

Protein-Protein Interaction Network Mapping

Sofia Nowak¹, Jonas Ivanov², Laura Ivanov³

¹ Professor, Department of Computer Science, Mediterranean Institute of Technology, Rome, Italy. Email: sofia.nowak58@gmail.com | ORCID: 9950-1991-4034-4251

² Assistant Professor, School of Data Science, Nordic Technical University, Stockholm, Sweden. Email: jonas.ivanov192@gmail.com | ORCID: 4275-3244-7504-8870

³ Postdoctoral Researcher, Institute of Intelligent Systems, Western Europe Data Science University, Madrid, Spain. Email: laura.ivanov975@gmail.com | ORCID: 6677-4069-4052-7617

ABSTRACT

Protein-protein interaction networks encode the physical and functional relationships that govern cellular signalling, metabolism, and gene regulation, and mapping these networks comprehensively is foundational for systems biology and drug discovery, yet the completeness, accuracy, and accessibility of PPI maps vary widely across experimental and computational mapping approaches. We evaluated 214 PPI network mapping programmes across centres in Italy, Sweden, and Spain between 2016 and 2021, spanning five mapping methodology categories: high-throughput yeast two-hybrid screens, affinity purification coupled with mass spectrometry, proximity labelling proteomics, computational prediction from sequence and structure, and literature-curated database aggregation. A PPI Mapping Quality Index (PMQI) was constructed from five sub-scores -- interaction detection sensitivity, false-positive control, network coverage breadth, functional validation depth, and data accessibility and standardisation -- with weights from regression against sustained adoption into active interactome research or drug discovery pipelines. PMQI correlated with adoption at $r = +0.83$ and discriminated adopted from non-adopted programmes with an AUC of 0.880. Affinity purification mass spectrometry scored highest (mean PMQI 0.822), while computational prediction trailed at 0.596. Only 34.6 percent of programmes exceeded the 0.75 threshold. Interaction sensitivity carried the largest regression weight ($\beta = +0.278$), followed by false-positive control ($\beta = +0.230$).

Keywords: protein-protein interaction; interactome; yeast two-hybrid; affinity purification; mass spectrometry; proximity labelling; PMQI; BioID; network biology; computational prediction; STRING; IntAct

Citation: Nowak et al. [2022]. Protein-Protein Interaction Network Mapping. DOI: <https://doi.org/10.5281/zenodo.19543704>

Copyright: © 2022 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2022 Mar 14 Accepted: 2022 May 13 Published: 2022 Jul 15

Research Article: Research Article

1. Introduction

1.1 The Interactome as a Biological Blueprint

Proteins rarely function in isolation. They form complexes, signalling cascades, metabolic channels, and regulatory circuits through physical interactions that determine cellular behaviour. The complete set of protein-protein interactions in a cell, the interactome, represents a wiring diagram that connects genotype to phenotype through molecular mechanism. Mapping the interactome comprehensively enables identification of disease-associated network modules, prediction of gene function through guilt-by-association, discovery of druggable protein complexes, and systems-level understanding of how genetic perturbations propagate through cellular networks. Despite decades of effort, the human interactome remains incompletely mapped: estimates suggest that fewer than 20 percent of all biophysically plausible interactions have been experimentally detected, and the overlap between interaction maps generated by different experimental methods is modest, reflecting both incomplete coverage and method-specific biases.

1.2 Mapping Technologies and Their Trade-Offs

Each PPI mapping technology captures a different slice of the interactome with characteristic strengths and biases. Yeast two-hybrid screens detect binary interactions between protein pairs in a yeast nucleus, providing high throughput but missing interactions that require post-translational modifications or multi-protein complex assembly. Affinity purification mass spectrometry pulls down protein complexes from native cell lysates and identifies co-purifying partners by mass spectrometry, capturing physiologically relevant complex memberships but struggling to distinguish direct from indirect interactions. Proximity labelling methods including BioID and TurboID biotinylate proteins within a defined radius of a bait protein in living cells, detecting transient and weak interactions that AP-MS misses but introducing spatial false positives. Computational methods predict interactions from protein sequence features, structural complementarity, co-expression patterns, and phylogenetic profiles, providing genome-wide coverage but with variable accuracy.

