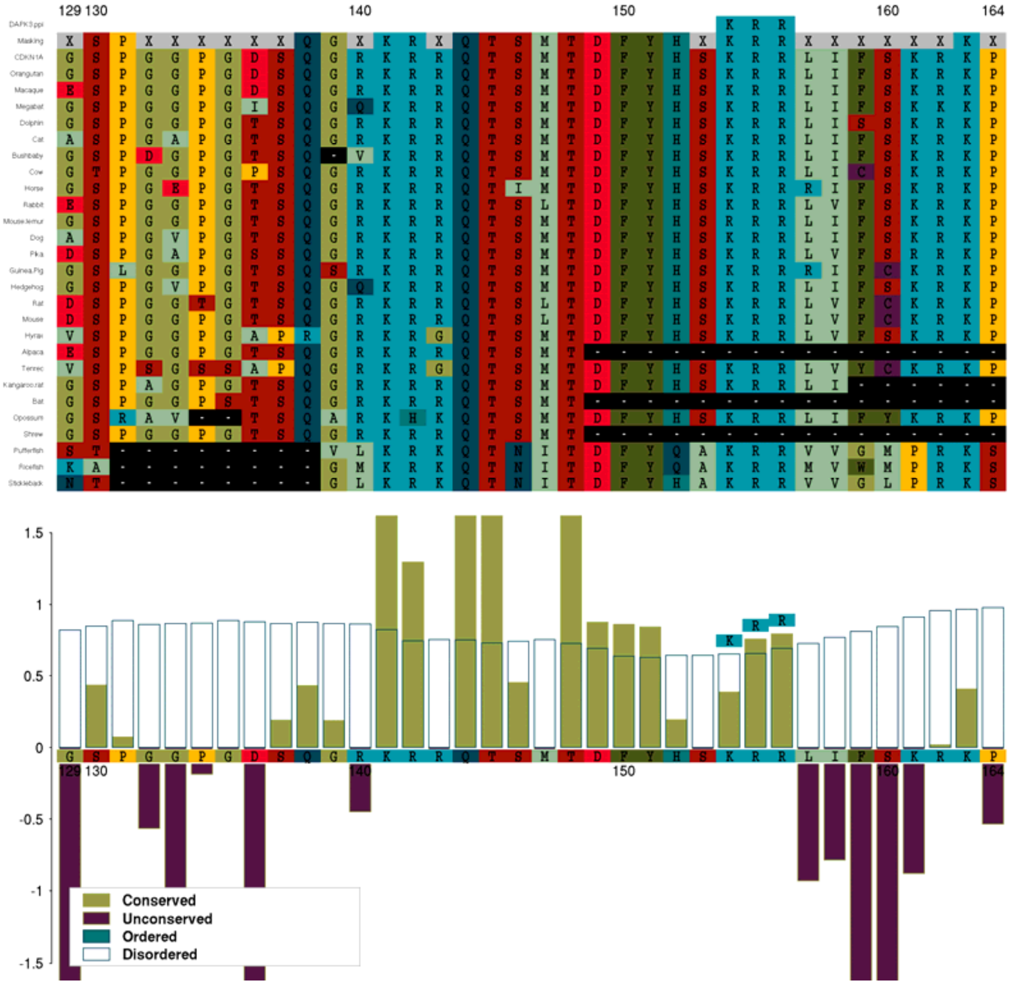
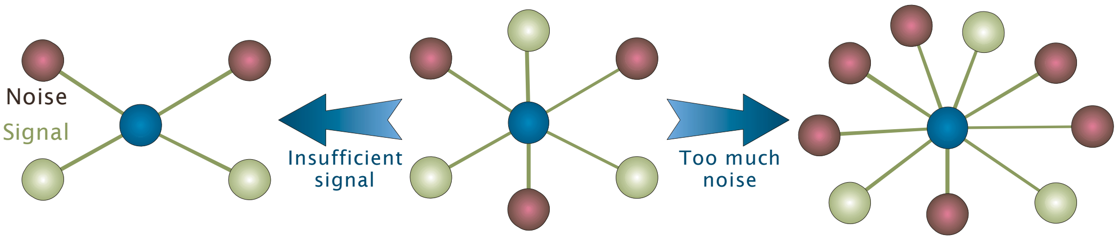
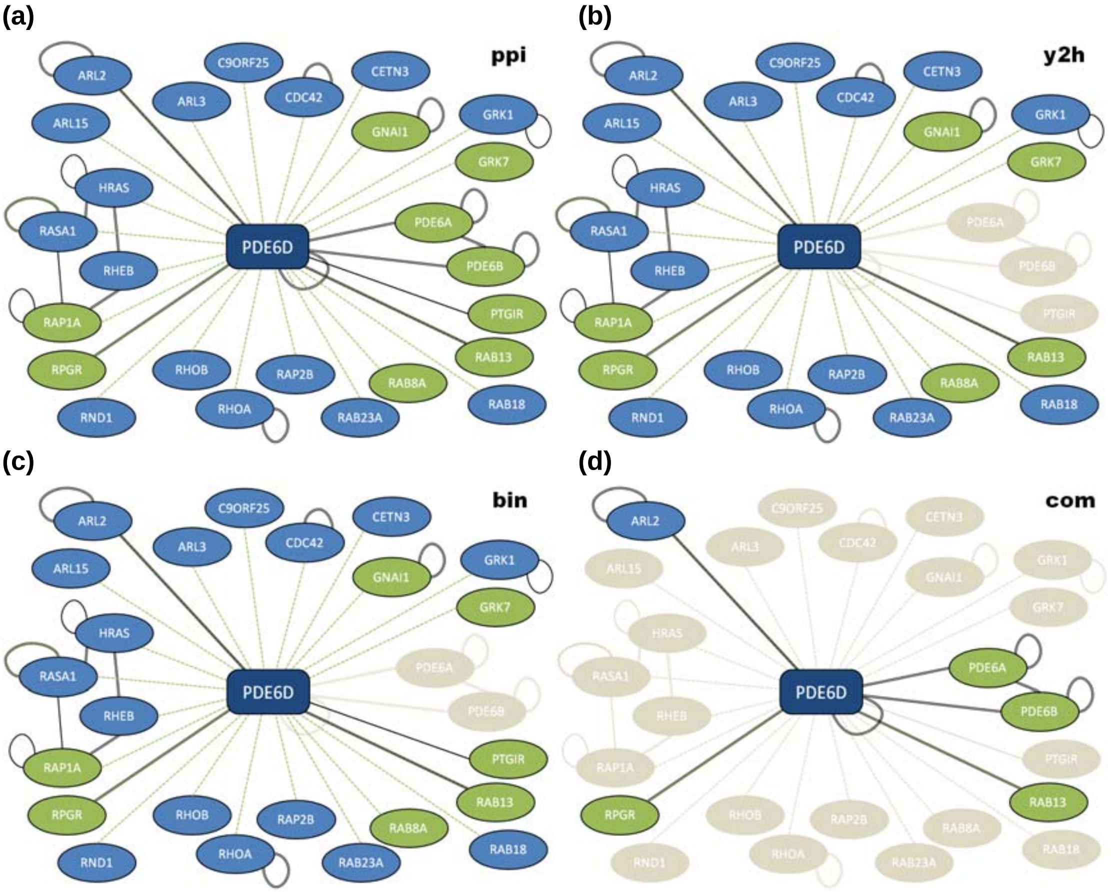
## Supplementary Figures



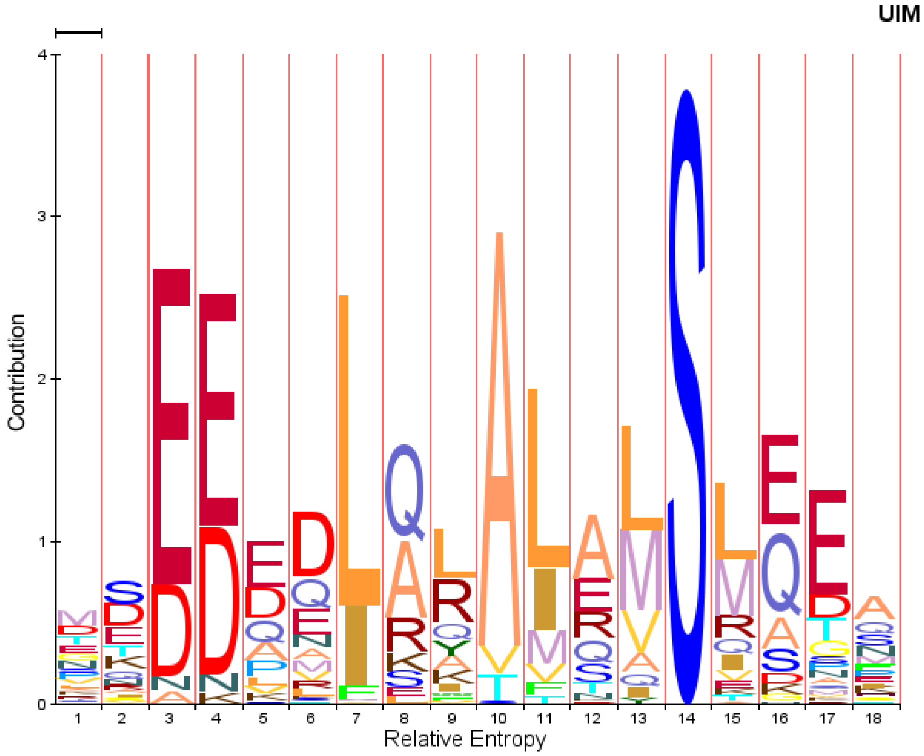
**Supplementary Fig. S1.** Region of CDKN1A showing conserved PCNA ligand motif.
The upper panel shows an alignment of a region of human CDKN1A with predicted vertebrate orthologues. The lower panel plots Relative Local Conservation (RLC) and IUPred disorder prediction scores for each residue. Residues designated “unconserved” (RLC < 0) and/or “disordered” (IUPred < 0.2) were masked out of the analysis (X in the “Masking” row of the alignment).



