**Supplementary Analysis and Methods.**

**1. PPID extraction**

The 15 columns of data required by the PSI-MI (Henning et al., 2004; Kerrien et al., 2007)[67,68] standard contained seven unique identifiers we used; interactor A and interactor B, interaction detection method identifiers, publication identifier (PubMed Id), tax ID of interactor A and interactor B and the interaction type identifier. The category of 'interaction detection method' and 'interaction type' are defined with Molecular Interaction (MI)-terms specified by the MI ontology (http://www.ebi.ac.uk/ols/index). The terms are represented in hierarchical relations, from most generic (e.g. interaction detection method (MI:0001): experimental interaction detection (MI:0045)) through more specific categories (e.g. interaction detection method, interactome parallel affinity capture (MI:0963)).

IntAct also provides information if a pair of interactors was detected as parts of a single complex, or a larger complex. Direct interaction information, can result from such experimental protocols as tandem affinity purification (TAP). When identifying larger protein complexes n-ary interactions are translated into binary ones. Therefore IntAct uses the spoke-model, pairing a “bait” protein with each of its “prey”. In this way, a single complex is spoke-expanded into multiple interactions. This procedure is known to generate false positive information5. ‘Raw-data’ was merged and pre-filtered for only human PPIs (meaning that both interactors were associated to the human taxID (9606)). Data was only kept if all seven columns of interest were non-empty. To obtain the most consistent dataset possible, spoke-expanded interactions from IntAct were removed.

To obtain only direct interactions, the 'interaction type' column was then filtered for the “association” (MI:0914), “physical association” (MI:0915) and 'direct interaction' term (MI:0407) and its 63 child-terms. Some of the source data uses an obsolete interaction type MI:0218, “physical interaction” which was also included, since it was updated to association and physical association, which we both include. PPIs based on the interaction types: “genetic interaction” (MI:0208) (including “suppression” (MI:0796) and “synthetic” (MI:0794)), “colocalization” (MI:0403), “genetic interference” (MI:0254) and “additive genetic interaction defined by inequality” (obsolete term, MI:0799) no not explicitly support direct interactions. Interactions only based on such evidence were excluded from the final set.

The pre-filtered, human PPI dataset contains 284679 interactions (both interactors are associated to the human taxid:9606). These include mirrored duplicates (a-b and b-a listed as two interactions), since their annotation order differed in distinct databases and depends on information supplied in additional columns. Since biological processes, specifically protein-protein-interactions are non-directional interactor IDs were sorted to obtain unique counts. The number of (1) unique non-mirrored interactions (a-b and b-a as one unique interaction), (2) associated interaction detection methods, (3) the interaction types and (4) reference publications can be seen in Table 1.

**2 PPI networks and analysis.**

Degree distribution in many real world networks is often thought to be scale-free (Barabasi 1999) [43]. However, few real world networks follow a power-law for all values in the distribution, more often distributions are heavy-tailed, with the power-law only applying to values greater than some minimum (Clauset et al., 2009)[60]. We therefore tested for scale-free structure in the PPI proteome networks using the R “PoweRlaw” package (version 0.50.0) (Gillespie, 2015)[61]. The package makes use of the power-law distribution for discrete data (Clauset et al., 2009)[60], which is dependent on the scaling parameter 'α' (typically lying in the range 2 <= α <= 3) and the minimum degree 'kmin' (with kmin >= 0). Since it is possible to fit a power-law to any data, a 'goodness-of-fit' test is needed to test the plausibility of fitting the power-law to the data. We used a bootstrap procedure (1000 iterations), based on the Kolmogorov-Smirnov (KS) test statistic as described in (Clauset et al., 2009; Gillespie, 2015)[60,61]. This test generates a P-value that can be used to quantify the plausibility of fit. If the P-value is ~ 0, then a power-law is unlikely to provided a plausible fit to the network data. If the P-value is large (that is to say greater than 0.1 and tending to 1.0) then the fit is possibly true, with any difference between the network data and the model probably due to statistical fluctuations (Clauset et al., 2009)[60]. We find evidence for scale-free structure in the PSP (P = 0.23), the presynaptic (P = 0.3) and PSP reduced (P = 0.33) PPI networks. Figures 3A, 3C and 3E show the log-log plots of the CDF of each PPI networks degree distribution (P(k)), versus its degree (k), with the best fitting power-law distribution to the network data highlighted in red.

