SUPPLEMENTARY NOTES

for

A system-level approach for deciphering the transcriptional response to prion infection

Mattia Zampieri [1,2], Giuseppe Legname [3] , Daniel Segre' [4,5] and Claudio Altafini [2,*]

1. Institute of Molecular Systems Biology, ETH, Wolfgang-Pauli-Str. 16, 8093 Zurich, Switzerland

2. SISSA-ISAS,International School for Advanced Studies via Bonomea 265, 34136 Trieste, Italy

3. Neurobiology Sector, via Bonomea 265, 34136 Trieste

4. Boston University Bioinformatics Program, 24 Cummington St., Boston

5. Boston University Biomedical Engineering, 44 Cummington St., Boston

August 11, 2011

Description of supplementary data

The following additional data are available with the on-line version of this paper. Additional data file 1 is a table listing the GEO ID of experiments included in the microarray analysis, for mouse, rat and human samples. Additional data file 2 is a table listing the algorithm predictions. For each gene the associated q.value and the estimated perturbation intensity are reported. Additional data file 3, 4 and 5 are tables listing functional enrichment results of inferred prion protein partners, prion-buffering genes and prion targets respectively. Additional data file 6 shows the drug's name with its relative mechanism of action and gene targets. Additional data file 7 reports the enriched pathways for the genes belonging to the network component of Fig.4. Additional data file 8 is a table listing the most frequently altered genes as reported in the existing literature. Color codes indicates common predicted genes with results in [4] (green) and already cited in previous publications (yellow). Additional data file 8 is the cytoscape format of the network in Fig.4. Finally, additional data file 9 contains the supplementary figures mentioned in the main text.

Additional analysis on Prnp and Prion inferred partners

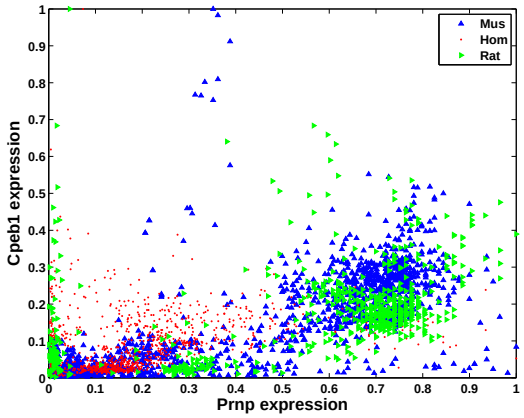
A specific comparison between the hypothesized 68 PrP^C ligands (www.genecards.org) with our list of 450 co-expressed genes, unveiled the enrichment of genes establishing physical bindings with PrP. In particular, we identified 11 PrP^C binding proteins yielding a p.value of 1.4187e-5 (hypergeometric test). These proteins (i.e. *Aldoc*, *Apbb*1, *Aplp*1, *Apoe*, *App*, *Clu*, *Eno*2, *Hprt*1, *Jph*3, *Tspan*7, *Cav*2) are mainly involved in neuron development, cell death, mrna polyadenilation, glycolysis/gluconeogenesis and Alzheimer's disease.

In case of prion infection, *Prnp* was in the most significant tail of the q.values distribution (2.8 percentile) and perturbation intensities (1.6 percentile), providing a strong validation to our approach. In addition, even in the case of infected mice overexpressing *Prnp* (data caollected from [4]) we could observe the high significance of the *Prnp* gene as a prions target (*q.value* equal to 0). On the contrary, the method applied to heterozygous infected mice was not able to recover any significance in the *Prnp* expression pattern. We believe this was due to the "additive" effects of the induced endogenous perturbation, that by reducing (approximately by half) the expression of the cellular form, masked the effect of the exogenous perturbation (i.e. prions propagation). This delayed the onset of the disease dramatically, and the process slowdown might require more accurate measurements to be properly analyzed.