**Supplementary Fig. S2.** Signal-to-noise trade-off.
Dataset quality has a big impact on results quality. Adding incorrect interactors will increase the noise. Inadvertently removing correct interactors, however, will reduce the signal. There is therefore a trade-off in the strictness of selection criteria for accepted interactors.



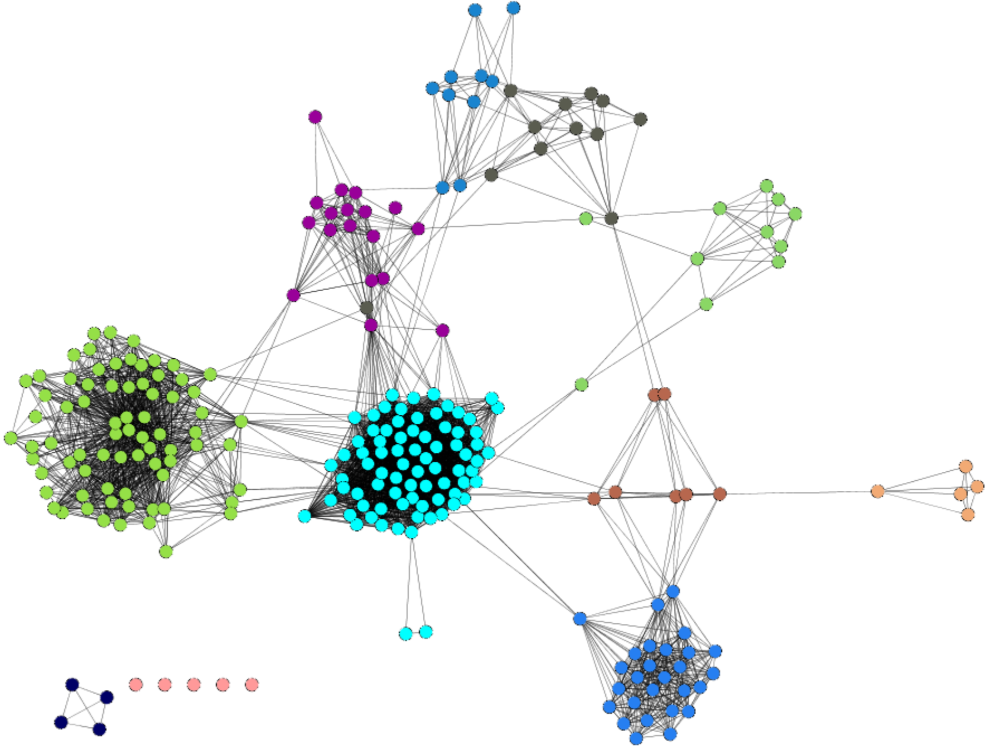
**Supplementary Fig. S3.** PPI compilation strategies applied to PDE6D interactome. Full compiled interactome for Retinal rod rhodopsin-sensitive cGMP 3',5'-cyclic phosphodiesterase subunit delta (PDE6D) with evidence types. Black double-line, “complex-enriched” interaction; green dashed line, yeast-two hybrid; single black line, other interaction evidence. Spokes returning the CAAX box motif are highlighted in green. For each compilation strategy, rejected interactors are faded.



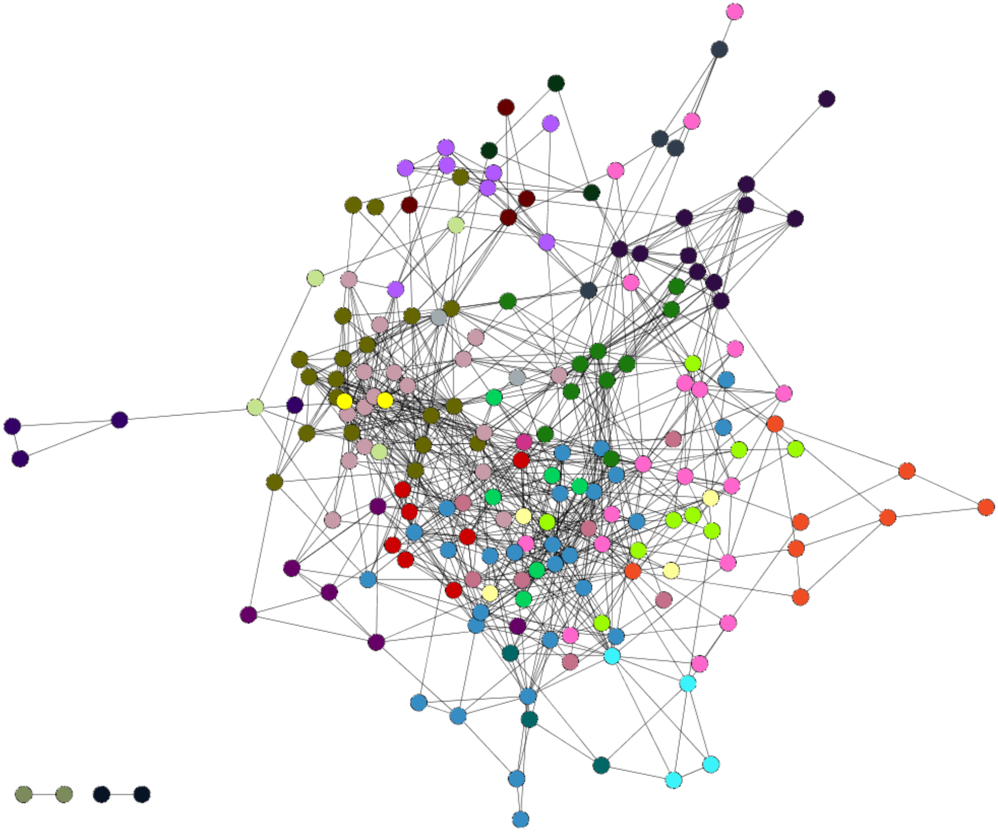
**Supplementary Fig. S4.** Ubiquitin Interacting Motif (UIM) sequence profile.
The LQxxL motif matches positions 7-11 of the UIM. Sequence Logo from http://pfam.sanger.ac.uk/.

**Supplementary Fig. S5.** Hubs and spokes returning the SxLxxNAxxF motif, part of the PABPC1 interacting region.
Hubs returning the motif are in the inside with spokes containing the motif around the outside. Lines joining hubs and spokes indicate spoke membership in the hub datasets. In addition to the True Positive hub PABPC1, spokes containing SxLxxNAxxF are also found in one Off-Target domain (RRM_1) and one random UPC dataset. Variants of this motif are also returned by the PABP domain (not shown).

**(a) TP Motif MCL Clusters**



**(b) Novel Motif MCL Clusters**



**Supplementary Fig. S6.** MCL clustering of (a) TP and (b) Novel SLiMs based on ComapriMotif similarity.

Each node is a motif, with each edge representing a CompariMotif similarity match. Each colour represents a different MCL cluster. Details of clusters can be found in Supplementary Table S3.