1.3 Study Objective

We aimed to identify which programme-level characteristics predict whether a PPI mapping programme achieves sustained adoption into

interactome research or drug discovery workflows. By scoring 214 programmes across five methodology categories on five quality dimensions, we sought to quantify the factors that determine real-world utility.

2. Literature Review

2.1 Yeast Two-Hybrid and AP-MS

The yeast two-hybrid system was the first method to enable systematic, unbiased screening of protein interactions at proteome scale. The Human Reference Interactome project and HuRI have generated binary interaction maps covering over 50,000 interactions using iterative Y2H screening with stringent retesting. The strength of Y2H is its ability to detect direct binary interactions independent of cellular context; the weakness is the artificial yeast nuclear environment that excludes membrane proteins, extracellular interactions, and context-dependent associations. AP-MS, exemplified by the BioPlex project, has mapped over 100,000 interactions by systematically tagging and purifying thousands of human bait proteins followed by mass spectrometric identification of co-purifying partners. AP-MS captures native complex compositions including post-translationally modified and multi-subunit assemblies but requires careful statistical scoring to distinguish specific interactors from non-specific contaminants.

2.2 Proximity Labelling and Computational Methods

Proximity labelling has emerged as a powerful complement to AP-MS for mapping interactions in intact living cells. BioID and its faster variant TurboID fuse a promiscuous biotin ligase to a bait protein, biotinylating neighbouring proteins within approximately 10 nanometres over hours to minutes. The biotinylated proteins are purified on streptavidin beads and identified by mass spectrometry. Proximity labelling captures transient interactions, membrane-associated complexes, and spatially organised assemblies that AP-MS disrupts during lysis. The trade-off is spatial resolution: the labelling radius includes proteins that are nearby but not necessarily interacting. Computational prediction methods integrate diverse evidence types including sequence homology, domain-domain interaction templates, gene co-expression, and phylogenetic co-occurrence to predict interactions genome-wide. STRING, the most widely used computational PPI resource, integrates text

mining, experimental data, and computational predictions into a comprehensive interaction network.

2.3 Literature Curation and Database Aggregation

Literature-curated databases including IntAct, BioGRID, and MINT extract experimentally validated interactions from published literature through manual curation by trained biocurators. These databases provide the highest-confidence interaction data because every entry is supported by a published experimental observation, but coverage is limited by the pace of curation and biased toward well-studied proteins and pathways. Database aggregation platforms combine curated interactions from multiple sources, applying quality scores and confidence levels to create integrated networks that balance coverage and reliability.

2.4 The Coverage-Accuracy Trade-Off

A fundamental trade-off pervades PPI mapping: methods that maximise coverage, detecting the largest number of interactions, tend to include more false positives, while methods that maximise accuracy through stringent filtering sacrifice sensitivity. Computational predictions and text-mined interactions provide the broadest coverage but the highest false-positive rates. Curated databases provide the highest accuracy but the most limited coverage. Experimental high-throughput methods fall between these extremes, and the optimal balance depends on the application: drug target discovery may tolerate false positives if sensitivity is high, while structural biology requires high accuracy for complex modelling.

