Comparing Alzheimer's and Parkinson's Diseases Using Graph Communities Structure

Alberto Calderone[1,4], Matteo Formenti[2,4], Federica Aprea[2,4], Michele Papa[3,4], Lilia Alberghina[2,4,5], Anna Maria Colangelo[2,4,5], Paola Bertolazzi[1,4]

1 Institute of Systems Analysis and Computer Science, National Research Council, Rome, Italy. **2** Laboratory of Neuroscience "R. Levi-Montalcini", Department of Biotechnology and Bioscience, University of Milano-Bicocca, Milano, Italy. **3** Laboratory of Neuronal Networks, Department of Mental and Physical Health and Preventive Medicine, Second University of Naples, Naples, Italy. **4** SYSBIO Centre of Systems Biology, University of Milano-Bicocca, Milano, Italy. **5** NeuroMI Milan Center for Neuroscience, University of Milano-Bicocca, Milano, Italy.

National Research Council of Italy
IASI
Institute for System Analysis and Computer Science

SYSBIO Centre of Systems Biology, University of Milano-Bicocca

Background

Neurological disorders exhibit a great variety of molecular alterations due to a complex interplay between polygenetic and environmental factors [1] . Classical reductionist approaches are focused on a few elements, providing a narrow overview of the etiopathogenic complexity of multifactorial diseases [2] . On the other hand, high-throughput technologies such as transcriptomics, proteomics, metabolomics and computational approaches allow the evaluation of many components of biological systems and their behaviors [2, 3] , thus allowing for system-level investigations. Molecular interactions are often represented as graphs which are the starting point of many computational approaches aiming to analyze, model, interpret and predict biological phenomena [4] . Network analysis of Parkinson's Disease (PD) and Alzheimer's Disease can highlight proteins or pathways common but differentially represented that can be discriminating between the two pathological conditions thus highlight similarities and differences.

Methods

1) We collected genes and proteins from two SBML models describing AD [5] and PD [6] and complemented these two lists with data downloaded from the KEGG database [7]. We used the human interactome [8] to extract two networks and we applied a score filer to reduce false positive interactions (Figure 1)

2) We use a state-of-the-art algorithm [9] to identify network communities in Alzheimer's disease (AD) and Parkinsons's disease (PD) protein interaction networks.

3) We assigned to each community a biological meaning using R and Gene Ontology.

4) To cross-compare AD and PD communities we calculated common GO terms and common edges between each community pair.

5) We constructed a Similarity Matrix that was then clustered using euclidean distance. This clustering step revealed areas in the Similarity Matrix that were statistically evaluated assigning to each GO term in the clusters a p-value calculated with respect to the entire Similarity Matrix. (Figure 2)

