# Network-based analyses of Huntington's Disease

***Ravi K.R. Kalathur and Matthias E. Futschik***

Centre for Molecular and Structural Biomedicine, University of Algarve, Faro, Portugal

## Introduction

Huntington's disease (HD) is a neurodegenerative disorder caused by a CAG repeat expansion in the gene coding for the *Huntingtin* protein**.** Although is a classic Mendelian disease following a dominant inheritance pattern, strong inter-individual variability in the disease progression suggests the existence of biological modifiers that could provide novel therapeutic targets.

For a better molecular characterization, we constructed a network of pathways for 700 genes from the HD Research Crossroads database (*http://hdresearchcrossroads.org*). These genes have been associated with HD in previous studies or screens. Besides pathway analysis, we examined these genes for differential expression and their promotor region for enrichment in transcription factor binding sites. Finally, we predict potential novel candidates for HD-relevant genes utilizing protein interaction data from a newly constructed database (HDNetDB; *http://hdnetdb.sysbiolab.eu*)

## Pathway and functional analysis

In the current version of the HD Research Crossroads database, 747 genes were shown to be directly or indirectly involved in HD. We searched for the Gene Ontology (GO) annotations (**Fig.** 1) and KEGG pathways that are enriched by these genes using DAVID data base. For the GO analysis, we chose GO level 5 and considered categories as significant that obtained false discovery rate (FDR) ≤ 0.05.

