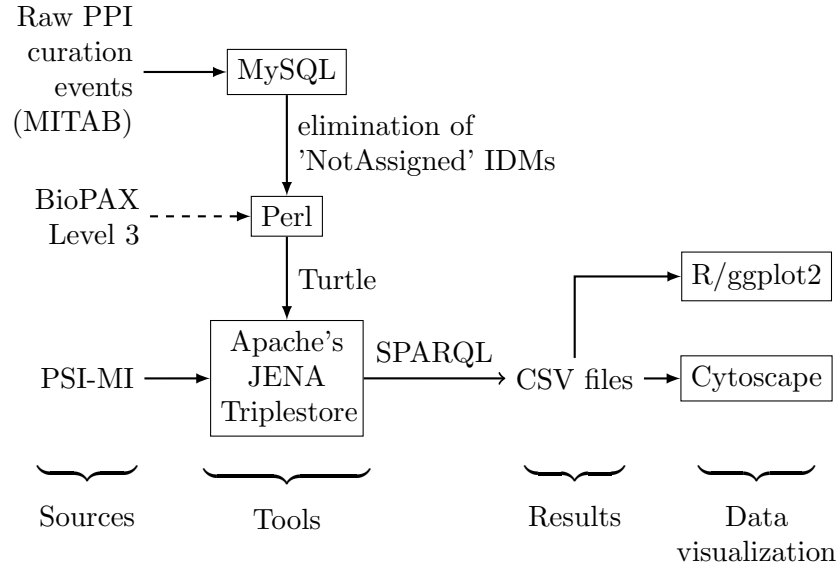


Figure S1: Scheme of the process used to identify redundancies in PPI curation events. Raw curation events were first imported in a MySQL database and then converted into an ontology following the BioPAX level 3 standard. This data was then merged with the PSI-MI ontology into a RDF triplestore powered by the Apache JENA engine. The detection of redundancies was done by a set of SPARQL queries that yield CSV output files. Data visualization was performed using the ggplot2 package of R or Cytoscape.



```

PREFIX : <http://www.biopax.org/examples/myExample#>
PREFIX bp: <http://www.biopax.org/release/biopax-level3.owl#>
PREFIX obo: <http://purl.obolibrary.org/obo/>
PREFIX oboInOwl: <http://www.geneontology.org/formats/oboInOwl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX OWL: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

```

```

SELECT DISTINCT
  (STRAFTER(STR(?p1), "#") AS ?protein1)
  (STRAFTER(STR(?p2), "#") AS ?protein2)
  (STRAFTER(STR(?pmid), "_") AS ?pubmed)
  (STRAFTER(STR(?dm_code1), "o/") AS ?mi)
  (STRAFTER(STR(?dm_name1), "#") AS ?idm)
WHERE
{
  ?ipp1 rdf:type bp:MolecularInteraction ;
        bp:participant ?p1 ;
        bp:participant ?p2 ;
        bp:xref ?pmid ;
        bp:evidence ?dm_name1 .

  ?dm_name1 bp:evidenceCode ?m_vocab1 .
  ?m_vocab1 bp:xref ?dm_code1 .

  FILTER ( STR(?p1) < STR(?p2) )
  FILTER NOT EXISTS {
    ?ipp2 rdf:type bp:MolecularInteraction ;
          bp:participant ?p1 ;
          bp:participant ?p2 ;
          bp:xref ?pmid ;
          bp:evidence ?dm_name2 .

    ?dm_name2 bp:evidenceCode ?m_vocab2 .
    ?m_vocab2 bp:xref ?dm_code2 .

    ?dm_code2 rdfs:subClassOf+ ?dm_code1 .
  }
}