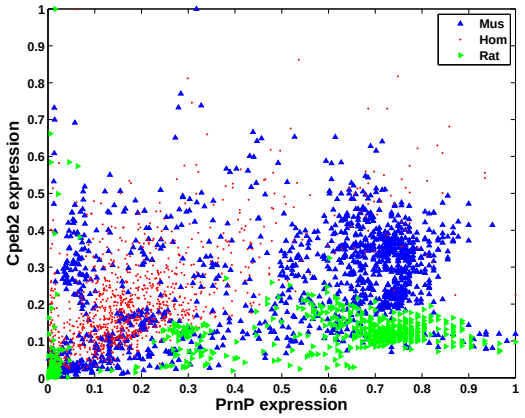
Expression patterns for Cpeb and Sirpa

Gene expression profiles of *Prnp* versus *Cpeb* family genes (i.e. Cpeb1, Cpeb2 and Cpeb4) and *Sirpa* are reported for the three compiled datasets (Fig.S1). We noticed that while for *Cpeb* genes a significant correlation was computed ($a-c$) in all three organisms, *Sirpa* (d) showed a very different expression pattern when passing from rat and mouse to human (Fig.S1, highlighted in bold). While in the first two mammals a direct proportionality was evident, in human an inverse relationship was emerging (On/Off switch like behavior).

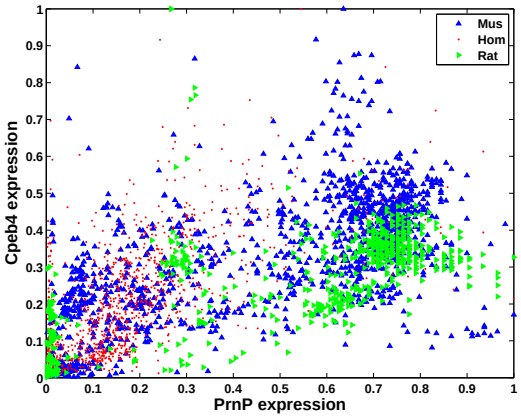
Figure S1: **Expression patterns**



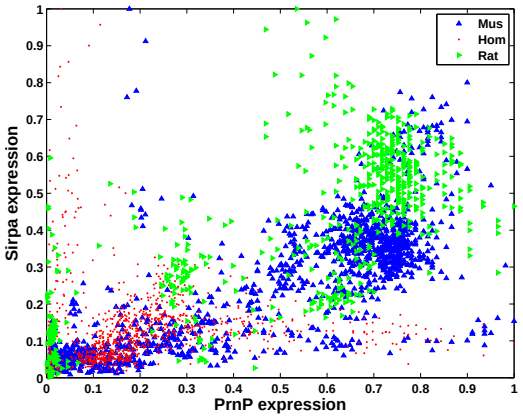
(a)



(b)



(c)



(d)

	Mouse		Human		Rat		All	
	Pearson	p.value	Pearson	p.value	Pearson	p.value	Pearson	p.value
Cpeb1	0.705248	0	0.28082	0	0.444861	0	0.559827	0
Cpeb2	0.45653	0	0.665226	0	0.243052	2.74E-010	0.522076	0
Cpeb4	0.668968	0	0.619358	0	0.623169	0	0.789792	0
Sirpa	0.789713	3.94E-295	**0.05948**	0.051327	0.728966	6.62E-110	0.78052	0

Network modularity

Co-regulation to the *Prnp* gene revealed prevalently genes encoding for proteins located on the membrane which are glycoproteins and related to nervous system development. For example, Eps15 which seemed to be involved in the regulation of mitogenic signals and in the control of cell proliferation, but in particular, in the internalization of ligand-inducible tyrosine kinase receptors (RTK), such as $Egfr$. In this study, *Ptgds* (potent inhibitor of platelet aggregation playing also a fundamental role in maturation and maintenance of the central nervous system) was inferred not only as a putative partner of PrP^C, but also as a target of PrP^{Sc}. Other genes seemed to match weak evidences of minor PrP functions, such as an effect of PrP on neuromuscular junction [6], or its hypothesized role in sleep and circadian rhythms [7]. Remaining significant biological processes were: ATPase activity, alcohol and nucleotide metabolic processes, carbohydrate catabolism, mitotic cell-cycle, response to light stimulus, hormone secretion and in particular long-term potentiation, previously linked to the cellular prion protein [5, 2]. Some of the predicted partners are involved in neurodegenerative diseases such as Ppp3ca, Chn1 [1], Rtn1, Rtn4, the *Sprn* gene, hypothesized to act as a modulator for the biological actions of the normal PrP, and *Prepl*, a serine peptidase interacting with *Snca* involved in Alzheimer disease (AD). Other genes are instead related to Huntington disease (HD), such as Dctn1, a key component of the mechanism of the axonal transport, and Calm1.