We compared each networks degree and scale-free structure relative against two randomised network models: the Eros-Renyi (ER) model, and one which follows a scale-free degree distribution (Chung & Lu, 2002; Cho et al., 2009)[62,63]. To generate each randomised network models we made use of the R 'igraph' package (Csardi & Nepusz, 2006)[12] (version 1.0.0), and used the same number of nodes and edges found in the observed PPI networks, and where appropriate the estimated scaling parameter (Est. α) as given in Table 3. We compared the P-values generated using the observed network data, to the P-values obtained by fitting the power-law to the randomised network model following a scale-free degree distribution (Chung & Lu, 2002; Cho et al., 2009)[62,63]. If our PPI networks were no different from random, we would expect the random P-values to be >= to the observed P-value. We generated and fitted 1000 randomised models and counted the fraction of P-values >= the observed P-value. We found < 5% of the randomised P-values (pre-synaptic 3.45%, PSP 0% and PSP reduced 0.28%, as shown in Table 3) to be significant, supporting the case that connections in our PPI networks are not random. This can also be seen in Supplementary Figures 1B, 1D and 1F, where we plot each networks log-log degree distribution against the randomised models including: the Eros-Renyi (ER) model, the scale-free model with power-law exponent obtained from the fitted PPI networks (i.e. α = Est. α, as given in Table 3), and the scale-free model with the power-law exponent of 2 (α = 2.0), for a 1000 iterations for each model.

In Supplementary Figure 2 we show the scaled density distributions for the other centrality measures including: Betweenness (Bet.), Clustering Coefficient (CC), Page Rank (PR), Semi-Local centrality (SL), and mean shortest network path between genes (SP) using the LCC for each network. Measures Page Rank (Figure 2A) and Semi-Local centrality (Figure 2E) show the clearest separation between the networks. Along side each centrality measure, and each network, we also show the scaled density relative to the random models: Eros-Renyi (ER), scale-free model with power-law exponent obtained from the fitted PPI network (i.e. α = Est. α, as given in Table 3), and a scale-free network with power-law exponent of 2 (α = 2.0), for a 1000 iterations for each model.

**3. Identification of optimal clustering algorithm.**

Clustering, or community detection, in networks has been well studied in the field of statistical physics (Forunato, 2010)[47], with particular attention to methods developed for social science networks. The underlying assumption(s) of what makes a community in social science, translates remarkably well to what we think of as a community (sub-complex, module) in PPI networks. This has lead to the wide application of the unsupervised cluster detection measure Modularity (Q) (Newman & Girvan, 2004)[13], in social and biological network studies. Modularity can be naturally applied to our synaptic PPI networks, were we use Modularity to identify molecular clusters in our synaptic PPI networks, which maximise the PPIs within each clusters relative to PPIs between external clusters.

The PAC values for the PSP PPI network (~~Figure 5B and Table 5~~) ranged from: 0.99 (lec), 0.94 (fc), 0.84 (wt), 0.35 (sgG1), 0.23 (louvain), 8.2E-5 (Spectral), 1.4E-3 (infomap), 0.01 (sgG2), 0.01 (sgG5). The PAC values indicate a preference for the PSP network to be split into its sub-structure, i.e. to reveal communities contained within super-sized complexes; evidence for this includes algorithms Spectral and infomap with relatively high community values (Cn ~~86 and 117~~ shown in Table 5 respectively) producing low PAC values (8.2E-5 and 1.4E-3 respectively). This is to be compared against the lec, fc and louvain algorithms generating low Cn values (10, 13 and 14) and high PAC values (0.99, 0.94, 0.23). The results further suggested the presence of a large mixing parameter (Yang et al., 2016)[24] underlying the presynaptic network.