```

Figure S2: SPARQL query to retrieve PPIs in heterodimers without explicit nor implicit redundancies (methods-no-parents.rq).

```

PREFIX : <http://www.biopax.org/examples/myExample#>
PREFIX bp: <http://www.biopax.org/release/biopax-level3.owl#>
PREFIX obo: <http://purl.obolibrary.org/obo/>
PREFIX oboInOwl: <http://www.geneontology.org/formats/oboInOwl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX OWL: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT DISTINCT
  (STRAFTER(STR(?p1), "#") AS ?protein1)
  (STRAFTER(STR(?p1), "#") AS ?protein2)
  (STRAFTER(STR(?pmid), "_") AS ?pubmed)
  (STRAFTER(STR(?dm_code1), "o/") AS ?mi)
  (STRAFTER(STR(?dm_name1), "#") AS ?idm)
WHERE
{
  ?ipp1 rdf:type bp:MolecularInteraction ;
    bp:participant1 ?p1 ;
    bp:participant2 ?p1 ;
    bp:xref ?pmid ;
    bp:evidence ?dm_name1 .

  ?dm_name1 bp:evidenceCode ?m_vocab1 .
  ?m_vocab1 bp:xref ?dm_code1 .

  FILTER NOT EXISTS {
    ?ipp2 rdf:type bp:MolecularInteraction ;
      bp:participant1 ?p1 ;
      bp:participant2 ?p1 ;
      bp:xref ?pmid ;
      bp:evidence ?dm_name2 .

    ?dm_name2 bp:evidenceCode ?m_vocab2 .
    ?m_vocab2 bp:xref ?dm_code2 .

    ?dm_code2 rdfs:subClassOf+ ?dm_code1 .
  }
}

```

Figure S3: SPARQL query to retrieve PPIs in homodimers without explicit nor implicit redundancies (methods-no-parents-equal.rq).

```

PREFIX : <http://www.biopax.org/examples/myExample#>
PREFIX bp: <http://www.biopax.org/release/biopax-level3.owl#>
PREFIX mi: <http://purl.obolibrary.org/obo/mi#>
PREFIX oboInOwl: <http://www.geneontology.org/formats/oboInOwl#>
PREFIX OWL: <http://www.w3.org/2002/07/owl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

```

```

SELECT DISTINCT
  (STRAFTER(STR(?p1), "#") AS ?protein1)
  (STRAFTER(STR(?p2), "#") AS ?protein2)
  (STRAFTER(STR(?pmid), "_") AS ?pubmed)
  (STRAFTER(STR(?dm_code1), "o/") AS ?mi)
  (STRAFTER(STR(?dm_name1), "#") AS ?idm)

WHERE
{
  ?ppi1 rdf:type bp:MolecularInteraction ;
    bp:participant ?p1 ;
    bp:participant ?p2 ;
    bp:xref ?pmid ;
    bp:evidence ?dm_name1 .

  ?dm_name1 bp:evidenceCode ?m_vocab1 .
  ?m_vocab1 bp:xref ?dm_code1 .

  ?dm_code1 :apid_suppl_is_binary true .

  FILTER ( STR(?p1) < STR(?p2) ) .
  FILTER NOT EXISTS {
    ?ppi2 rdf:type bp:MolecularInteraction ;
      bp:participant ?p1 ;
      bp:participant ?p2 ;
      bp:xref ?pmid ;
      bp:evidence ?dm_name2 .
    ?dm_name2 bp:evidenceCode ?m_vocab2 .
    ?m_vocab2 bp:xref ?dm_code2 .

    ?dm_code2 :apid_suppl_is_binary false .

    ?dm_code1 rdfs:subClassOf+ ?dm_code2 .
  }
}
GROUP BY ?p1 ?p2 ?pmid ?dm_name1

```

Figure S4: SPARQL query to retrieve only binary PPIs in heterodimers without explicit nor implicit redundancies (methods-no-parents-only-bin.rq).

```

PREFIX : <http://www.biopax.org/examples/myExample#>
PREFIX bp: <http://www.biopax.org/release/biopax-level3.owl#>
PREFIX mi: <http://purl.obolibrary.org/obo/mi#>
PREFIX oboInOwl: <http://www.geneontology.org/formats/oboInOwl#>
PREFIX OWL: <http://www.w3.org/2002/07/owl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

```

```

SELECT DISTINCT

```

```

  (STRAFTER(STR(?p1), "#") AS ?protein1)
  (STRAFTER(STR(?p1), "#") AS ?protein2)
  (STRAFTER(STR(?pmid), "_") AS ?pubmed)
  (STRAFTER(STR(?dm_code1), "o/") AS ?mi)
  (STRAFTER(STR(?dm_name1), "#") AS ?idm)