# Supplementary Tables

## **Supplementary Table S1.** Significant True Positive SLiM Predictions

| **Known SLiM** | | | | **Significant True Positive SLiMFinder predictions** | | | |
|---|---|---|---|---|---|---|---|
| **Name** | **Description** | | **Hubs** | **ppi** | **y2h** | **bin** | **com** |
| LIG_1433 | Dominant 14-3-3 ligand motif (Variation of ELM:LIG_14-3-3_1). Expanded RS.S.P motif - [HKR][ST].[ST].P. | Genes | SFN | [ILV].[KR][ST].[ST] (3.77e-07) | | [IL].[KR][ST].S (3.22e-05) | [ILV].[KR][ST].[ST] (2.89e-07) |
| | | | YWHAB | [LM].[HR][ST].S (7.41e-06) | RS.S.P (0.024) | | [LM].[HR][ST].S (7.69e-07) |
| | | | YWHAG | R[ST]R[ST].S (1.44e-06) | | | RSR[ST].S (0.005) |
| | | | YWHAH | [LM].[HKR][ST].[ST] (2.64e-07) | | R[ST].S[IMV] (8.65e-04) | R[ST].S[FLMV] (1.09e-07) |
| | | | YWHAQ | [FHY][HR].R..S (6.28e-06) | | | [FHY][HR].R..S (7.42e-06) |
| | | | YWHAZ | FR.R..S (3.41e-04) | RS.S.P (0.009) | | RS.S (2.35e-04) |
| | | Domain | 14-3-3 | RS.S (5.23e-08) | RS.S.P (2.59e-05) | [ILM].[HKR][ST].S (9.02e-04) | RS.S (1.25e-06) |
| LIG_AP_GAE | Simplified Gamma-adaptin ear ligand motif, [DE][DE][DE]F.[DE]F (ELM LIG_AP_GAE_1) | Genes | AP1G1 | FADF.S (0.041) | | | |
| | | | GGA2 | FADF.S (0.007) | | | |
| | | Domain | GAT | FADF.S (0.015) | | | |
| LIG_BRCT | SxxF Phosphomotif interacting with BRCA1 | Gene | MAPK10 | T.S..F.F (0.015) | | | |
| LIG_CtBP | PxDLS CtBP interaction ligand (ELM LIG_CtBP) | Genes | CTBP1 | P[ILM]DL (2.52e-05) | P[IL]DL (7.05e-04) | P[IL]DL (0.004) | P[IL]DL (2.55e-05) |
| | | | CTBP2 | | D..[IL]DL (0.007) | D..[IL]DL (0.007) | |
| LIG_CYC | Cyclin recognition motif, [RK].L (ELM LIG_CYCLIN_1, MnM PBMCY00002A) | Domain | Cyclin_N | EE..R.L (2.57e-04) | | | EE..R.L (0.012) |
| LIG_Dyn | KxTQT Dynein Light Chain ligand (ELM LIG_Dynein_DLC8_1) | Gene | DYNLL2 | K.TQ (0.049) | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| LIG_EH | Canonical Eps15 homology (EH) domain binding motif, NPF (ELM LIG_EH_1) | Genes | EPS15 | NP[FY] (0.007) | | | S..NP[FV] (0.002) |
| | | | EPS15L1 | S..NP[FV] (7.67e-04) | n/a | n/a | |
| LIG_GoLoco | Part of G-protein G-alpha domain binding motif (MnM PBMGPC00002A) | Domain | G-alpha | | Q..R.D.QR (0.009) | Q..R.D.QR (0.017) | |
| LIG_PABPC1 | PABPC1 binding region | Gene | PABPC1 | S.L..NA..F (3.38e-07) | | L..NA..F.P (3.51e-10) | S.L..NA..F (1.64e-07) |
| | | Domain | PABP | L..NA..F.P (3.89e-05) | | L..NA..F.P (1.19e-07) | L..NA..F.P (1.84e-05) |
| LIG_PCNA | Q..[IL]..FF PCNA ligand (ELM LIG_PCNA) | Gene | PCNA | Q..[IL]..FF (4.28e-08) | | | Q..[IL]..FF (1.54e-08) |
| LIG_PDZ | Canonical C-terminal PDZ motif [ST].[ILV]$. (ELM LIG_PDZ_1) See also generic PTS1 targetting motif TRG_PTS1 | Genes | CASK | [ST].V$ (2.48e-05) | | | |
| | | | DLG1 | E[ST]..$ (0.00e+00) | [ST].V$ (4.30e-13) | [ST].V$ (5.28e-13) | E[ST].V$ (1.40e-08) |
| | | | DLG2 | E[ST]..$ (0.00e+00) | [ST].V$ (2.87e-08) | [ST].V$ (2.87e-08) | E[ST]..$ (6.46e-06) |
| | | | DLG3 | [ST].[LV]$ (0.00e+00) | [ST].[LV]$ (8.82e-11) | [ST].[LV]$ (2.33e-09) | |
| | | | DLG4 | [ST].[LV]$ (0.00e+00) | [ST].[LV]$ (1.22e-15) | E[ST].V$ (1.11e-16) | [ST].V$ (4.00e-15) |
| | | | DLGAP1 | [ST].[LV]$ (0.004) | | | |
| | | | ERBB2IP | [ST].V$ (0.020) | | | [ST].V$ (0.011) |
| | | | GIPC1 | S.V$ (0.007) | | S.V$ (0.006) | |
| | | | GOPC | S.V$ (7.12e-05) | | S.V$ (5.23e-05) | |
| | | | INADL | [ST].[LV]$ (3.46e-11) | | [ST].V$ (2.78e-12) | |
| | | | MAGI1 | [ST].[LV]$ (1.16e-05) | T.[LV]$ (5.03e-06) | [ST].[LV]$ (4.15e-06) | |
| | | | MAGI2 | [ST].[LV]$ (3.03e-05) | | | |
| | | | MAGI3 | [ST].V$ (3.85e-09) | | [ST].V$ (3.85e-09) | |
| | | | PDZD3 | T.[FL]$ (9.06e-06) | | T.[FL]$ (9.06e-06) | n/a |
| | | | PDZK1 | [ST].[FLM]$ (1.02e-12) | [ST].L$ (3.50e-07) | [ST].[FL]$ (3.16e-10) | T.[FM]$ (2.81e-05) |
| | | | RGS3 | E[ST]..$ (0.014) | | E[ST]..$ (0.002) | |
| | | | SLC9A3R1 | T.L$ (0.00e+00) | T.[LM]$ (2.00e-07) | T.[LM]$ (9.33e-15) | [ST].[LM]$ (5.77e-15) |
| | | | SLC9A3R2 | T.[LM]$ (0.00e+00) | T.[LM]$ (2.79e-10) | T.[LM]$ (4.77e-15) | T.[LM]$ (5.31e-10) |
| | | | SNTA1 | E[ST]..$ (9.57e-13) | ET..$ (0.044) | E[ST]..$ (1.04e-07) | E[ST]..$ (0.029) |
| | | | SNTB1 | E[ST]..$ (4.94e-08) | n/a | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | SNTB2 | E..[IV]$ (1.67e-05) | | | |
| | | | SNTG1 | ET..$ (0.037) | n/a | ET..$ (0.023) | |
| | | | TJP1 | T[HR][LV]$ (1.67e-04) | | D..V$ (0.008) | |
| | | | USH1C | T.L$ (0.007) | T.L$ (0.002) | T.L$ (0.003) | |
| | | Domains | EBP50_C-term | T.L$ (0.00e+00) | T.[LM]$ (2.91e-14) | T.[LM]$ (0.00e+00) | [ST].[LM]$ (0.00e+00) |
| | | | PDZ | n/a | [ST].[FLMV]$ (3.25e-13) | [ST].[FILMV]$ (2.18e-14) | [ST].[FILMV]$ (2.40e-08) |
| LIG_PP1 | PP1 docking motif [RK].{0,1}[IV][^P][FW] (ELM LIG_PP1) | Gene | PPP1CA | K[KR]V.F (4.03e-06) | K[KR]V.F (1.25e-07) | K[KR]V.F (5.28e-08) | |
| LIG_PTB | NPxY Phopshotyrosine binding (PTB) motif (ELM LIG_PTB_1) | Gene | DAB2 | P..F.NP.Y (5.48e-04) | | | P..F.NP.Y (1.11e-04) |
| LIG_SH2 | SH2 domain ligand. Strongest YxN (GRB2) and YxxQ (STAT3) and YxxP (MnM) motifs only. Many other SH2 motifs might be part of the MOD_pY Generic motif. | Genes | GRAP2 | DY.N[LM] (0.007) | | | DY.N[LM] (1.90e-04) |
| | | | GRB2 | DY[MV]N (2.06e-04) | | | DY[MV]N (3.08e-04) |
| | | | LCK | | | | E..P.Y.NL (0.039) |
| | | | LCP2 | D[HY]EN (0.009) | n/a | n/a | |
| | | | PTPN1 | | n/a | | [DE]Y.N.S (0.010) |
| | | | SYK | YEN (8.11e-05) | n/a | | YEN (0.020) |
| | | Domain | SH2 | n/a | | | [DE]Y[MV]N (0.014) |
| LIG_SH3 | Canonical PxxP SH3 ligand motif. | Gene | FYN | P..P.[KR][KR] (0.005) | | | |
| LIG_WW_1 | PPxY WW ligand motif (ELM LIG_WW_1) | Domains | WW | PP.Y (0.042) | | PP.Y (0.027) | |
| MOD_CAAXbox | Generic CAAX box prenylation motif Cx[ILMV]x$ (ELM MOD_CAAXbox) | Gene | PDE6D | C.[IL].$ (3.33e-06) | | C.L.$ (0.045) | |
| | | Domain | GMP_PDE_delta | C.[IL].$ (0.001) | | | |
| MOD_CK1 | Sxx[ST] Motif recognised by CK1 for Ser/Thr phosphorylation (ELM | Genes | CAMK2G | S..[ST]..[WY]G (0.007) | n/a | | |
| | | | RPS6KA1 | | n/a | | S.HS.Y (0.010) |
| | | | SRC | S..T..TP (0.002) | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | MOD_CK1_1) | Domain | Pkinase_C | | | | D..[ST]I.S (0.013) |
| MOD_CK2 | CK2 phosphorylation motif. Core ELM [ST]..E motif expanded to [ST]..[DE] (MiniMotif Miner). | Genes | CSNK2A1 | SD.E (0.034) | | | |
| | | | MAPKAPK2 | R..S.[IV]E (0.001) | | | |
| | | Domain | PI3_PI4_kinase | SD.E (0.015) | | | SD.E (1.34e-04) |
| MOD_GSK3 | [ST]xxx[ST] Site recognised by GSK3 for Ser/Thr Phosphorylation (ELM MOD_GSK3_1) | Genes | GSK3B | | | | TP..[ST]P (0.032) |
| | | | PIK3R1 | | | | DSG..S (0.002) |
| MOD_PKB | R.R..[ST][^P] PKB Phosphorylation site (ELM MOD_PKB_1) | Genes | AKT1 | R.R..S (2.19e-06) | | | R.R[ST].S (0.004) |
| | | | SGK1 | R.R[ST].S (0.029) | | | |
| | | Domain | Pkinase_C | R.R..S (2.35e-05) | | | |
| MOD_PKC | PKC phosphorylation motif, [ST].[KR] [MnM] | Gene | PRKAB1 | | | A.GT.R (0.018) | |
| | | Domain | AMPKBI | | | A.GT.R (0.018) | |
| MOD_STP | Common recurring phosphorylation motif [ST]P | Genes | CDK2 | TPT (0.049) | | | |
| | | | GSK3B | P.SP (0.002) | | P.SP (0.029) | TP..[ST]P (0.032) |
| | | | MAPK14 | P.SP (0.004) | | | |
| | | | MAPK9 | T.SP.F.F (0.033) | | | |
| | | | NEK2 | | TPY (0.007) | TPY (0.036) | |
| | | | SRC | S..T..TP (0.002) | | | |
| | | | YWHAH | SP..SP (6.21e-05) | | | P.SP..[FY] (0.009) |
| | | | YWHAQ | SP..SP (0.009) | | | |
| | | Domains | 14-3-3 | P..[ST]P (7.12e-04) | | | P..[ST]P (0.009) |
| | | | Pkinase | n/a | P..[ST]P (0.002) | n/a | n/a |
| | | | Pkinase_C | P.[ST]P (5.09e-04) | | P.SP..[FWY] (0.001) | |
| MOD_SUMO | Canonical sumoylation motif, [AILMV]K.E (ELM MOD_SUMO) | Gene | SUMO1 | [ILV]K.E..[DE] (0.013) | | | |
| Yxx# | Multifunctional Yxx[ILMVF] motif, which includes ITAM, ITIM, ITSM, SH2 and endocytic | Genes | CSK | Y..V.K (0.025) | | | |
| | | | GRAP2 | DY.N[LM] (0.007) | | | DY.N[LM] (1.90e-04) |
| | | | GRB2 | DY[MV]N[MV] (0.010) | Y.EV[FH] (0.016) | Y.EV[FH] (0.001) | DY[MV]N[MV] (7.64e-04) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | targetting motifs. | | LCK | | | | E..P.Y.NL (0.039) |
| | | | PRKG1 | N..Y..[FY] (0.035) | | | |
| | | | PTK2B | E..Y..M..F (0.040) | | | |
| | | | PTPN11 | YA.[ILMV] (3.21e-05) | | | [ILV].YA.L (1.26e-04) |
| | | | PTPN6 | YA.[ILV] (7.39e-04) | [IL].YA.L (0.004) | [IL].YA.L (0.032) | |
| | | | RIPK1 | | | | Y..L.H..Q (0.029) |
| | | | ZAP70 | Y[DE].[LV] (0.009) | n/a | | |
| | | Domains | DSPc | NY..[MV]..H (0.007) | | | |
| | | | ITAM | | | | P.Y.N[FL] (0.042) |
| | | | SH2 | n/a | | | DY[MV]N[MV] (0.020) |