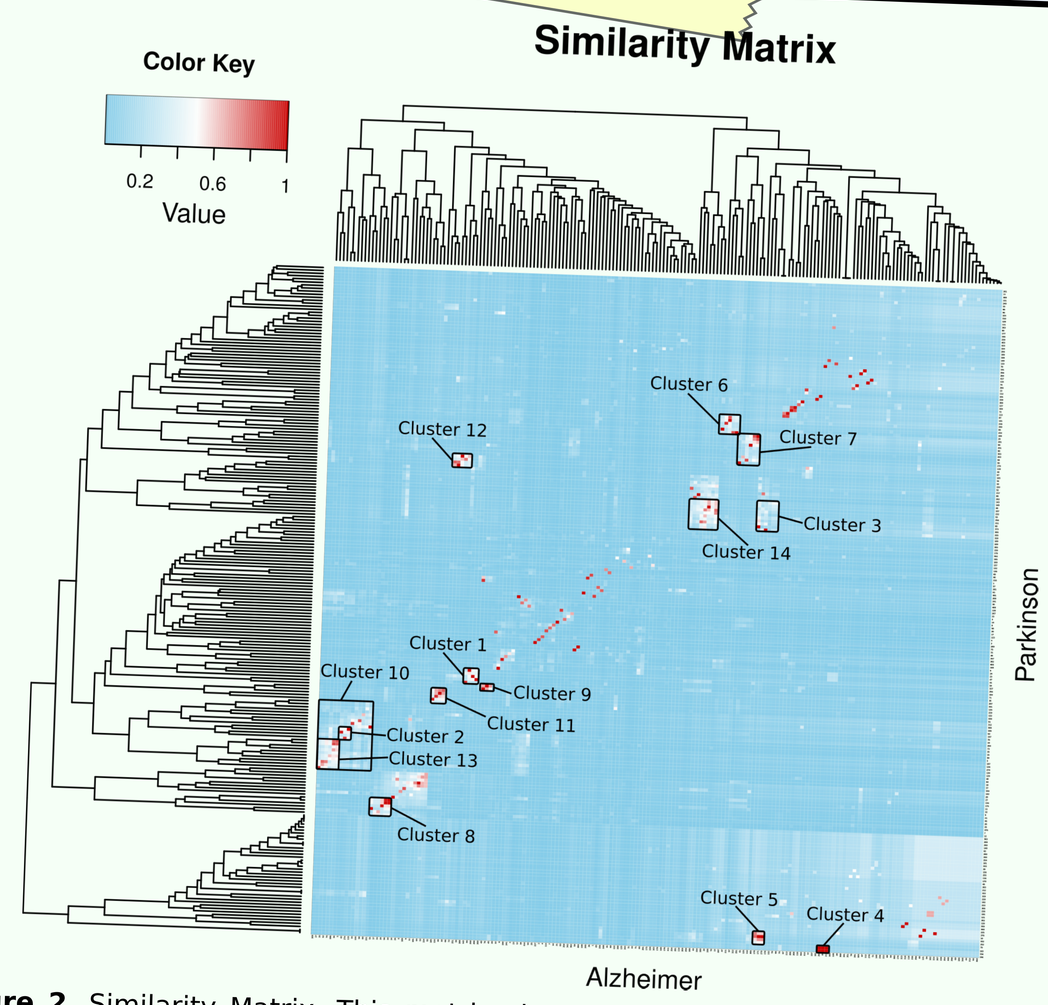

Figure 2. Similarity Matrix. This matrix shows statistically significant communities found in Alzheimer's and Parkinson's diseases according to their Gene Ontology overlap. Black boxes are clusters that might reveal strong significance. Single red dots are communities that are almost exclusively overlapped between the two pathologies. As an example, community 174 of AD includes enzymes catalyzing the synthesis of tetrahydropterin (BH4) while Glucose Metabolism processes have a significant p-values in clusters 8, 9 and 11.

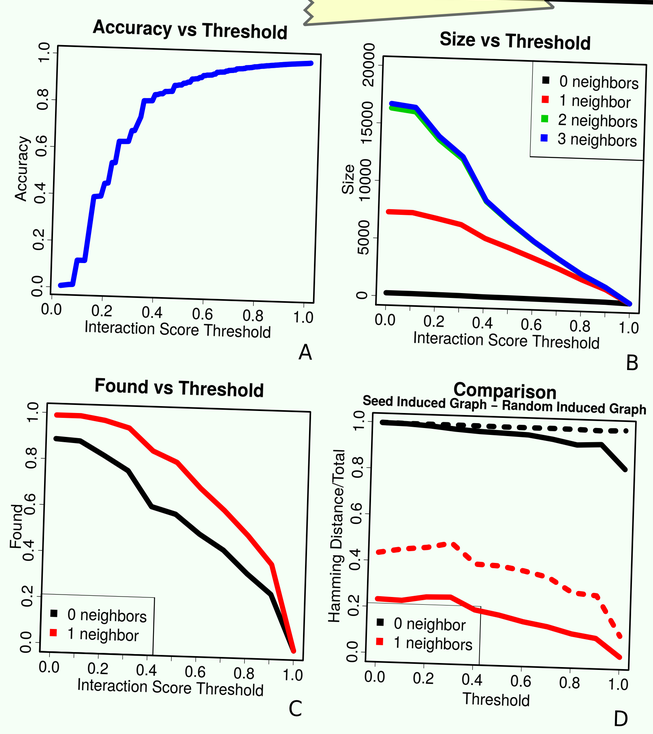

Figure 1. A) Accuracy against Reactome. Accuracy threshold greater than 0.4 captures Reactome's interactions. B) Network expansion on a starting set of about 400 vertices. The best choice was taking only the first neighbors. C) Recall. Average fraction of seed proteins captured in both networks at each threshold. D) Similarity between networks and random networks. Dashed lines show distance from random networks, continuous lines show distance between AD and PD networks. Distance 0, identical networks; distance 1, completely different network

Results

We obtained a list of 827 significant Gene Ontology terms from AD list, and a list of 550 terms from PD list. The simple intersection between these two lists resulted in a list that contained 368 common terms. Despite this richness of terms, known processes involved in both pathologies were missing or not significant.

We refined the basic Gene Ontology analysis by assessing the network communities containing GO terms with a significant Benjamini corrected p-value (p-value ≤ 0.05).

Furthermore, we analyzed a Similarity Matrix to extract clustered similar communities between the two pathologies. Our strategy allowed us to identify some common unknown processes and some known processes such as DNA repair, RNA metabolism and glucose metabolism which were not detected with simple GO enrichment.

In particular, we were able to capture the connection between mitochondrial dysfunction and metabolism (glucose and glutamate/ glutamine).

1. Alberghina L, Colangelo AM: The modular systems biology approach to investigate the control of apoptosis in Alzheimer's disease neurodegeneration. BMC Neurosci 2006, 7 Suppl 1:S2.
2. Minguez P, Parca L, Diella F, Mende DR, Kumar R, Helmer-Citterich M, Gavin A-C, van Noort V, Bork P: Deciphering a global network of functionally associated post-translational modifications. Mol Syst Biol 2012, 8:1–14.
3. Barabási A-L, Oltvai ZN: Network biology: understanding the cell's functional organization. Nat Rev Genet 2004, 5:101–13.
4. Villoslada P, Steinman L, Baranzini SE: Systems biology and its application to the understanding of neurological diseases. Ann Neurol 2009, 65:124–39.
5. Mizuno S, Iijima R, Ogishima S, Kikuchi M, Matsuoka Y, Ghosh S, Miyamoto T, Miyashita A, Kuwano R, Tanaka H: AlzPathway: a comprehensive map of signaling pathways of Alzheimer's disease. BMC Syst Biol 2012, 6:52.
6. Fujita K a., Ostaszewski M, Matsuoka Y, Ghosh S, Glaab E, Trefois C, Crespo I, Perumal TM, Jurkowski W, Antony PM a, Diederich N, Buttini M, Kodama A, Satagopam VP, Eifes S, Del Sol A, Schneider R, Kitano H, Balling R: Integrating pathways of Parkinson's disease in a molecular interaction map. Mol Neurobiol 2014, 49:88–102.
7. Kanehisa M, Goto S: KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 2000, 28:27–30.
8. Calderone A, Castagnoli L, Cesareni G: Mentha: a Resource for Browsing Integrated Protein-Interaction Networks. Nat Methods 2013, 10:690.
9. Rosvall M, Axelsson D, Bergstrom CT: The map equation. Eur Phys J Spec Top 2010, 178:13–23.