```

```

WHERE

```

```

{
  ?ppi1 rdf:type bp:MolecularInteraction ;
        bp:participant1 ?p1 ;
        bp:participant2 ?p1 ;
        bp:xref ?pmid ;
        bp:evidence ?dm_name1 .

  ?dm_name1 bp:evidenceCode ?m_vocab1 .
  ?m_vocab1 bp:xref ?dm_code1 .

  ?dm_code1 :apid_suppl_is_binary true .

  FILTER NOT EXISTS {
    ?ppi2 rdf:type bp:MolecularInteraction ;
          bp:participant1 ?p1 ;
          bp:participant2 ?p1 ;
          bp:xref ?pmid ;
          bp:evidence ?dm_name2 .
    ?dm_name2 bp:evidenceCode ?m_vocab2 .
    ?m_vocab2 bp:xref ?dm_code2 .

    ?dm_code2 :apid_suppl_is_binary false .

    ?dm_code1 rdfs:subClassOf+ ?dm_code2 .
  }
}
GROUP BY ?p1 ?p2 ?pmid ?dm_code1 ?dm_name1

```

Figure S5: SPARQL query to retrieve only binary PPIs in homodimers without explicit nor implicit redundancies (methods-no-parents-only-bin-equal.rq).

```

PREFIX : <http://www.biopax.org/examples/myExample#>
PREFIX bp: <http://www.biopax.org/release/biopax-level3.owl#>
PREFIX obo: <http://purl.obolibrary.org/obo/>
PREFIX oboInOwl: <http://www.geneontology.org/formats/oboInOwl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX OWL: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

```

```

SELECT DISTINCT
  (STRAFTER(STR(?p1), "#") AS ?protein1)
  (STRAFTER(STR(?p2), "#") AS ?protein2)
  (STRAFTER(STR(?pmid), "_") AS ?pubmed)
  (STRAFTER(STR(?dm_code1), "o/") AS ?mi1)
  (STRAFTER(STR(?dm_name1), "#") AS ?idm1)
  (STRAFTER(STR(?dm_code2), "o/") AS ?mi2)
  (STRAFTER(STR(?dm_name2), "#") AS ?idm2)
  (STRAFTER(STR(?db1), "#") AS ?database1)
  (STRAFTER(STR(?db2), "#") AS ?database2)
WHERE
{
  ?ipp1 rdf:type bp:MolecularInteraction ;
    bp:participant ?p1 ;
    bp:participant ?p2 ;
    bp:xref ?pmid ;
    bp:datasource ?db1 ;
    bp:evidence ?dm_name1 .

  ?dm_name1 bp:evidenceCode ?m_vocab1 .
  ?m_vocab1 bp:xref ?dm_code1 .

  ?ipp2 rdf:type bp:MolecularInteraction ;
    bp:participant ?p1 ;
    bp:participant ?p2 ;
    bp:xref ?pmid ;
    bp:datasource ?db2 ;
    bp:evidence ?dm_name2 .

  ?dm_name2 bp:evidenceCode ?m_vocab2 .
  ?m_vocab2 bp:xref ?dm_code2 .

  FILTER (STR(?p1) < STR(?p2)) .
  FILTER EXISTS {
    ?dm_code2 rdfs:subClassOf+ ?dm_code1 .
  }
}

```

Figure S6: SPARQL query to retrieve redundancies in PPIs in heterodimers with source databases (methods-parents-with-dbs.rq).

```

PREFIX : <http://www.biopax.org/examples/myExample#>
PREFIX bp: <http://www.biopax.org/release/biopax-level3.owl#>
PREFIX obo: <http://purl.obolibrary.org/obo/>
PREFIX oboInOwl: <http://www.geneontology.org/formats/oboInOwl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX OWL: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