**Supplementary Table S2.** PCNA interactors returning SLiM predictions.

| Gene | Evidence | ppi[a] | bin[a] | com[a] | y2h[a] |
|---|---|---|---|---|---|
| ANXA2 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western | None | - | None | - |
| APEX1 | BioGrid:reconstituted complex;<br>HPRD:in vivo | None | - | D[FIL].N | - |
| APEX2 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo | None | None | [ILV]..[FWY]F<br>[ILV]..[FHY]F | - |
| BARD1 | IntAct:pull down | None | - | None | - |
| BAZ1B | HPRD:in vitro | None | - | - | - |
| CCNA2 | BioGrid:affinity capture-western;<br>DIP:coimmunoprecipitation | None | - | None | - |
| CCNB1 | HPRD:complex;<br>HPRD:in vivo | None | - | D[FIL].N | - |
| CCND1 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:complex;<br>HPRD:in vitro | None | - | None | - |
| CCND3 | BioGrid:reconstituted complex;<br>HPRD:in vitro | None | - | None | - |
| CCNO | BioGrid:affinity capture-ms;<br>BioGrid:reconstituted complex;<br>DIP:coimmunoprecipitation;<br>DIP:fluorescence microscopy;<br>DIP:two hybrid;<br>HPRD:in vitro;<br>HPRD:yeast 2-hybrid | None | None | None | None |
| CDAN1 | DIP:affinity chromatography technology;<br>DIP:biochemical;<br>DIP:fluorescence microscopy;<br>DIP:x-ray crystallography | None | - | - | - |
| CDC2 | HPRD:complex;<br>HPRD:in vitro | None | - | None | - |
| CDC25C | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | D[FIL].N | None |
| CDC6 | BioGrid:two-hybrid;<br>HPRD:yeast 2-hybrid | None | None | - | None |
| CDK2 | BioGrid:affinity capture-western;<br>DIP:affinity chromatography technology;<br>DIP:biophysical;<br>DIP:coimmunoprecipitation;<br>HPRD:complex;<br>HPRD:in vitro | None | - | None | - |
| CDK4 | BioGrid:affinity capture-western;<br>HPRD:complex | None | - | None | - |
| CDK5 | HPRD:complex;<br>HPRD:in vitro | None | - | None | - |
| CDK6 | HPRD:in vivo | None | - | - | - |
| CDKN1A | BioGrid:affinity capture-western;<br>BioGrid:co-crystal structure;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>DIP:coimmunoprecipitation;<br>DIP:cosedimentation through density gradient;<br>HPRD:complex; | None | None | None | None |

| | | | | | |
|---|---|---|---|---|---|
| | HPRD:in vitro;<br>IntAct:far western blotting;<br>IntAct:two hybrid pooling;<br>MINT:anti bait coimmunoprecipitation;<br>MINT:two hybrid | | | | |
| CDKN1C | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo | None | None | I..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F | None |
| CDKN2A | IntAct:anti bait coip;<br>IntAct:fluorescence imaging;<br>IntAct:pull down | None | - | None | - |
| CDT1 | HPRD:in vitro;<br>HPRD:in vivo;<br>MINT:anti bait coimmunoprecipitation;<br>MINT:gst pull down | None | - | [ILV]..[FWY]F<br>[ILV]..[FHY]F | - |
| CHAF1A | BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid;<br>IntAct:anti bait coip;<br>MINT:gst pull down | None | None | None | None |
| CHTF18 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | None | - |
| CLSPN | BioGrid:affinity capture-western | None | - | I..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F | - |
| CSTF1 | IntAct:pull down | None | - | None | - |
| DHX9 | HPRD:complex;<br>HPRD:in vitro | None | - | D[FIL].N | - |
| DNA2 | Reactome:reaction | None | - | - | - |
| DNMT1 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro | None | - | D[FIL].N<br>D[FIL].N<br>[ILV]..[FHY]F | - |
| DNTT | BioGrid:co-purification;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | None | None |
| DNTTIP2 | HPRD:in vitro | None | - | - | - |
| EIF1B | IntAct:anti bait coip | None | - | None | - |
| EP300 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | D[FIL].N | - |
| EPB41 | IntAct:anti bait coip | None | - | None | - |
| ERCC5 | BioGrid:reconstituted complex;<br>DIP:affinity chromatography technology;<br>DIP:biochemical;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | I..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F | - |
| FEN1 | BioGrid:affinity capture-western;<br>BioGrid:co-purification;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>DIP:affinity chromatography technology;<br>DIP:biochemical;<br>DIP:two hybrid;<br>HPRD:in vitro; | Q..[IL]..FF | None | Q..[IL]..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F<br>Q..L..FF | None |