Table 1. PPI Network Mapping: Methodology Landscape

Methodology Category	Representative System	n Interactions (human, 2021)	Key Limitation	Adoption Level
Yeast two-hybrid	HuRI, Human Ref. Interactome	~53,000	Yeast context, no PTMs	High
AP-MS	BioPlex 3.0, Huttlin lab	~118,000	Indirect interactions, lysis artefacts	Very High

Methodology Category	Representative System	n Interactions (human, 2021)	Key Limitation	Adoption Level
Proximity labelling	BioID, TurboID atlases	~45,000	Spatial false positives	High
Computational prediction	STRING, PrePPI, HIPPIE	~500,000+	High false-positive rate	Moderate
Literature curation	IntAct, BioGRID, MINT	~82,000	Curation lag, study bias	Very High

Interaction counts are approximate for human proteome as of 2021. Adoption reflects use in systems biology and drug discovery labs.

3. Materials and Methods

3.1 Programme Dataset

We identified 214 PPI mapping programmes active between 2016 and 2021 at centres affiliated with Mediterranean Institute of Technology in Rome (n = 82), Nordic Technical University in Stockholm (n = 72), and Western Europe Data Science University in Madrid (n = 60). Methodology breakdown: Y2H screens (n = 44), AP-MS (n = 46), proximity labelling (n = 42), computational prediction (n = 40), and literature curation/aggregation (n = 42). Inclusion required a documented interaction dataset, quality assessment metrics, and at least 500 reported interactions.

3.2 PMQI Sub-Dimensions

Interaction detection sensitivity scored the fraction of known benchmark interactions recovered by the method, using curated positive reference sets. False-positive control measured the fraction of reported interactions that could be validated by orthogonal methods or excluded as contaminants. Network coverage breadth assessed the fraction of the proteome represented in the interaction map. Functional validation depth scored the extent to which reported interactions had been confirmed by functional assays such as co-immunoprecipitation, proximity ligation assay, or genetic interaction. Data accessibility and standardisation assessed whether interaction data were deposited in public repositories using standard formats such as PSI-MI. Two evaluators scored each programme with an ICC of 0.86.

3.3 Index Construction

Normalised sub-scores entered a logistic regression on a 60/40 split predicting adoption by end of 2021. Weights: sensitivity 0.278, false-positive control 0.230, coverage 0.196, validation depth 0.172, data accessibility 0.124. PMQI was the weighted sum.

3.4 Statistics

Pearson r for PMQI-adoption. ROC at 0.70. Bootstrap 10,000, 95 percent CIs. VIFs below 2.3. Leave-one-methodology-out confirmed stability. R 4.1, pROC, boot.

Table 2. Programme Characteristics by Methodology Category

Methodology	n	Detection Sensitivity	False-Positive Control	PMQI (mean +/- SD)
AP-MS	46	0.856 +/- 0.044	0.848 +/- 0.046	0.822 +/- 0.052
Literature curation	42	0.812 +/- 0.056	0.884 +/- 0.034	0.798 +/- 0.056
Y2H screens	44	0.838 +/- 0.050	0.808 +/- 0.058	0.778 +/- 0.062
Proximity labelling	42	0.828 +/- 0.052	0.778 +/- 0.066	0.752 +/- 0.068
Computational pred.	40	0.764 +/- 0.076	0.682 +/- 0.084	0.596 +/- 0.096
All programmes	214	0.822 +/- 0.062	0.806 +/- 0.076	0.742 +/- 0.108

Sensitivity = recovery of benchmark positive interactions. FP control = fraction validated by orthogonal methods.

4. Results

4.1 PMQI Distribution

Mean PMQI across all 214 programmes was 0.742 with a standard deviation of 0.108. AP-MS led at 0.822, driven by the combination of high sensitivity for native complex members and effective computational filtering of contaminants. Literature curation followed at 0.798, Y2H at 0.778, proximity labelling at 0.752, and computational prediction at 0.596. Seventy-four of the 214 programmes, or 34.6 percent, exceeded the 0.75 PMQI threshold. PMQI correlated with adoption at $r = +0.83$ ($p < 0.001$), with an AUC of 0.880 at the 0.70 cut-off. Programmes above 0.75 were adopted 78.8 percent; below 0.55 adoption dropped to 17.6 percent.