Assuming that the same level of details is captured for each gene in the dataset, a predicted interaction is more robust if conserved among different species. By searching for the connected components in the portions of the network conserved among mouse, rat and human (e.g. there is a path connecting any couple of nodes in the graph), 287 disjoint sub-graphs emerged.

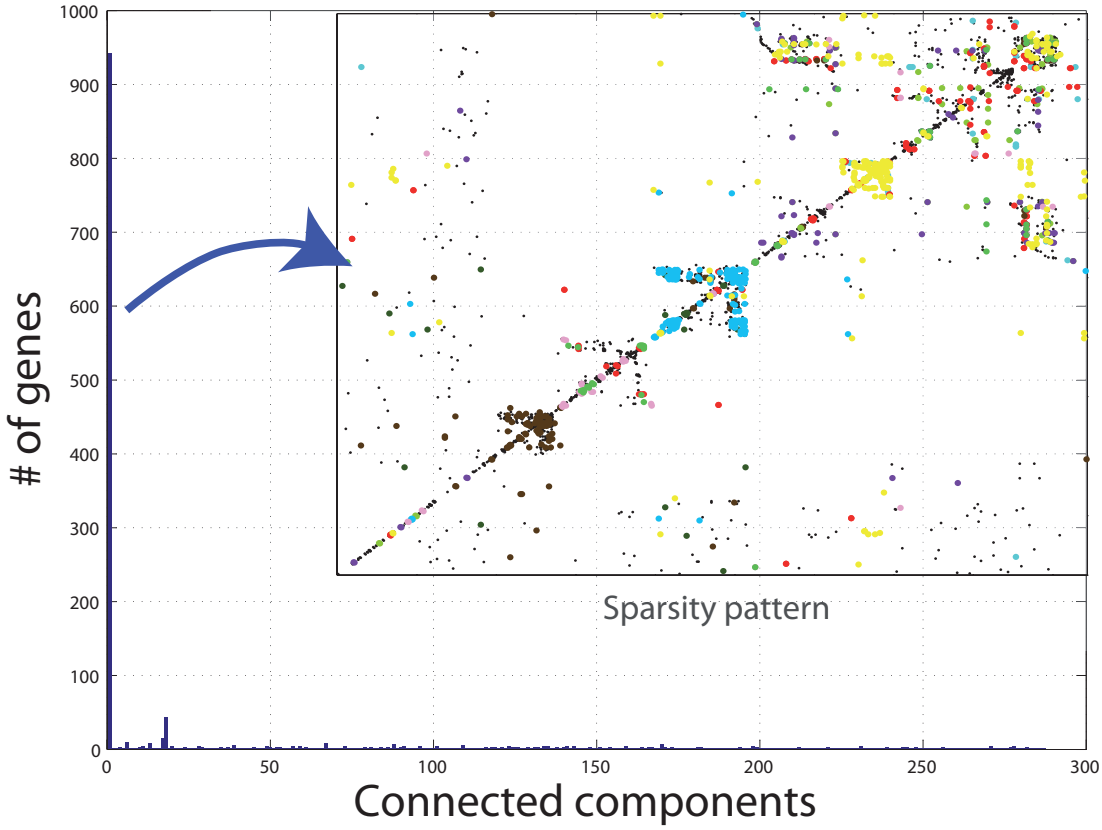


Figure S2: **Size of connected components.** The histogram represents the number of nodes in each component.