| | | | | | |
|---|---|---|---|---|---|
| | HPRD:in vivo;<br>HPRD:yeast 2-hybrid;<br>IntAct:fret;<br>IntAct:two hybrid array;<br>IntAct:two hybrid pooling;<br>Reactome:direct_complex;<br>Reactome:reaction | | | | |
| FTSJ1 | IntAct:anti bait coip | None | - | None | - |
| GADD45A | BioGrid:affinity capture-western;<br>BioGrid:far western;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>DIP:coimmunoprecipitation;<br>DIP:two hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | None | None |
| GADD45B | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | None | None |
| GADD45G | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>DIP:affinity chromatography technology;<br>DIP:coimmunoprecipitation;<br>DIP:two hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | None | None |
| HDAC1 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | None | - |
| HLA-B | IntAct:anti bait coip | None | - | None | - |
| HLA-C | IntAct:anti bait coip | None | - | None | - |
| HLTF | DIP:anti tag coimmunoprecipitation;<br>DIP:pull down | None | - | None | - |
| HPRD13123 | HPRD:in vivo | - | - | - | - |
| HPRD16797 | HPRD:in vivo | - | - | - | - |
| HUS1 | BioGrid:affinity capture-western;<br>BioGrid:two-hybrid;<br>HPRD:yeast 2-hybrid | None | None | None | None |
| IKBKE | IntAct:anti bait coip | None | - | None | - |
| ING1 | BioGrid:affinity capture-western;<br>BioGrid:co-fractionation;<br>DIP:biochemical;<br>DIP:coimmunoprecipitation;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | None | - |
| KCTD13 | BioGrid:far western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | None | - |
| KIAA0101 | BioGrid:affinity capture-western;<br>BioGrid:two-hybrid | Q..[IL]..FF | None | Q..[IL]..FF<br>I..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F | None |
| LIG1 | BioGrid:affinity capture-ms;<br>HPRD:in vitro | Q..[IL]..FF<br>Q.[ST][IL]..FF | - | Q..[IL]..FF<br>Q.[ST][IL]..FF<br>I..FF<br>[ILV]..[FWY]F | - |

| | | | | | |
|---|---|---|---|---|---|
| | | | | [ILV]..[FHY]F | |
| MCC | IntAct:anti bait coip | None | - | D[FIL].N | - |
| MCL1 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | D[FIL].N | None |
| MME | IntAct:anti bait coip | None | - | None | - |
| MSH2 | BioGrid:affinity capture-ms;<br>HPRD:in vitro | None | - | None | - |
| MSH3 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>BioGrid:far western;<br>BioGrid:protein-peptide;<br>HPRD:in vitro;<br>HPRD:in vivo | Q..[IL]..FF | - | Q..[IL]..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F<br>Q..L..FF | - |
| MSH6 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>BioGrid:far western;<br>BioGrid:protein-peptide;<br>HPRD:in vitro;<br>HPRD:in vivo | Q..[IL]..FF<br>Q.[ST][IL]..FF<br>TL.SFF | - | Q..[IL]..FF<br>Q.[ST][IL]..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F<br>TL.SFF<br>Q..L..FF | - |
| MUTYH | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo | Q..[IL]..FF | - | Q..[IL]..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F<br>Q..L..FF | - |
| MYBBP1A | BioGrid:affinity capture-ms;<br>HPRD:in vitro | None | - | None | - |
| OTUB1 | BioGrid:affinity capture-ms | None | - | None | - |
| PARP1 | BioGrid:affinity capture-western;<br>HPRD:in vivo | None | - | D[FIL].N | - |
| PCNA | BioGrid:co-purification;<br>DIP:x-ray crystallography;<br>HPRD:complex;<br>HPRD:in vitro | None | - | None | - |
| POLA1 | HPRD:complex;<br>Reactome:reaction | None | - | None | - |
| POLA2 | Reactome:reaction | None | - | - | - |
| POLB | HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | - | None |
| POLD1 | BioGrid:affinity capture-ms;<br>BioGrid:invitro;<br>BioGrid:invivo;<br>HPRD:complex;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>IntAct:anti bait coip;<br>IntAct:far western blotting;<br>Reactome:reaction | None | - | None | - |
| POLD2 | BioGrid:far western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>IntAct:anti bait coip;<br>Reactome:reaction | None | - | None | - |
| POLD3 | BioGrid:affinity capture-ms;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>IntAct:anti tag coip;<br>IntAct:confocal microscopy;<br>IntAct:far western blotting;<br>IntAct:fret;<br>MINT:two hybrid;<br>Reactome:reaction | Q..[IL]..FF<br>Q.[ST][IL]..FF | None | Q..[IL]..FF<br>Q.[ST][IL]..FF<br>I..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F | None |

| | | | | | |
|---|---|---|---|---|---|
| POLD4 | HPRD:in vitro;<br>HPRD:yeast 2-hybrid;<br>IntAct:elisa;<br>IntAct:far western blotting;<br>Reactome:reaction | None | None | - | None |
| POLDIP2 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | None | None |
| POLE | BioGrid:affinity capture-ms;<br>HPRD:in vitro;<br>Reactome:reaction | None | - | D[FIL].N | - |
| POLE2 | Reactome:reaction | None | - | - | - |
| POLH | BioGrid:reconstituted complex;<br>BioGrid:two-hybrid;<br>HPRD:in vitro;<br>HPRD:yeast 2-hybrid;<br>MINT:gst pull down;<br>MINT:two hybrid | TL.SFF | None | [ILV]..[FWY]F<br>[ILV]..[FHY]F<br>TL.SFF | None |
| POLI | MINT:gst pull down;<br>MINT:two hybrid | None | None | [ILV]..[FWY]F<br>[ILV]..[FHY] | None |
| POLL | BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | None | - |
| POLM | HPRD:in vitro | None | None | - | - |
| PPP1CC | IntAct:antibody array | None | - | - | - |
| PRIM1 | Reactome:reaction | None | - | - | - |
| PRIM2 | HPRD:complex;<br>Reactome:reaction | None | - | None | - |
| PRKDC | BioGrid:affinity capture-ms;<br>HPRD:in vitro | None | - | None | - |
| PTMA | HPRD:in vitro | None | - | - | - |
| PTP4A3 | IntAct:anti bait coip | None | - | None | - |
| RAD17 | DIP:fluorescence microscopy | None | - | - | - |
| RAD18 | HPRD:complex | None | - | None | - |
| RAD9A | DIP:coimmunoprecipitation;<br>HPRD:in vitro | None | - | D[FIL].N | - |
| RELA | IntAct:tap;<br>MINT:anti tag coimmunoprecipitation | None | - | None | - |
| RFC1 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:complex;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>Reactome:reaction | None | - | I..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F | - |
| RFC2 | BioGrid:affinity capture-ms;<br>BioGrid:reconstituted complex;<br>HPRD:complex;<br>HPRD:in vitro;<br>Reactome:reaction | None | - | None | - |
| RFC3 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>HPRD:complex;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>Reactome:reaction | None | - | None | - |
| RFC4 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>DIP:coimmunoprecipitation;<br>HPRD:complex;<br>HPRD:in vitro; | None | - | None | - |

| | | | | | |
|---|---|---|---|---|---|
| | HPRD:in vivo;<br>Reactome:reaction | | | | |
| RFC5 | BioGrid:affinity capture-ms;<br>BioGrid:reconstituted complex;<br>HPRD:complex;<br>HPRD:in vitro;<br>Reactome:reaction | None | - | None | - |
| RPA1 | BioGrid:affinity capture-ms;<br>HPRD:complex;<br>HPRD:in vitro;<br>Reactome:reaction | None | - | D[FIL].N<br>[ILV]..[FWY]F | - |
| RPA2 | Reactome:reaction | None | - | - | - |
| RPA3 | Reactome:reaction | None | - | - | - |
| RUVBL2 | BioGrid:affinity capture-ms | None | - | None | - |
| SEC23IP | HPRD:in vitro | None | - | - | - |
| SMARCA4 | BioGrid:co-fractionation | None | - | - | - |
| SUB1 | HPRD:yeast 2-hybrid | None | None | - | None |
| TCOF1 | HPRD:in vitro | None | - | - | - |
| TIMELESS | BioGrid:affinity capture-western | None | - | None | - |
| TOP1 | HPRD:complex | None | - | None | - |
| TRAF6 | IntAct:anti bait coip | None | - | None | - |
| TSC22D1 | IntAct:anti bait coip | None | - | None | - |
| UBA52 | DIP:anti bait coimmunoprecipitation;<br>DIP:anti tag coimmunoprecipitation;<br>DIP:comigration in sds page;<br>MINT:his pull down | None | - | None | - |
| UBB | DIP:anti bait coimmunoprecipitation;<br>DIP:anti tag coimmunoprecipitation;<br>DIP:comigration in sds page;<br>HPRD:in vivo;<br>MINT:his pull down | None | - | None | - |
| UBC | BioGrid:affinity capture-western;<br>DIP:anti bait coimmunoprecipitation;<br>DIP:anti tag coimmunoprecipitation;<br>DIP:comigration in sds page;<br>MINT:his pull down | None | - | None | - |
| UBE2B | HPRD:complex | None | - | None | - |
| UNG | BioGrid:affinity capture-ms;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro | Q..[IL]..FF<br>Q.[ST][IL]..FF<br>TL.SFF | - | Q..[IL]..FF<br>Q.[ST][IL]..FF<br>[ILV]..[FWY]F<br>[ILV]..[FHY]F<br>TL.SFF<br>Q..L..FF | - |
| WRN | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>MINT:fluorescence microscopy | None | - | [ILV]..[FWY]F<br>[ILV]..[FHY]F | - |
| XRCC1 | BioGrid:affinity capture-western;<br>BioGrid:co-localization;<br>BioGrid:fret;<br>BioGrid:reconstituted complex;<br>HPRD:in vitro;<br>HPRD:in vivo;<br>HPRD:yeast 2-hybrid | None | None | D[FIL].N | None |
| XRCC5 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | None | - |
| XRCC6 | BioGrid:affinity capture-ms;<br>BioGrid:affinity capture-western;<br>HPRD:in vitro;<br>HPRD:in vivo | None | - | None | - |
| YBX1 | BioGrid:affinity capture-western;<br>BioGrid:reconstituted complex; | None | - | None | - |

HPRD:in vitro;
HPRD:in vivo

a. None, no significant motif returned in this spoke; -, spoke not in dataset.