We employed a reclustering approach to those algorithms (lec, fc, wt, sgG1, louvain) with high PAC scores, > 0.1, to split their super-sized complexes into sub-complexes. Here we define super-sized complexes as those >= 10% of the number of vertcies in the network. For example, for the PSP PPI network with 4495 vertices this would imply communities of size >= 449. For each super-sized complex originally uncovered by the clustering algroithm, we reapply the algorithm to those edges and vertcies found internal to that complex. In general, the reclustering approach lowers the PAC scores found for the original algorithms: lec (0.99 -> 4.8E-3), fc (0.94 → 3.2E-3), wt (0.84 → 7.8E-3), sgG1 (0.35 → 0.017), louvain (0.23 → 0.01).

A different approach was applied to the Spectral algorithm. The Spectral algorithm generated the lowest PAC score (8.2E-5), however a large fraction (~37%) of these are single vertex communities: while more than half the communities discovered are of size <= 5 vertices. We can either view these singlet communities as outliers, or that they should belong to their ‘parent’ community before a division into children communities is made. The later view is more appropriate if we are interested in uncovering functional sub-complexes from the network. We therefore applied different stopping values to the Spectral algorithm of 1%, 2.5% and 5% of the number of vertcies in the network. For example, for the PSP PPI network with 4495 vertices this would imply stop splitting communities to sizes of 44, 112 and 224.

The PAC values for the presynaptic PPI network (Figure 5B and Table 5) ranged from: 0.98 (lec), 0.97 (fc), 0.55 (wt), 0.34 (sgG1), 0.26 (lourvain), 0.1 (Spectral), 0.02 (infomap), 0.008 (sgG2), 0.003 (sgG5). The PAC values indicate a preference for the presynaptic network to be split into its sub-structure, i.e. to reveal communities contained within super-sized complexes; evidence for this includes algorithms sgG5 and infomap with relatively high Cn values (86 and 117 shown in Table 5 respectively) producing low PAC values (0.003 and 0.02 respectively). This is to be compared against the lec, fc and lourvain algorithms generating low Cn values (10, 13 and 14 respectfully) and high PAC values (0.98, 0.97, 0.26). The results further suggested the presence of a large mixing parameter (Yang et al., 2016)[24] underlying the presynaptic network. A similar picture is also seen in the PSP and PSP reduced networks as given in Tables 6 and 7. The PAC values for the PSP reduced network (Figure 6B) ranged from: 0.99 (lec) and (fc), 0.77 (sgG1), 0.65 (wt), 0.59 (lorvain), 0.17 (infomap), 0.11 (Spectral), 0.007 (sgG2), 0.002 (sgG5). The low PAC values indicate a preference of the PSP reduced network to be split into more complexes.

Both pre- and post-synaptic PPI networks tended towards associative clustering ~~algorithms,~~ which revealed substructure (low PAC values in Tables 5-7). The mixing parameter (µ), the fraction of network edges between communities, is also considered an important global clustering parameter when analysing synthetic networks with well defined heterogeneous degree distributions and community structure (Yang et al., 2016)[24]. Low values of µ, i.e. ~0.2, indicate algorithms can uncover well the community structure from data. At µ ≥ 0.5, indicates the community structure found by the algorithm has been lost, implying there is a large overlap occurring between the communities. Applying µ to our clustered PPI presynaptic and PSP networks we find the opposite trend to the PAC measure, with the mixing coefficient favouring algorithms with a low number of large communities; where we note SVI is a purely probabilistic method, where the clustering result given by assigning each gene to the community with the highest probability found (Gopalan & Blei, 2013)[11]. Considering only the optimal setting for each the algorithms, the infomap (followed by Spectral) algorithm gave the optimal combination of PAC and mixing parameter values for both pre- and post-synaptic PPI networks. Therefore for all other studies we used both infomap and Spectral, along with a dedicated probabilistic clustering algorithm SVI as a comparison, to study properties of influential proteins within the pre- and post-synaptic PPI networks.