This distribution revealed a major component, on which we focused our attention. The first observation we made is the presence of several genes related to various neurodegenerative diseases (see Box 1 and Table S6). We visualized the adjacency matrix, opportunely reordered using the "minimum degree algorithm" [3], with the aim of minimizing the sparsity of the off diagonal terms. Such transformation is helpful in grouping around the diagonal densely connected components, that tend to avoid interactions with "distant" parts of the graph. By highlighting only the genes belonging to one of the significant enriched pathways (inset of Fig.S2) in the above connected component (e.g. Gap junction, Neurodegenerative Diseases, Cell cycle, Phosphatidylinositol signaling system, Proteasome, Ribosome, Long term depression potentiation, Leukocyte transendothelial migration, Natural killer cell mediated cytotoxicity, Oxidative phosphorylation) we could observe the tendency of most genes to influence other transcripts related to the same process. On the contrary for neurodegenerative related genes promiscuous interactions were more frequent.

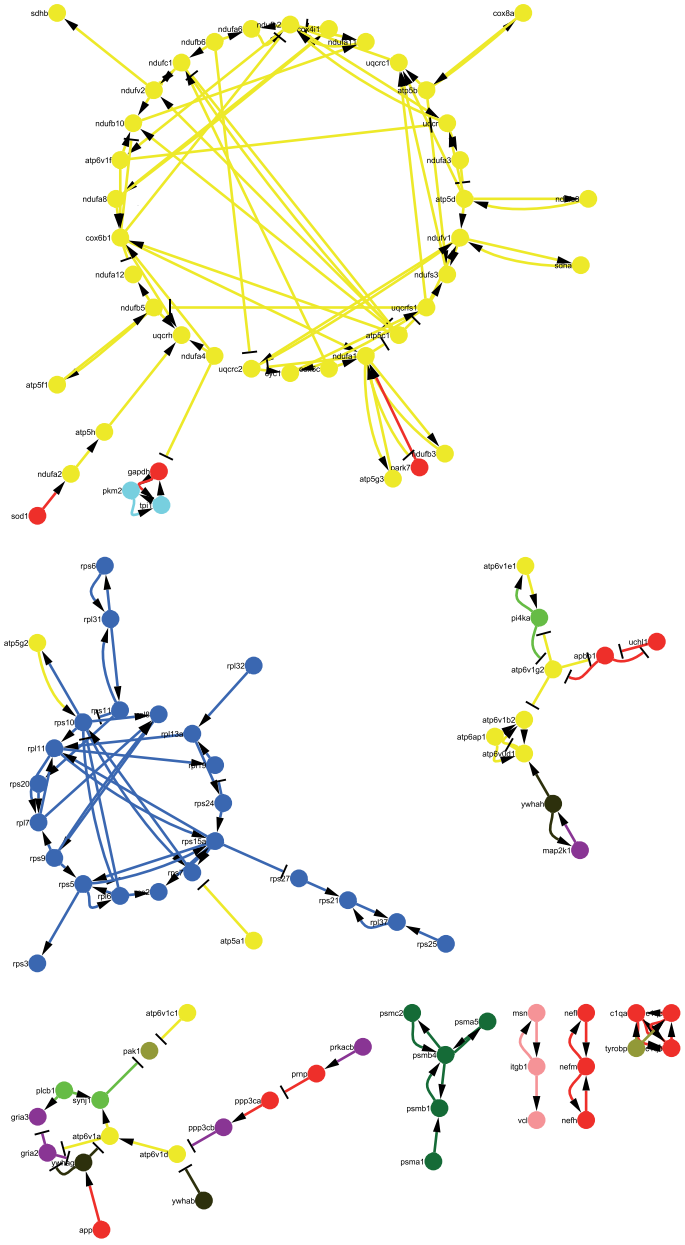


Figure S3: **Network details**

For example Sod1, here linked to oxidative phosphorylation (see Fig.S3 and the Cytoscape additional data file), is a ubiquitous, predominantly cytosolic protein that converts the superoxide radicals to hydrogen peroxide, in the mithocondrion. *Gadph*, was also indirectly related to oxidative phosphorylation and directly to gap junction. This was not surprising given that regardless of its glycolytic activity it is also involved in membrane trafficking in the early secretory pathway. Furthermore, neurofilaments ($Nefh$) are usually associated to two more intermediate filament proteins L, M which are involved in the maintenance of neuronal caliber. This complex was perfectly reproduced as a disjoint subnetwork. Finally, *Prnp* was mostly related to long-term depression/potentiation in agreement with several experimental observations [5, 2].