**Supplementary Table S3.** MCL clusters of TP and novel motifs returned from Real datasets, connected by CompariMotif similarity.

| MCL[a] | Rank[b] | Nodes[c] | Edges[c] | Density[c] | Members[d] | Novel[e] | TP[e] |
|---|---|---|---|---|---|---|---|
| Real | 1 | 82 | 2309 | 0.348 | C.K.K, FR.R..S, FR.R[ST].S, K.[KR]..A..CR, K.[KR]..A..[KR], L.R..S, L.RS.S[FM], L.R[ST].S[MV], L.[HR][ST].S, Q[KR].R, R..S.P, R..S.S, R..S.[IV]E, R..SAP, R.NS, R.R..S, R.R.R, R.RH.S, R.RS, R.RS.S, R.RT.S, R.R[ST].S, R.R[ST].[ST], RS.S, RS.S..[FH], RS.S.P, RS.S[LMV], RSR[ST].S, RT.S, R[ST].S..[DE], R[ST].S.P, R[ST].SAP, R[ST].S[FLMV], R[ST].S[IMV], R[ST].[ST][FLMV], R[ST]R[ST].S, R[ST]R[ST].[ST], R[ST][HR][ST].S, [FHY][HR].R..S, [FH]QRS.S, [FH]R.R..S, [FH][KR].R..[ST], [FLM].[HR][ST].S, [FL]R.R..S, [FY][HR].R..S, [HKR].RS.S, [HKR].RS.[ST], [HKR].R[ST].S, [HKR].R[ST].S[FLV], [HKR].[HKR][ST].[ST], [HKR][ST].S.[ST], [HKR][ST].[ST], [HKR][ST].[ST].P, [HR].R..S..[DE], [HR].R..[ST].P, [HR].R[ST].S, [HR].R[ST].[ST], [HR].R[ST][HKR]S, [HR].[HKR][ST].S, [HR].[HR][ST].S, [HR][ST]..S, [HR][ST].[ST], [HR][ST]R[ST].S, [ILMV].[HKR][ST].[ST], [ILM].[HKR][ST].S, [ILV].[KR][ST].[ST], [IL].[KR][ST].S, [IL].[KR][ST].S.P, [IL][HR][KR][ST].S, [KR]N[HK], [KR]N[KR], [KR]S.S..D, [KR][ST].S.P, [KR][ST].S[LMV], [KR][ST].[ST], [KR][ST][KR][ST], [LMV].[HR][ST].S, [LM].[HKR][ST].S, [LM].[HKR][ST].[ST], [LM].[HR]..S.P, [LM].[HR][ST].S, [ST].SAP | 7 | 75 |
| Real | 2 | 74 | 1131 | 0.209 | D..L$, D..V$, D..[LM]$, DT..$, DT.L$, DT.[LM]$, E..T..$, E..T.L$, E..V$, E..[ILV]$, E..[IV]$, E..[LV]$, E.T..$, E.T.L$, EDT..$, ES..$, ES.V$, ET..$, ET.V$, E[ST]..$, E[ST].V$, E[ST].[LV]$, K.T.L$, L.T..$, L.T.V$, L.[ST]..$, L.[ST].V$, LE..V$, R.[ST]..$, R.[ST].V$, S.IE..V$, S.V$, T.L$, T.M$, T.V$, T.[FL]$, T.[FM]$, T.[LM]$, T.[LV]$, TRL$, TT.V$, T[HKR].$, T[HKR][LV]$, T[HR].$, T[HR][LV]$, T[KR].$, T[KR]L$, T[ST].[LV]$, V..T..$, [DE]..[LM]$, [DE]S..$, [DE][ST]..$, [DE][ST].[ILMV]$, [DE][ST].[ILV]$, [DE][ST].[LMV]$, [DE][ST].[LV]$, [ILV].[ST].V$, [IL].[ST]..$, [IL].[ST].V$, [IL]E..V$, [KR]..T.V$, [KR].T..$, [KR].[ST]..$, [KR].[ST].[LV]$, [KR]L$, [LV].[ST].V$, [ST].L$, [ST].V$, [ST].[FILMV]$, [ST].[FLMV]$, [ST].[FLM]$, [ST].[FL]$, [ST].[LM]$, [ST].[LV]$ | 0 | 74 |
| Real | 3 | 37 | 400 | 0.300 | A.EF.P, G..WA..H, G.A..[FH], HL..IA, KN..AA, L..NA..F, L..NA..F.P, L..NA.EF, L..NAK.F, LN..A..F, LN..A..F.P, LN.NA, LN.NA..F, M.LL..N, N..A..F.P, N..A.EF, N..A.EF.P, N..AA, N.NA..F, N.NA..F.P, NA..F, NA..F.P, NA.EF, NP.A..F, NP.A..F.P, N[FL].AL, S..NPNA, S.L..N..EF, S.L..NA, S.L..NA..F, S.L..NA.E, S.L.PNA, S.LN.NA, V..[IV]LN, [FLV]NP.A, [ILMV].[DE]N, [ST].[IL]N.N | 10 | 27 |
| Real | 4 | 20 | 138 | 0.363 | FADF.S, K.V.F, K[KR]V.F, L.F.D, N..V.[FH], RP.T[FV].[IL], RV.F, RV.[FW]A, V.[FIV]C, VTF, VTF..[DE], V[ST]F.[DE], [KR].V.F.[DE], [KR].V[ST]F, [KR].[ILV].F, [KR].[ILV].[FW], [KR]K.V.F, [KR]V[ST]F, [KR][KR]..F, [LV].[FW].[DE] | 3 | 17 |
| Real | 5 | 39 | 251 | 0.169 | A.Q.A, A.[ILM]A, A.[LV].D[FL], A.[LV].[FH], AA.[KR], AA[FILMV], AA[FILV], AA[ILV], AM..A, A[FILMV]..A, A[FIV]..C, A[ILMV]..A, A[IM]..A, A[LV]..[FHY], CKA, H..[FY].K, I..A.E, M[FY]..L, NY..[MV]..H, Q..A.L, R..[FLV]..A, Y.EV[FH], YA.[ILMV], YA.[ILV], YA[ST][LV], [DE].[FILMV]..A, [DE][HY]..F..A, [FILV]..A[LV], [FIV]E.A, [FLMV]..A..A, [FLMV].[HKR]A, [FLV]..ALQ, [ILV].YA.L, [ILV].YA.[IL], [IL].EA[LMV], [IL].YA.L, [IV].Y..L.[FH], [IV].Y..L.[HK], [LV]..A.K | 29 | 10 |
| Real | 6 | 17 | 69 | 0.254 | D..[IM].F.L, G..G..F.L..E, G..G.GF.L, G..V.L.C, G.GF.L..E, GFG[FL], GFHL, H.LV, KG..G..F, KG..G..F.L, KG..G.G..L, KG..G.GF, W..G[FL], [DE]..[LM].[FH].L, [DE].M..GL.D, [MV].[FHY].L.[DE], [MV].[FH].L.[DE] | 17 | 0 |
| Real | 7 | 16 | 61 | 0.254 | D..N..[FL], D..N[FI], DY.N[LMV], DY.N[LM], DY[MV]N, DY[MV]N[MV], D[HY]EN, EK.YE, EN[LMV], YEN, YEN[LV], YVNV, Y[DE].[LV], [DE]Y.N.S, [FIL].D..N, [KR]EK[LM] | 6 | 10 |
| Real | 8 | 19 | 73 | 0.213 | A.Q.L, A.Q.[LV], ALWD, A[FIV]..L, A[FIV]..[LV], E..V..L.K.L, K..E[LV].D, L.D[LM]..D, L.[DE]A, LA.Q, Q.[FILMV].[DE], Q.[FLV].[DE], | 19 | 0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | Q.[FL]..NN, Q.[ILV]N, Q[FY][LV].D, Q[KR]..[DE], [ILMV].N..D, [MV]Q.L.D, [MV]Q.L.D.L | | |
| Real | 9 | 10 | 30 | 0.333 | P..[ST]P, P.SP, P.SP..[FWY], P.SP..[FY], P.[ST]P, S..T..TP, SP..SP, SP..[ST]P, TPY, [ST]P..[ST]P..[ST] | 0 | 10 |
| Real | 10 | 12 | 36 | 0.273 | D..[IL]DL, D[FIL].N, I..[KR][LM]..[DE], PLDL, P[ILMV]DL, P[ILM]DL, P[IL]DL, [DE][FIL]..Q[LV], [DE][ILMV]..[HKR][FILV], [DE][ILV]A..L, [DE][IL].A, [IL].NM | 7 | 5 |