**4. Annotation enrichment tests.**

The probability genes intersection between two annotation sets, at the network level, was calculated using hypergeometric distribution (Pocklington et al., 2006)[31]:

~~–~~ (3)

Where is taken as the network size,and the number of annotations of types and respectively, and the number of genes overlapping between the two annotations sets. P-values calculated using eqn (3) where corrected using the Benjamini and Yekutieli (B-Y) [59] procedure, and tested against the more stringent Bonferroni correction at the 0.05 (\*), 0.01 (\*\*) and 0.001 (\*\*\*) significance levels, The hypergeometric distribution was also used to calculate the significance of enrichment of each cluster for each annotation:

-(4)

Where in (4) is the total number of genes in the network; the number of genes in the community; the total number of functional annotated genes in the network, and the number of functional annotated genes per community. P-values, ≤10-2, were tested for their strength of significance (sig), by recording the percentage of P-values found from every community/annotation combination, lower than or equal to the observed P-value, when 1000 random permutations of the annotation labels were made. P-values found with a strength of significance < 1% were considered statistically significant. P-values values were also tested against the more stringent Bonferroni correction at the 0.05 (\*), 0.01 (\*\*) and 0.001 (\*\*\*) significance levels, and highlighted throughout the enrichment tables (see Supplementary Tables S2-S4).

We also tested the significance of the overlap between two annotation sets within a community, or Bridging region (see Section 3.5), relative to the annotation set sizes at the network level:

-(5)

Where and are the number of annotations of types and , and the number of genes overlapping between the two annotation sets in community, or Bridging region, of size . P-values calculated using eqn (5) where corrected using the Benjamini and Yekutieli (B-Y) [59] procedure, and tested against the more stringent Bonferroni correction at the 0.05 (\*), 0.01 (\*\*) and 0.001 (\*\*\*) significance levels,

Equation (5) is similar in spirit to calculating the probability of the intersection distance between two distributions as given in eqn (13) pg 8 A. T. Kalinka, The probability of drawing intersesions: extending the hypergeometric distribution, arXiv:1305.0717v5 (2014). Where we'd set v1 = v2, and where we've focused on the population overlap relative to the the size of the community/region, and overlap found in it.

**5. False Discovery Rate calculation for Disease-Disease pairs.**

To calculate each disease-disease pairs q-value, we first calculated the fraction of randomised z-scores for a given disease pair AB, found greater than threshold parameter :

- (8)

Where in (8) gives the false discovery rate (FDR) for disease pair AB, the threshold z-score which we vary in the range [0,10] in steps of 0.1, the r­th randomised z-score for the disease pair AB, R (=1000) the total number of random iterations and the function 'theta' corresponds to the Heaviside step function.

The FDR over all disease pairs is then given by:

- (9)