Prion diseases

Prion diseases are transmissible diseases that include Creutzfeldt-Jakob disease, bovine spongiform encephalopathy and scrapie. They are caused by the accumulation of an altered cellular prion protein (PrP^C) conformation named PrP^{SC} that subsequently acts as a catalyst for the recruitment and modification of the normal form.

Alzheimer's disease

Alzheimer is characterizable by extracellular β-amyloid plaques, neurofibrillary tangles and neuronal loss. Mutations in the Aβ precursor protein (APP) seem to be at the origin of the alterations in the processing of Aβ that lead to the toxic form. Experimental evidences suggest that Aβ is at the beginning of the amyloid cascade and gives rise to tau (MAPT) mislocation, misfolding and toxicity. Strictly related genes are: APOE and BACE1.

Huntington's disease

Huntington is caused by a mutation, consisting in an expanded CAG tract in the huntingtin protein. The poly-Q expansion leads to a conformational change resulting in protein aggregates.

Amyotrophyc lateral sclerosis

Amyotrophyc lateral sclerosis is mainly characterized by degeneration of motor neurons. Mutations of SOD1 gene result in a destabilization of the normal dimer that leads to the formation of amyloid structure. A fundamental gene for the disease progression is NEFH.

Parkinson's disease

The loss of dopaminergic neurons in the substantia nigra pars compacta and the presence of cytoplasmic inclusion bodies (Lewy bodies) in the neurons are characteristic markers of the disease. Among the proteins that have been related to PD, there are α-synuclein (SNCA), PARK7 and UCHL1.

Box 1: Brief neurodegenerative diseases description.

References

[1] G. M. Cosseddu, O. Andréoletti, C. Maestrale, B. Robert, C. Ligios, F. Piumi, U. Agrimi, and D. Vaiman. Gene expression profiling on sheep brain reveals differential transcripts in scrapie-affected/not-affected animals. *Brain Res*, 1142:217–222, Apr 2007.

[2] J. Curtis, M. Errington, T. Bliss, K. Voss, and N. MacLeod. Age-dependent loss of ptp and ltp in the hippocampus of prp-null mice. *Neurobiol Dis*, 13(1):55–62, Jun 2003.

[3] T. A. Davis, J. R. Gilbert, S. I. Larimore, and E. G. Ng. A column approximate minimum degree ordering algorithm. *SIAM Journal on Matrix Analysis and Applications*, 17:886–905, 2000.

[4] D. Hwang, I. Y. Lee, H. Yoo, N. Gehlenborg, J. H. Cho, B. Petritis, D. Baxter, R. Pitstick, R. Young, D. Spicer, N. D. Price, J. G. Hohmann, S. J. Dearmond, G. A. Carlson, and L. E. Hood. A systems approach to prion disease. *Mol Syst Biol*, 5:252–252, 2009.

[5] L. E. Maglio, V. R. Martins, I. Izquierdo, and O. A. Ramirez. Role of cellular prion protein on ltp expression in aged mice. *Brain Res*, 1097(1):11–18, Jun 2006.

[6] P. B. Nico, B. Lobão-Soares, M. C. Landemberger, W. Marques, C. I. Tasca, C. F. de Mello, R. Walz, C. G. Carlotti, R. R. Brentani, A. C. Sakamoto, and M. M. Bianchin. Impaired exercise capacity, but unaltered mitochondrial respiration in skeletal or cardiac muscle of mice lacking cellular prion protein. *Neurosci Lett*, 388(1):21–26, Nov 2005.

[7] I. Tobler, S. E. Gaus, T. Deboer, P. Achermann, M. Fischer, T. Rülicke, M. Moser, B. Oesch, P. A. McBride, and J. C. Manson. Altered circadian activity rhythms and sleep in mice devoid of prion protein. *Nature*, 380(6575):639–642, Apr 1996.