### Functional analysis

**Biological processes**

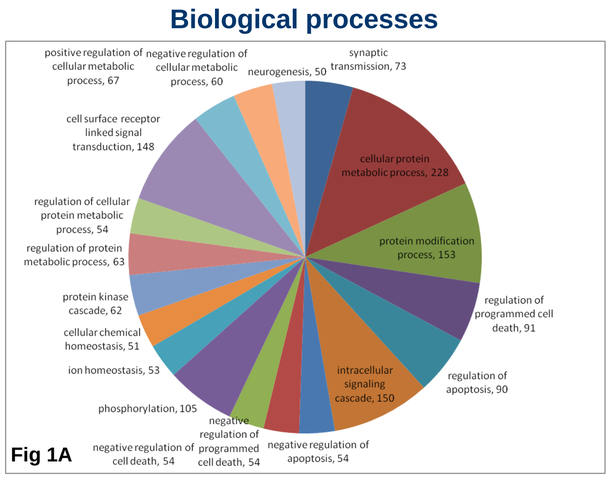



Fig 1A

**Molecular function**

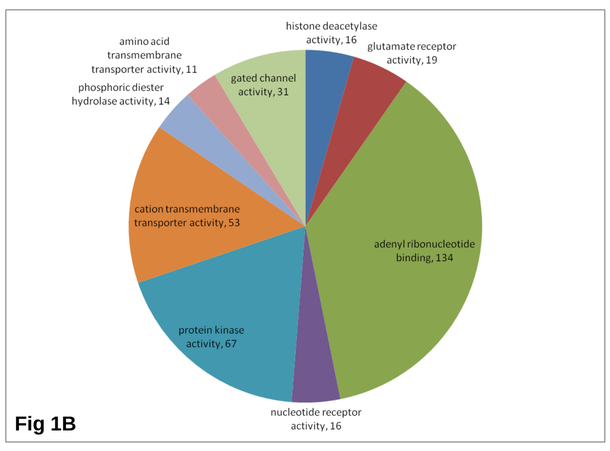



Fig 1B

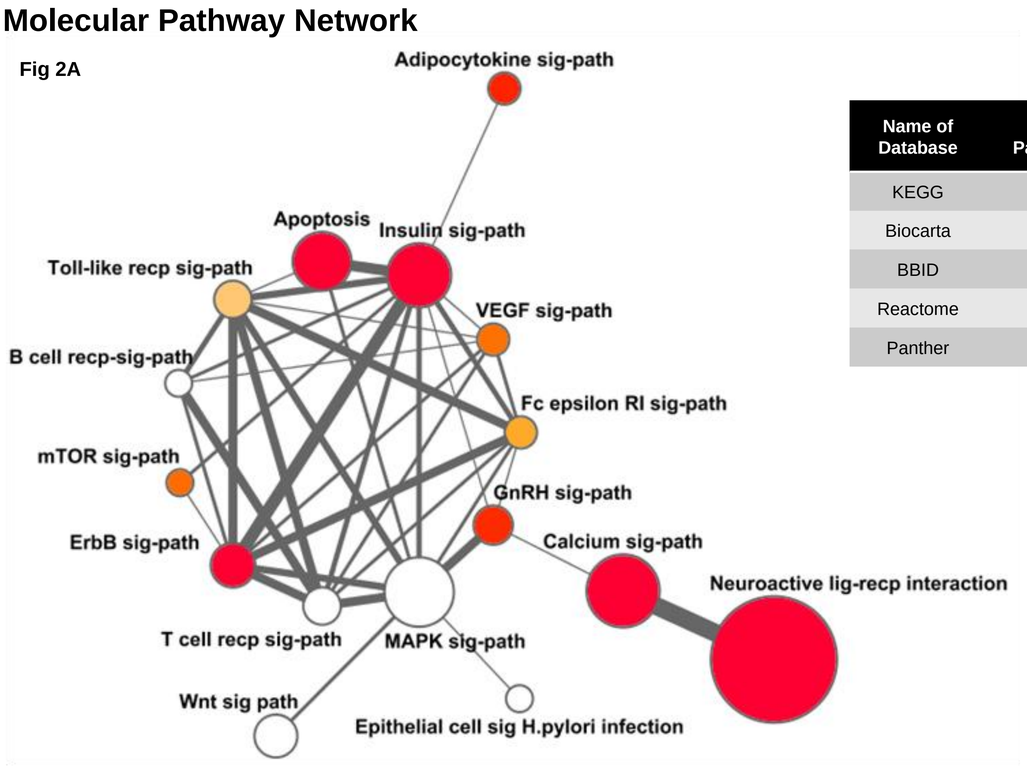
| GO category | Genes | FDR |
|---|---|---|
| Synaptic transmission | 73 | 6.35E-25 |
| Protein modification process | 153 | 7.74E-16 |
| Regulation of apoptosis | 90 | 3.99E-15 |
| Intracellular signaling cascade | 150 | 3.78E-15 |
| Cellular chemical homeostasis | 51 | 1.9E-10 |
| Protein kinase cascade | 62 | 1.58E-09 |

| GO category | Count | FDR |
|---|---|---|
| Protein kinase activity | 67 | 2.62E-08 |
| Cation transmembrane transporter activity | 53 | 7.29E-04 |
| Glutamate receptor activity | 19 | 8.19E-14 |
| Histone deacetylase activity | 16 | 7.58E-16 |
| Nucleotide receptor activity | 16 | 8.07E-09 |

**Figure 1.Gene Ontology enrichment analysis. A and B:** Pie charts show distribution of HD-associated genes across biological processes and molecular functions and corresponding tables display the number of genes and statistical significance of enrichment (FDR).

### Molecular Pathway Network

Fig 2A




| Name of Database | No: of Pathways |
|---|---|
| KEGG | 43 |
| Biocarta | 24 |
| BBID | 10 |
| Reactome | 19 |
| Panther | 42 |

### Disease Network

Fig 2B

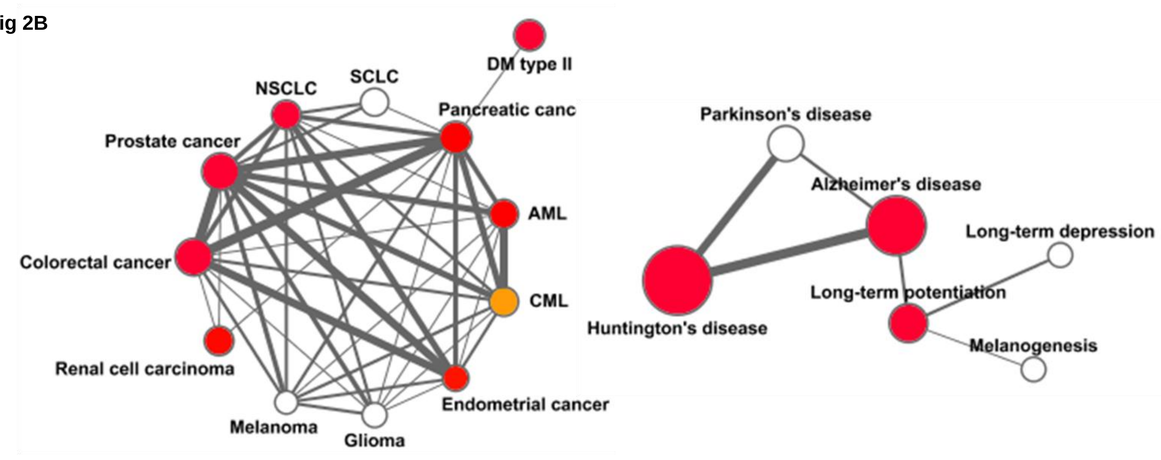



**Figure 2. Huntington's disease pathway network analysis. A**: Network of KEGG molecular pathways and **B**: of KEGG diseases gene set enriched in HD-associated genes. The statistical significance of enrichment is represented by the color (from red indicating high enrichment to white indicating no enrichment.)