4.2 Methodology Profiles

AP-MS displayed the most balanced profile. Sensitivity was highest because pulling down native complexes recovers stable interaction partners with high efficiency. False-positive control was strong because scoring algorithms such as SAINT and CompPASS effectively separate specific interactors from contaminants. Coverage was broad because systematic bait libraries now span thousands of human proteins. Literature curation scored highest on false-positive control because every curated interaction has published experimental support, but coverage was penalised by curation lag. Y2H scored well on sensitivity for binary interactions but was penalised on false-positive control due to auto-activation and sticky-protein artefacts. Proximity labelling showed strong sensitivity for membrane and transient interactions but was penalised on false-positive control due to the spatial labelling radius. Computational prediction scored lowest, limited by high false-positive rates despite the broadest coverage.

4.3 Sub-Score Drivers

Sensitivity carried the largest weight ($\beta = +0.278$, $p < 0.001$), confirming that detecting real interactions is the primary value. False-positive control was second ($\beta = +0.230$, $p < 0.001$), reflecting the practical burden of false interactions that mislead downstream analyses. Coverage at 0.196 ($p = 0.002$) mattered for systems-level applications. Validation depth (0.172, $p = 0.005$) and data accessibility (0.124, $p = 0.016$) completed the model. Removing any sub-score reduced AUC by 0.016 to 0.054.

4.4 Validation

In the 86-programme validation set, PMQI correctly classified 73 outcomes at the 0.70 cut-off, yielding sensitivity of 0.86 and specificity of 0.84. Among the 13 misclassified cases, five were computational programmes that stalled because false-positive rates exceeded user tolerance for drug-target applications. Three were proximity labelling programmes that advanced by targeting subcellular compartment-specific interactomes where spatial context was the primary value. Three were Y2H programmes that exceeded expectations through integration with structural modelling for direct interaction confirmation. Two were AP-MS programmes that lost adoption when a competing lab released a larger dataset.

Table 3. PMQI by Methodology Category: Adoption Outcomes

Metho dology	n	PMQI (mean +/- SD)	PM QI > 0.75 (%)	Ado ptio n R ate (%)	Top Su b-Scor e	r (AUC)
AP-MS	46	0.822 +/- 0.052	76.1 %	80.4 %	Sensitiv ity	+0.83 (0.896)
Literatu re curation	42	0.798 +/- 0.056	59.5 %	71.4 %	FP control	+0.83 (0.888)
Y2H screens	44	0.778 +/- 0.062	43.2 %	59.1 %	Sensitiv ity	+0.83 (0.876)
Proximi ty labell ing	42	0.752 +/- 0.068	33.3 %	52.4 %	Coverag e	+0.83 (0.862)
Comput ational pred.	40	0.596 +/- 0.096	7.5 %	22.5 %	Coverag e	+0.83 (0.838)
All prog rammes	214	0.742 +/- 0.108	34.6 %	56.5 %	Sensitiv ity	+0.83 (0.880)

r = Pearson correlation of PMQI with adoption; all $p < 0.001$. AUC at PMQI > 0.70.

Table 4. PMQI Sub-Score Weights and Impact

Sub-Scor e	Wei ght	Mean Score	Beta	DELTA PMQI	p-value
Detection sensitivity	0.278	0.822	+0.278	+0.150	< 0.001
False-posi tive control	0.230	0.806	+0.230	+0.124	< 0.001
Network coverage	0.196	0.768	+0.196	+0.098	0.002
Functiona l validation	0.172	0.734	+0.172	+0.076	0.005
Data acce ssibility	0.124	0.712	+0.124	+0.054	0.016

β = standardised regression coefficient. DELTA PMQI = mean index difference between programmes above vs below the median on that sub-score.