```

```

SELECT DISTINCT
  (STRAFTER(STR(?p1), "#") AS ?protein1)
  (STRAFTER(STR(?p1), "#") AS ?protein2)
  (STRAFTER(STR(?pmid), "_") AS ?pubmed)
  (STRAFTER(STR(?dm_code1), "o/") AS ?mi1)
  (STRAFTER(STR(?dm_name1), "#") AS ?idm1)
  (STRAFTER(STR(?dm_code2), "o/") AS ?mi2)
  (STRAFTER(STR(?dm_name2), "#") AS ?idm2)
  (STRAFTER(STR(?db1), "#") AS ?database1)
  (STRAFTER(STR(?db2), "#") AS ?database2)
WHERE
{
  ?ipp1 rdf:type bp:MolecularInteraction ;
    bp:participant1 ?p1 ;
    bp:participant2 ?p1 ;
    bp:xref ?pmid ;
    bp:datasource ?db1 ;
    bp:evidence ?dm_name1 .

  ?dm_name1 bp:evidenceCode ?m_vocab1 .
  ?m_vocab1 bp:xref ?dm_code1 .

  ?ipp2 rdf:type bp:MolecularInteraction ;
    bp:participant1 ?p1 ;
    bp:participant2 ?p1 ;
    bp:xref ?pmid ;
    bp:datasource ?db2 ;
    bp:evidence ?dm_name2 .

  ?dm_name2 bp:evidenceCode ?m_vocab2 .
  ?m_vocab2 bp:xref ?dm_code2 .

  FILTER EXISTS {
    ?dm_code2 rdfs:subClassOf+ ?dm_code1 .
  }
}

```

Figure S7: SPARQL query to retrieve redundancies in PPIs in homodimers with source databases (methods-parents-with-dbs-equal.rq).

### Human PPIs

| Redundancies             | Apparent number of EEs |       |       |      |       |
|--------------------------|------------------------|-------|-------|------|-------|
|                          | 2                      | 3     | 4     | 5    | 5+    |
| All                      | 89551                  | 35106 | 16648 | 5254 | 12633 |
| No explicit              | 72902                  | 13193 | 14269 | 3475 | 7170  |
| No explicit nor implicit | 39151                  | 19437 | 4813  | 2249 | 4904  |