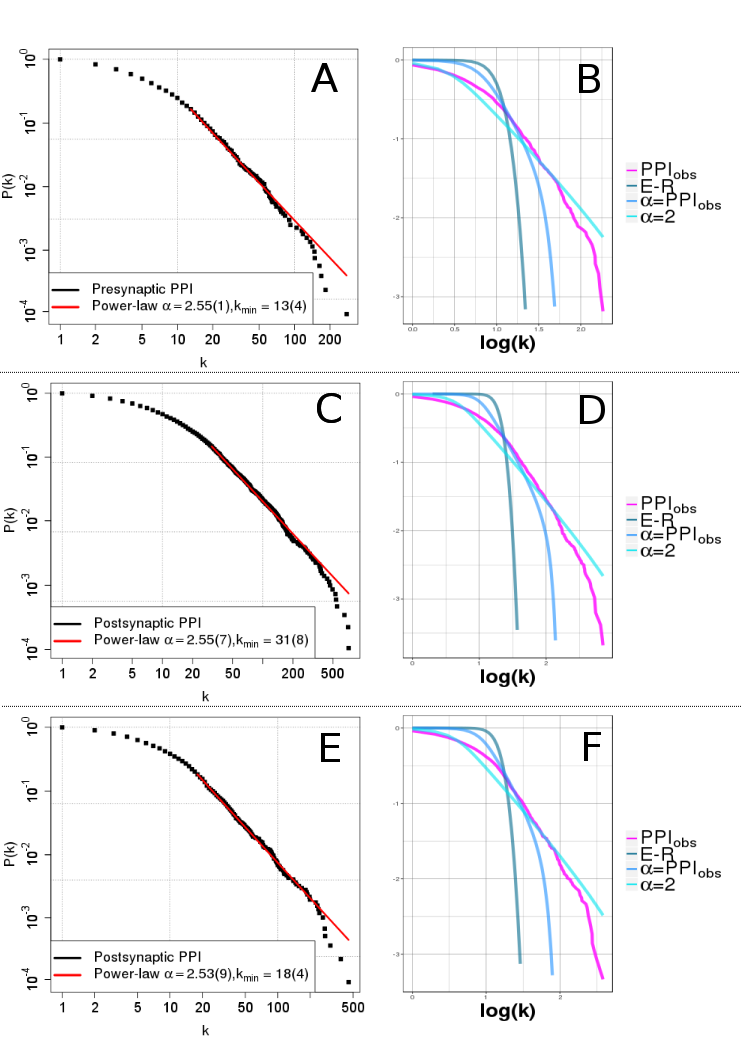
Where P denotes the set of all disease-disease pair combinations . The q-value for a given disease-disease pair is defined as the minimal FDR obatined from (9), at which the particular disease pair reaches significance.

**6. Disease –disease association**.

All diseases were found significantly localised on the presynaptic, PSP and PSP reduced PPI networks, with P-values ranging from 2.6x10-13 to 3.5x10-239 (Menche et al., 2015)[56]. Of the 55 disease-disease pairs (11 are trivial): 26 pairs on the presynaptic network were found significantly (i.e. q-value < 0.05) overlapping; 26 pairs on the PSP network were found to significantly overlap and 2 (AD-HD and ID-HD) were found to be significantly separated; on the PSP reduced network 24 pairs were found to be significant overlapping (for full details see the full randomised model given in Supplementary Table S7).