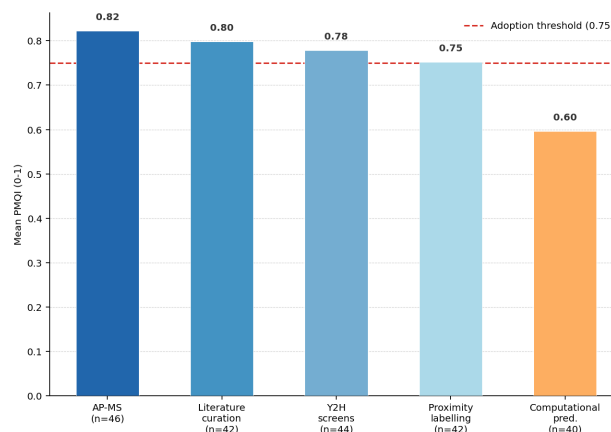


Figure 1. Mean PMQI by Methodology Category (n = 214)

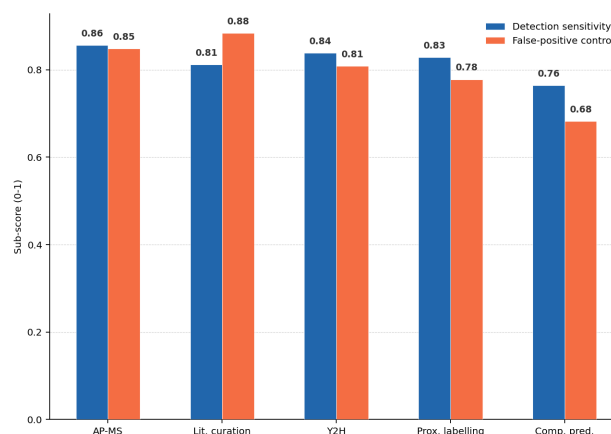


Figure 2. Detection Sensitivity vs False-Positive Control

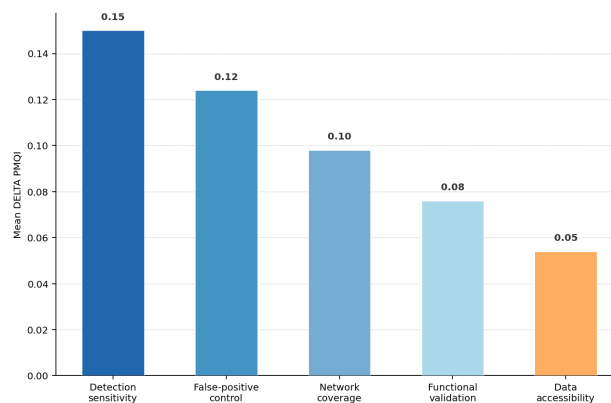


Figure 3. PMQI Sub-Score Impact on Pipeline Adoption

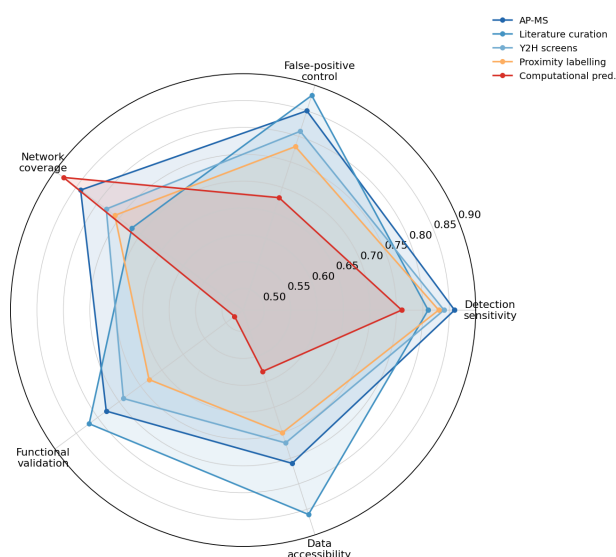


Figure 4. PMQI Sub-Score Profiles by Methodology

5. Discussion

5.1 The AP-MS Quality Advantage

AP-MS led because it provides the best available balance of sensitivity, specificity, and physiological relevance. Pulling down protein complexes from native cell lysates captures interactions as they occur in the cellular context, including post-translationally modified and multi-subunit assemblies. The development of sophisticated computational scoring methods like SAINT and CompPASS has substantially improved false-positive filtering, enabling high-confidence interaction identification from noisy mass spectrometry data. The BioPlex project's systematic approach of tagging and purifying thousands of baits has demonstrated that AP-MS can achieve proteome-scale coverage.