| Real | 11 | 24 | 74 | 0.134 | D.F.VK, D.I..[FIL], D.I..[FLV], DI..I..[LV], E..L..W, G[IM][FY], I..FF, I.E.[FL], I.EL, LN..[FI], LQ..L, M..Q..[IL], Q..F.C, Q..I[FM], Q..[IL]..FF, Q.[ST][IL]..FF, QQ.[FILMV], RQ..[FLV], [DE]VI, [FI]G.[FV]D, [ILV]..[FWY]F, [IL]N..[FV], [IV].[DE].F, [KR]..[DE].Q..L | 20 | 4 |
| Real | 12 | 16 | 43 | 0.179 | F.N.R, F.NP.Y, INPD, N.[KR]I, NPF, NP[FV], NP[FY], P..F.N, P..F.NP.Y, S..NP[FV], TNPF, WF.N.R, [FM]..NI, [FY].FN.K, [FY].N..Y, [FY].[WY]..N | 8 | 8 |
| Real | 13 | 7 | 13 | 0.310 | E.LA, EE..R.L, EE.E.[FLV], EE.E.[FV], R..VE.[LM], [FH]EE[LV], [IV]..EV[DE] | 6 | 1 |
| Real | 14 | 4 | 6 | 0.500 | L..EN..L, L..EN..[IL], N..L..[DE], [LM].N..[FLV] | 4 | 0 |
| Real | 15 | 8 | 13 | 0.232 | A..[ILMV].Q, A..[IM]..Q, G.VLA.M, LA..I, N..A..[IV]L, Q.A..M, RA.[ST][LMV], [IV]..QG | 7 | 1 |
| Real | 16 | 5 | 7 | 0.350 | EK[FI].R.[HK], K..RWT, P..K..[KR], P..P.[KR][KR], P.K..R | 4 | 1 |
| Real | 17 | 8 | 11 | 0.196 | D.QL, K..I.D, Q[FILV]..LQ, Q[LM]R, R.IRD, R.Q.R..R, [DE]E.L.Q[IM], [HR].[IL]R | 8 | 0 |
| Real | 18 | 3 | 3 | 0.500 | D.G[FL].C, DMG, V.D.G | 3 | 0 |
| Real | 19 | 3 | 3 | 0.500 | DRL..V, DRL..V.G, LQ.V | 3 | 0 |
| Real | 20 | 3 | 3 | 0.500 | H[HY]H.P, H[LV][HR].H, [FH]HH.P | 3 | 0 |
| Real | 21 | 4 | 4 | 0.333 | M.LL[KR].N, [HK].N.W, [KR]..GW, [KR].N..[KR] | 4 | 0 |
| Real | 22 | 4 | 4 | 0.333 | HWK, HWKQ, P.HW, [HY][FWY]D | 4 | 0 |
| Real | 23 | 5 | 5 | 0.250 | D[KR].L.I.G, Q..G..VNG, VNG, [FM].[IM]N, [IV]NG | 5 | 0 |
| Real | 24 | 5 | 5 | 0.250 | F.[HK]..H[IV], LE..R..[IL], R.G..R, R.GF.[KR], [KR]..[FLM].R | 5 | 0 |
| Real | 25 | 3 | 2 | 0.333 | D[IM].F..F, E..E..DF..L, E..[DE].M.F..[FL] | 3 | 0 |
| Real | 26 | 3 | 2 | 0.333 | E.DI.V, SD.E, [DE].[DE]..[MV]..G | 2 | 1 |
| Real | 27 | 3 | 2 | 0.333 | L..K..Q.E, L.KM.Q, [DE]..HL..KE | 3 | 0 |
| Real | 28 | 2 | 1 | 0.500 | VAR, [MV]VA | 2 | 0 |
| Real | 29 | 2 | 1 | 0.500 | ERG.K, P..T.V..E.G | 2 | 0 |
| Real | 30 | 2 | 1 | 0.500 | S..[ST]..[WY]G, S.HS.Y | 0 | 2 |
| TP | 1 | 78 | 2234 | 0.372 | FR.R..S, FR.R[ST].S, L.R..S, L.RS.S[FM], L.R[ST].S[MV], L.[HR][ST].S, R..S.P, R..S.S, R..S.[IV]E, R..SAP, R.NS, R.R..S, R.RH.S, R.RS, R.RS.S, R.RT.S, R.R[ST].S, R.R[ST].[ST], RA.[ST][LMV], RS.S, RS.S..[FH], RS.S.P, RS.S[LMV], RSR[ST].S, RT.S, R[ST].S..[DE], R[ST].S.P, R[ST].SAP, R[ST].S[FLMV], R[ST].S[IMV], R[ST].[ST][FLMV], R[ST]R[ST].S, R[ST]R[ST].[ST], R[ST][HR][ST].S, S..[ST]..[WY]G, S.HS.Y, [FHY][HR].R..S, [FH]QRS.S, [FH]R.R..S, [FH][KR].R..[ST], [FLM].[HR][ST].S, [FL]R.R..S, [FY][HR].R..S, [HKR].RS.S, [HKR].RS.[ST], [HKR].R[ST].S, [HKR].R[ST].S[FLV], [HKR].[HKR][ST].[ST], [HKR][ST].S.[ST], [HKR][ST].[ST], [HKR][ST].[ST].P, [HR].R..S..[DE], [HR].R..[ST].P, [HR].R[ST].S, [HR].R[ST].[ST], [HR].R[ST][HKR]S, [HR].[HKR][ST].S, [HR].[HR][ST].S, [HR][ST]..S, [HR][ST].[ST], [HR][ST]R[ST].S, [ILMV].[HKR][ST].[ST], [ILM].[HKR][ST].S, [ILV].[KR][ST].[ST], [IL].[KR][ST].S, [IL].[KR][ST].S.P, [IL][HR][KR][ST].S, [KR]S.S..D, [KR][ST].S.P, [KR][ST].S[LMV], [KR][ST].[ST], [KR][ST][KR][ST], [LMV].[HR][ST].S, [LM].[HKR][ST].S, [LM].[HKR][ST].[ST], [LM].[HR]..S.P, [LM].[HR][ST].S, [ST].SAP | 0 | 78 |
| TP | 2 | 74 | 1131 | 0.209 | D..L$, D..V$, D..[LM]$, DT..$, DT.L$, DT.[LM]$, E..T..$, E..T.L$, E..V$, E..[ILV]$, E..[IV]$, E..[LV]$, E.T..$, E.T.L$, EDT..$, ES..$, ES.V$, ET..$, ET.V$, E[ST]..$, E[ST].V$, E[ST].[LV]$, K.T.L$, L.T..$, L.T.V$, L.[ST]..$, L.[ST].V$, LE..V$, R.[ST]..$, R.[ST].V$, S.IE..V$, S.V$, T.L$, T.M$, T.V$, T.[FL]$, T.[FM]$, T.[LM]$, T.[LV]$, TRL$, TT.V$, T[HKR].$, T[HKR][LV]$, T[HR].$, T[HR][LV]$, T[KR].$, T[KR]L$, T[ST].[LV]$, V..T..$, [DE]..[LM]$, [DE]S..$, [DE][ST]..$, [DE][ST].[ILMV]$, [DE][ST].[ILV]$, [DE][ST].[LMV]$, [DE][ST].[LV]$, [ILV].[ST].V$, [IL].[ST]..$, [IL].[ST].V$, [IL]E..V$, [KR]..T.V$, [KR].T..$, [KR].[ST]..$, [KR].[ST].[LV]$, [KR]L$, [LV].[ST].V$, [ST].L$, [ST].V$, [ST].[FILMV]$, [ST].[FLMV]$, [ST].[FLM]$, [ST].[FL]$, [ST].[LM]$, [ST].[LV]$ | 0 | 74 |
| TP | 3 | 27 | 301 | 0.429 | A.EF.P, L..NA..F, L..NA..F.P, L..NA.EF, L..NAK.F, LN..A..F, LN..A..F.P, LN.NA, LN.NA..F, N..A..F.P, N..A.EF, N..A.EF.P, N.NA..F, N.NA..F.P, NA..F, NA..F.P, NA.EF, NP.A..F, NP.A..F.P, S..NPNA, S.L..N..EF, S.L..NA, | 0 | 27 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | S.L..NA..F, S.L..NA.E, S.L.PNA, S.LN.NA, [ST].[IL]N.N | | |