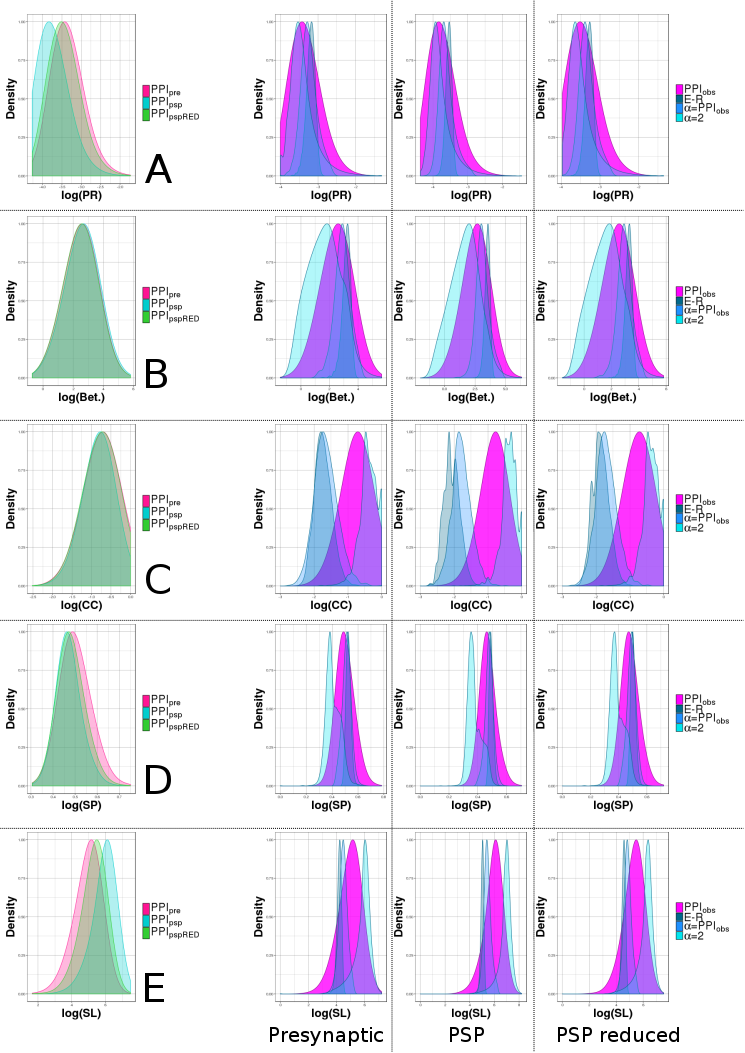
**Supplementary Figure 1**

(A) The data points represent the CDF of the degree distribution, P(k), versus degree, k, for the Presynaptic PPI network. The red line show the fitted power-law distribution with estimates of the scaling parameter α (2.55) and kmin (13), with a goodness-of-fit of 0.32 after 1000 iterations of the bootstrapping procedure [3]. (B) Degree distribution of the Presynaptic network (PPIobs) relative to the following random models after 1000 iterations: Eros-Renyi (ER), scale-free network with power-law exponent obtained from the fitted PPI network (α = PPIobs = 2.55), and scale-free network with power-law exponent of 2 (α = 2.0). (C ) The data points represent the CDF of the degree distribution, P(k), versus degree, k, for the PSP PPI network. The red line show the fitted power-law distribution with estimates of the scaling parameter α (2.51) and kmin (24), with a goodness-of-fit of 0.59 after 1000 iterations of the bootstrapping procedure [3]. (D) Degree distribution of the PSP network (PPIobs) relative to the following random models after 1000 iterations: Eros-Renyi (ER), scale-free network with power-law exponent obtained from the fitted PPI network (α = PPIobs = 2.51), and scale-free network with power-law exponent 2 (α = 2.0)). (E) The data points represent the CDF of the degree distribution P(k), versus degree, k, for the PSP reduced PPI network. The red line show the fitted power-law distribution with estimates of the scaling parameter α (2.59) and kmin (19), with a goodness-of-fit of 0.23 after 1000 iterations of the bootstrapping procedure [3]. (F) Degree distribution of the PSP reduced network (PPIobs) relative to the following random models after 1000 iterations: Eros-Renyi (ER), scale-free network with power-law exponent obtained from the fitted PPI network (α = PPIobs = 2.59), and scale-free network with power-law exponent of 2 (α = 2.0).



**Supplementary Figure 2**

Scaled density plots for the Presynaptic (pink), PSP (blue) and PSP redcued (green) PPI networks versus log of the following centrality measures: (A) PageRank (PR), (B) Betweenness (Bet.), (C ) Clustering Coefficient (CC), (D) mean of the shortest network path between genes (SP), and (E) Semi-local centrality (SL). Also shown for each centrality measure and each PPI network (PPIobs) the scaled density relative to the following random models after 1000 iterations: Eros-Renyi (ER), scale-free network with power-law exponent obtained from the fitted PPI network (i.e. α = PPIobs = Est. α, as given in Table 3), and scale-free network with power-law exponent of 2 (α= 2.0).



**Supplementary Figure 3**

Distribution of the disease pair overlap/separation () as calculated using eqn (6) for (A) AD versus HTN, (B) PD versus AD, ( C) PD versus HTN, (D) PD versus SCH, (E) SCH versus HTN and (F) SCH versus AD for 1000 randomised studies on the PSP PPI network (N: 4752 E: 48792). Highlighted on each distribution is the corresponding z-score for the disease pair overlap/separation as calculated using eqn (7), see also Supplementary Table S7.