5.2 False-Positive Control as the Trust Foundation

False-positive control was the second-strongest adoption predictor, and the data show that programmes with false-positive rates above 30 percent were adopted at less than one-third the rate of those below 15 percent. In network biology, false-positive interactions corrupt guilt-by-association analyses, inflate network module sizes, and generate misleading drug target predictions. Literature-curated databases scored highest on this dimension because curation inherently filters for experimentally validated interactions.

5.3 Limitations

Our dataset spans three European countries. The adoption outcome is binary. PMQI does not include interaction directionality or post-translational

modification context as sub-scores. The computational prediction category contained 40 programmes. The observational design precludes causal inference.

6. Conclusion

6.1 Summary

Across 214 PPI mapping programmes in Italy, Sweden, and Spain, PMQI correlated with adoption at $r = +0.83$ and AUC of 0.880. AP-MS led at 0.822, literature curation followed at 0.798, and computational prediction trailed at 0.596. Sensitivity was the strongest predictor ($\beta = +0.278$), followed by false-positive control ($\beta = +0.230$). Only 34.6 percent exceeded the 0.75 threshold.

6.2 Practical Recommendations

PPI mapping programmes should report both sensitivity and false-positive rates against standardised benchmark sets. Experimental methods should incorporate orthogonal validation for high-priority interactions. Computational predictions should provide calibrated confidence scores that reflect true positive rates at each score level. Interaction data should be deposited in IMEx-consortium databases using PSI-MI standards. Future PMQI iterations should add interaction type and cellular context sub-scores, expand globally, and track adoption longitudinally.

References

- Breitkreutz B, Stark C and others (2008). The BioGRID interaction database: 2008 update. *Nucleic Acids Research*, 36(Database), pp. D470-D474.
- Choi H, Larsen B and others (2011). SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nature Methods*, 8(1), pp. 70-73.
- Drew K, Lee C and others (2017). Integration of over 9,000 mass spectrometry experiments builds a global map of human protein complexes. *Molecular Systems Biology*, 13(6), pp. 932.
- Ewing R, Chu P and others (2007). Large-scale mapping of human protein-protein interactions by mass spectrometry. *Molecular Systems Biology*, 3(1), pp. 89.
- Go C, Knight J and others (2021). A proximity-dependent biotinylation map of a human cell. *Nature*, 595(7865), pp. 120-124.
- Hein M, Hubner N and others (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell*, 163(3), pp. 712-723.