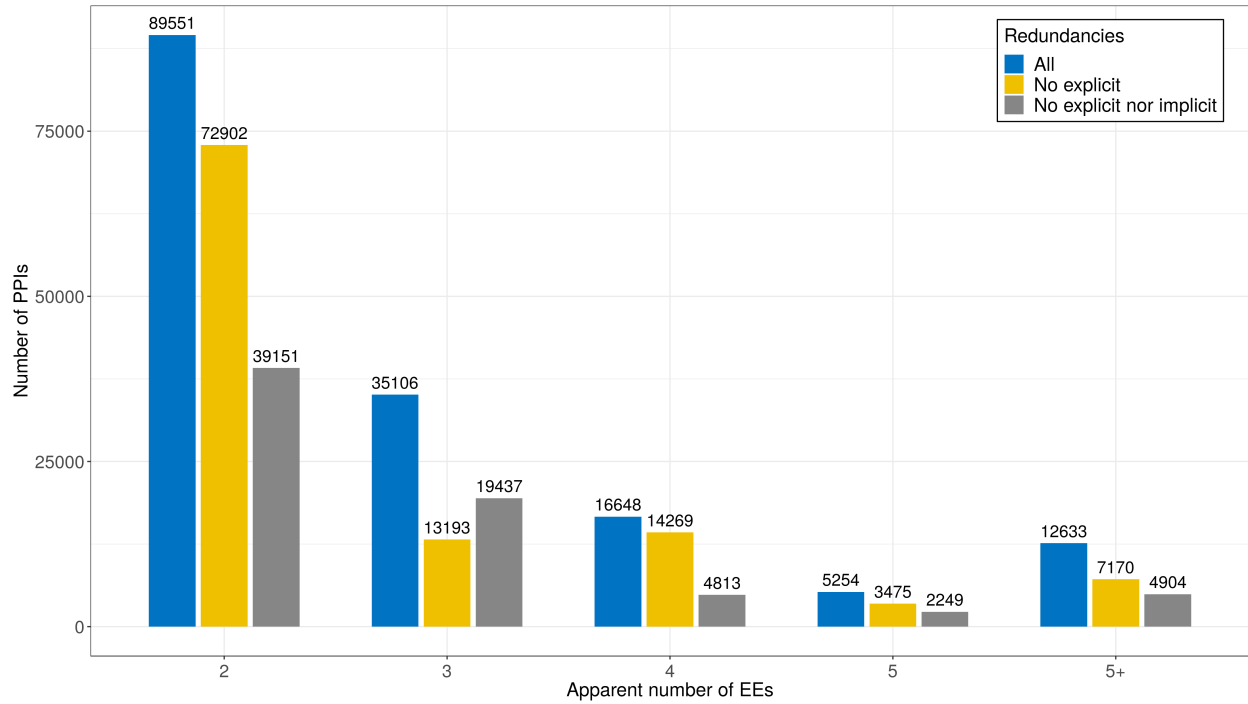


Figure S8: Number of human PPIs apparently supported by more than one experimental evidence (EE) before and after taking into account explicit and implicit redundancy.

### Yeast PPIs

| Redundancies             | Apparent number of EEs |       |      |      |      |
|--------------------------|------------------------|-------|------|------|------|
|                          | 2                      | 3     | 4    | 5    | 5+   |
| All                      | 28102                  | 10557 | 4344 | 2205 | 7105 |
| No explicit              | 26134                  | 5093  | 2799 | 1422 | 4787 |
| No explicit nor implicit | 12042                  | 3435  | 1708 | 1058 | 3068 |

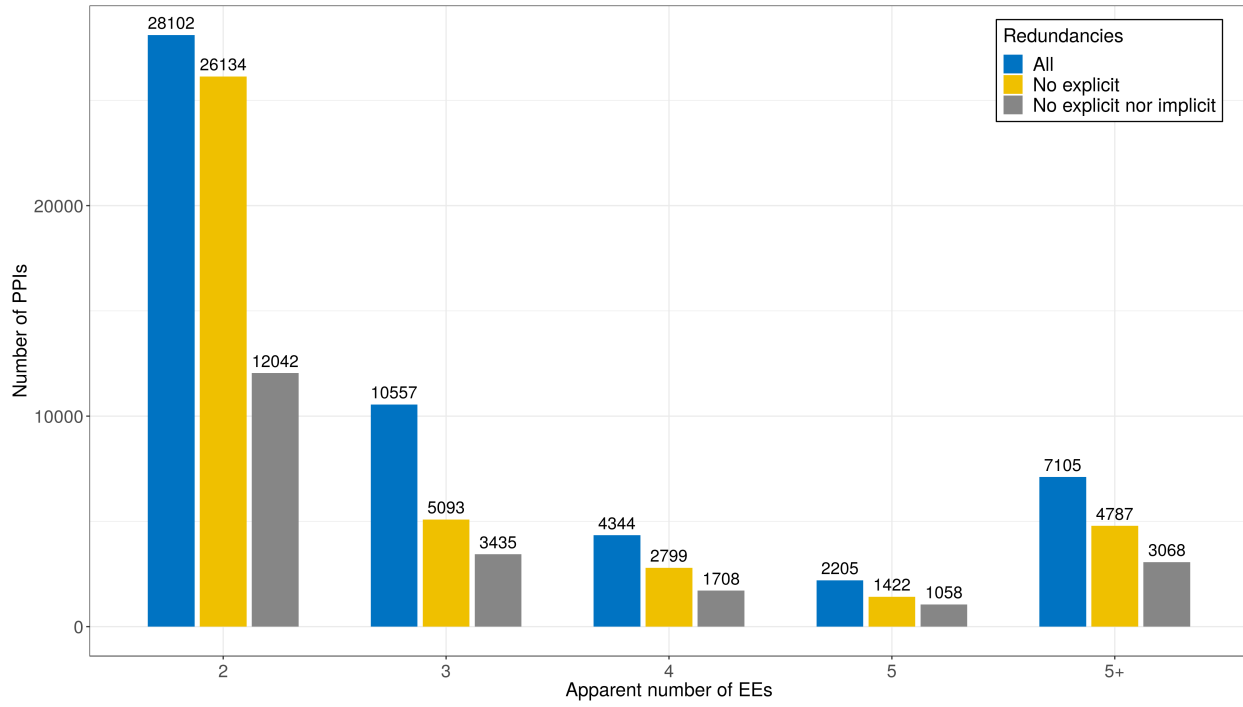


Figure S9: Number of yeast PPIs apparently supported by more than one experimental evidence (EE) before and after taking into account explicit and implicit redundancy.