| TP | 4 | 18 | 108 | 0.353 | EE..R.L, FADF.S, K.V.F, K[KR]V.F, RP.T[FV].[IL], RV.F, RV.[FW]A, VTF, VTF..[DE], V[ST]F.[DE], [KR].V.F.[DE], [KR].V[ST]F, [KR].[ILV].F, [KR].[ILV].[FW], [KR]K.V.F, [KR]V[ST]F, [KR][KR]..F, [LV].[FW].[DE] | 0 | 18 |
| TP | 5 | 12 | 43 | 0.326 | DY.N[LMV], DY.N[LM], DY[MV]N, DY[MV]N[MV], D[HY]EN, SD.E, YEN, YEN[LV], YVNV, Y[DE].[LV], [DE]Y.N.S, [FY].N..Y | 0 | 12 |
| TP | 6 | 10 | 33 | 0.367 | NY..[MV]..H, Y.EV[FH], YA.[ILMV], YA.[ILV], YA[ST][LV], [ILV].YA.L, [ILV].YA.[IL], [IL].YA.L, [IV].Y..L.[FH], [IV].Y..L.[HK] | 0 | 10 |
| TP | 7 | 7 | 21 | 0.500 | F.NP.Y, NPF, NP[FV], NP[FY], P..F.NP.Y, S..NP[FV], TNPF | 0 | 7 |
| TP | 8 | 11 | 32 | 0.291 | P..P.[KR][KR], P..[ST]P, P.SP, P.SP..[FWY], P.SP..[FY], P.[ST]P, S..T..TP, SP..SP, SP..[ST]P, TPY, [ST]P..[ST]P..[ST] | 0 | 11 |
| TP | 9 | 5 | 10 | 0.500 | D..[IL]DL, PLDL, P[ILMV]DL, P[ILM]DL, P[IL]DL | 0 | 5 |
| TP | 10 | 4 | 6 | 0.500 | I..FF, Q..[IL]..FF, Q.[ST][IL]..FF, [ILV]..[FWY]F | 0 | 4 |
| Novel | 1 | 20 | 100 | 0.263 | A.Q.A, AM..A, A[FILMV]..A, A[ILMV]..A, A[IM]..A, CKA, F.[HK]..H[IV], H..[FY].K, I..A.E, L.[DE]A, R..[FLV]..A, [DE].[FILMV]..A, [DE][HY]..F..A, [FILV]..A[LV], [FIV]E.A, [FLMV]..A..A, [FLMV].[HKR]A, [FLV]..ALQ, [IL].EA[LMV], [LV]..A.K | 20 | 0 |
| Novel | 2 | 11 | 37 | 0.336 | G..G..F.L..E, G..G.GF.L, G..V.L.C, G.GF.L..E, GFG[FL], GFHL, KG..G..F, KG..G..F.L, KG..G.G..L, KG..G.GF, W..G[FL] | 11 | 0 |
| Novel | 3 | 27 | 89 | 0.127 | A.Q.L, A.Q.[LV], A[FIV]..L, A[FIV]..[LV], D.F.VK, D.G[FL].C, G[IM][FY], K..E[LV].D, L.D[LM]..D, LA.Q, LQ..L, M..Q..[IL], Q..F.C, Q..I[FM], Q.[FILMV].[DE], Q.[FLV].[DE], Q.[FL]..NN, QQ.[FILMV], Q[FY][LV].D, Q[KR]..[DE], RQ..[FLV], V.D.G, [FIL].D..N, [IV].[DE].F, [KR]..[DE].Q..L, [MV]Q.L.D, [MV]Q.L.D.L | 27 | 0 |
| Novel | 4 | 17 | 50 | 0.184 | A.[ILM]A, A.[LV].D[FL], A.[LV].[FH], AA.[KR], AA[FILMV], AA[FILV], AA[ILV], ALWD, A[FIV]..C, A[LV]..[FHY], G.VLA.M, HL..IA, L.F.D, N..AA, N[FL].AL, Q..A.L, [FLV]NP.A | 17 | 0 |
| Novel | 5 | 8 | 18 | 0.321 | C.K.K, K.[KR]..A..CR, K.[KR]..A..[KR], KN..AA, Q[KR].R, R.R.R, [KR]N[HK], [KR]N[KR] | 8 | 0 |
| Novel | 6 | 10 | 23 | 0.256 | D..[IM].F.L, E..L..W, H.LV, M[FY]..L, N..V.[FH], V.[FIV]C, [DE]..[LM].[FH].L, [DE].M..GL.D, [MV].[FHY].L.[DE], [MV].[FH].L.[DE] | 10 | 0 |
| Novel | 7 | 7 | 14 | 0.333 | E.LA, EE.E.[FLV], EE.E.[FV], EN[LMV], R..VE.[LM], [FH]EE[LV], [IV]..EV[DE] | 7 | 0 |
| Novel | 8 | 11 | 23 | 0.209 | D..N..[FL], L..EN..L, L..EN..[IL], M.LL..N, M.LL[KR].N, N..L..[DE], [HK].N.W, [ILMV].[DE]N, [KR]..GW, [KR].N..[KR], [LM].N..[FLV] | 11 | 0 |
| Novel | 9 | 6 | 10 | 0.333 | D.I..[FIL], D.I..[FLV], DI..I..[LV], I.EL, [DE]VI, [FI]G.[FV]D | 6 | 0 |
| Novel | 10 | 4 | 6 | 0.500 | Q..G..VNG, Q.[ILV]N, VNG, [IV]NG | 4 | 0 |
| Novel | 11 | 6 | 9 | 0.300 | A..[ILMV].Q, A..[IM]..Q, LA..I, N..A..[IV]L, Q.A..M, [IV]..QG | 6 | 0 |
| Novel | 12 | 4 | 5 | 0.417 | EK[FI].R.[HK], K..RWT, P..K..[KR], P.K..R | 4 | 0 |
| Novel | 13 | 9 | 13 | 0.181 | D.QL, D[KR].L.I.G, K..I.D, Q[FILV]..LQ, Q[LM]R, R.IRD, R.Q.R..R, [DE]E.L.Q[IM], [HR].[IL]R | 9 | 0 |
| Novel | 14 | 8 | 11 | 0.196 | D[FIL].N, F.N.R, N.[KR]I, P..F.N, WF.N.R, [FM]..NI, [FY].FN.K, [FY].[WY]..N | 8 | 0 |
| Novel | 15 | 7 | 9 | 0.214 | INPD, LN..[FI], V..[IV]LN, [FM].[IM]N, [ILMV].N..D, [IL].NM, [IL]N..[FV] | 7 | 0 |
| Novel | 16 | 3 | 3 | 0.500 | DRL..V, DRL..V.G, LQ.V | 3 | 0 |
| Novel | 17 | 3 | 3 | 0.500 | H[HY]H.P, H[LV][HR].H, [FH]HH.P | 3 | 0 |
| Novel | 18 | 7 | 9 | 0.214 | E.DI.V, I..[KR][LM]..[DE], [DE].[DE]..[MV]..G, [DE][FIL]..Q[LV], [DE][ILMV]..[HKR][FILV], [DE][ILV]A..L, [DE][IL].A | 7 | 0 |
| Novel | 19 | 4 | 4 | 0.333 | LE..R..[IL], R.G..R, R.GF.[KR], [KR]..[FLM].R | 4 | 0 |
| Novel | 20 | 4 | 4 | 0.333 | HWK, HWKQ, P.HW, [HY][FWY]D | 4 | 0 |
| Novel | 21 | 4 | 3 | 0.250 | DMG, D[IM].F..F, E..E..DF..L, E..[DE].M.F..[FL] | 4 | 0 |
| Novel | 22 | 4 | 3 | 0.250 | E..V..L.K.L, L..K..Q.E, L.KM.Q, [DE]..HL..KE | 4 | 0 |
| Novel | 23 | 2 | 1 | 0.500 | VAR, [MV]VA | 2 | 0 |
| Novel | 24 | 2 | 1 | 0.500 | EK.YE, [KR]EK[LM] | 2 | 0 |
| Novel | 25 | 2 | 1 | 0.500 | G..WA..H, G.A..[FH] | 2 | 0 |
| Novel | 26 | 2 | 1 | 0.500 | ERG.K, P..T.V..E.G | 2 | 0 |

a. MCL clustering run: Real, TP and novel; TP, TP only; Novel, novel only.

b. Rank of cluster based on Score = Node number x Density.

c. Number of Nodes (N) and Edges (E) in cluster. Density (D) is calculated as D = E / N(N-1).

d. Patterns of predicted SLiMs in cluster.

e. Numbers of Novel and TP SLiMs found in each cluster.