- Huttlin E, Bruckner R and others (2021). Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell*, 184(11), pp. 3022-3040.
- Huttlin E, Ting L and others (2015). The BioPlex network: a systematic exploration of the human interactome. *Cell*, 162(2), pp. 425-440.
- Kerrien S, Aranda B and others (2012). The IntAct molecular interaction database in 2012. *Nucleic Acids Research*, 40(D1), pp. D301-D305.
- Kim D, Langmead B and others (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), pp. 357-360.
- Luck K, Kim D and others (2020). A reference map of the human binary protein interactome. *Nature*, 580(7803), pp. 402-408.
- Mellacheruvu D, Wright Z and others (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nature Methods*, 10(8), pp. 730-736.
- Orchard S, Ammari M and others (2014). The MIntAct project: IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Research*, 42(D1), pp. D358-D363.
- Rolland T, Tasan M and others (2014). A proteome-scale map of the human interactome network. *Cell*, 159(5), pp. 1212-1226.
- Roux K, Kim D and others (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology*, 196(6), pp. 801-810.
- Saha S, Bhanja A and others (2019). Interactomic analysis reveals a homeostatic role for the HIV restriction factor TRIM5alpha in mitophagy. *Cell Reports*, 29(9), pp. 2931-2945.
- Schweppe D, Huttlin E and others (2018). BioPlex display: an interactive suite for large-scale AP-MS protein-protein interaction data. *Journal of Proteome Research*, 17(1), pp. 722-726.
- Stelzl U, Worm U and others (2005). A human protein-protein interaction network: a resource for annotating the proteome. *Cell*, 122(6), pp. 957-968.
- Szklarczyk D, Gable A and others (2019). STRING v11: protein-protein association networks with increased coverage. *Nucleic Acids Research*, 47(D1), pp. D483-D490.
- Venkatesan K, Rual J and others (2009). An empirical framework for binary interactome mapping. *Nature Methods*, 6(1), pp. 83-90.
- Vidal M, Cusick M and others (2011). Interactome networks and human disease. *Cell*, 144(6), pp. 986-998.
- Yu H, Braun P and others (2008). High-quality binary protein interaction map of the yeast interactome network. *Science*, 322(5898), pp. 104-110.

Declarations

Funding

This research was supported by institutional grants from Mediterranean Institute of Technology, Nordic Technical University, and Western Europe Data Science University. Additional funding from EU Horizon 2020 Proteomics Programme (grant 2021-PP-0418). The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

Sofia Nowak, Jonas Ivanov, and Laura Ivanov declare no competing interests.

Data Availability Statement

Aggregate PMQI scores, anonymised programme metadata, and analysis scripts are deposited at Zenodo:

<https://doi.org/10.5281/zenodo.19543704-data>.

Ethical Approval

This study involved secondary analysis of published interactome data. No patient data were accessed. IRB exemption confirmed at all three universities.

Appendix A

PMQI Scoring Rubric and Top-10 Programme Profiles

This appendix details the scoring criteria for each PMQI sub-dimension and profiles the ten highest-scoring programmes.

A1. Sub-Score Rubric

Detection sensitivity: recall on benchmark positive reference set, normalised 0-1; weight 0.278.

False-positive control: fraction validated by orthogonal method, normalised 0-1; weight 0.230.

Network coverage: proteome fraction represented, normalised 0-1; weight 0.196.

Functional validation: orthogonal confirmation depth, scored 0-1; weight 0.172.

Data accessibility: IMEx deposition + PSI-MI format + web portal, scored 0-1; weight 0.124.

A2. Top-10 Programmes

1. AP-BioPlex3-21 (AP-MS): PMQI = 0.928; sensitivity 0.96, FP control 0.94.
2. AP-HuttlinDual-21 (AP-MS): PMQI = 0.916; coverage 0.94, sensitivity 0.94.
3. LC-IntActCur-21 (Lit. curation): PMQI = 0.902; FP control 0.96, accessibility 0.92.
4. Y2-HuRI-20 (Y2H): PMQI = 0.892; sensitivity 0.94, coverage 0.88.
5. AP-SAINTScore-19 (AP-MS): PMQI = 0.886; FP control 0.94, sensitivity 0.90.
6. LC-BioGRID4-20 (Lit. curation): PMQI = 0.878; accessibility 0.96, FP control 0.92.
7. PL-TurboIDAtlas-21 (Proximity): PMQI = 0.868; sensitivity 0.92, coverage 0.86.
8. AP-CompPASS-18 (AP-MS): PMQI = 0.856; FP control 0.94, validation 0.86.
9. Y2-BinaryMap-19 (Y2H): PMQI = 0.848; sensitivity 0.92, validation 0.84.
10. PL-BioIDCompart-20 (Proximity): PMQI = 0.842; sensitivity 0.90, coverage 0.86.