## Expression and transcription factor analysis

To identify transcription factors perturbed in HD, we identified genes differently regulated in HD caudate nucleus genes using the expression data from Hodges A *et. al.* 2006 (**Fig 3A**). Subsequently, we obtained the promoter sequences (-1000 to +200 bp) using TRANSFAC and identified transcription factor whose binding sites tend to be enriched in the promotor regions of HD-associated genes. are highly significant enriched (p-value ≥ 0.0001;**Fig 3B**).

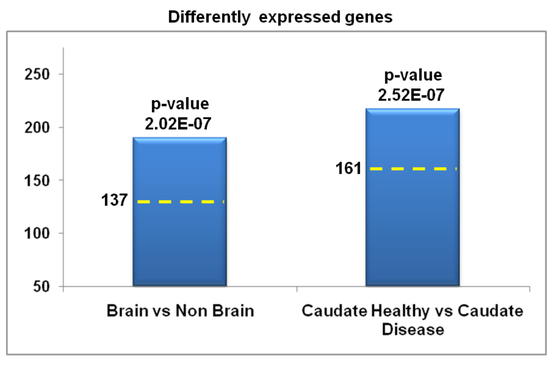



Fig 3A

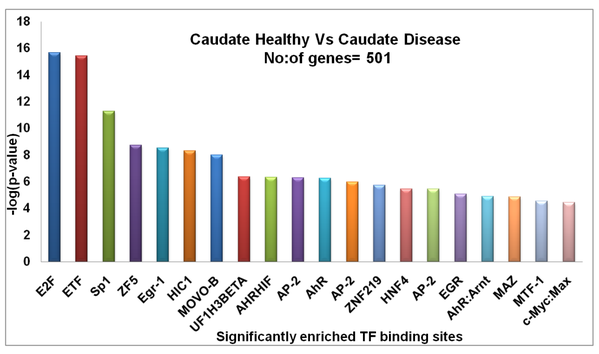



Fig 3B

**Figure 3: Expression and transcription factor analysis: A:** Number of gene differentially regulated in brain with respect to non-brain tissue and caudate nucleus and HD caudate nucleus. Expected number is marked by yellow dotted line; p-value was calculated by Fisher's exact test. **B**: TFs promotors sites enriched in differentially expressed genes. Y-axis: -log p-value.

### Role of identified transcription factors in brain and Huntington's disease

Our analysis yielded a number of transcription factors which may be have an important role in brain morphogenesis and possibly in HD:

- **E2F1** is involved in cell proliferation and apoptosis, controlling genes regulating S phase entry and DNA synthesis. E2F1 plays a major role in neurotoxicity and also triggers neural cells to undergo glutamate induced apoptosis in ischemic brain tissue by forming complex with PAR-4 *(Lu C et al. NAR. 2008)*.
- **EGR-1** is a nuclear protein and functions as a transcriptional regulator. It plays a major role at neuromuscular junctions. Acetylcholine receptors epsilon subunit synthesis is activated by NRG-1 induces expression of EGR-1 in myotubes. Inactivation of EGR-1 causes deficit in long-term recognition memory. *(Bozon B et al. Hippocampus 2002)*.
- **ETF** is one of the first transcription factors expressed at the beginning of mammalian development and plays a role in neural development by regulating the expression of PAX3 *(Yasunami M et al. JBC. 1995)*.

## Novel potential targets

Analysis of human protein interactions from HDNetDB (*http://hdnetdb.sysbiolab.eu*) showed that ~21 % of the interactors of a HD-associated proteins are associated with HD themselves, whereas only 6% of the total number of genes in the interactome are HD associated genes (**Fig 4A**). Equally, there exists strong tendency of HD-associated proteins to interact (**Fig 4B**). This trend might be used to predict novel HD-associated proteins: **Figure 5** shows two exemplary proteins which have relatively high number of HD-associated proteins among its interaction partners and thus may be associated with HD as well.

| | Human PPI interactome | Randomized scale-free network | Randomized network |
|---|---|---|---|
| HD-ass.protein - HD –ass.protein | 1188 | 958 | 470 |
| HD ass. protein- non HD-ass protein | 10599 | 11057 | 5055 |

Fig 4A

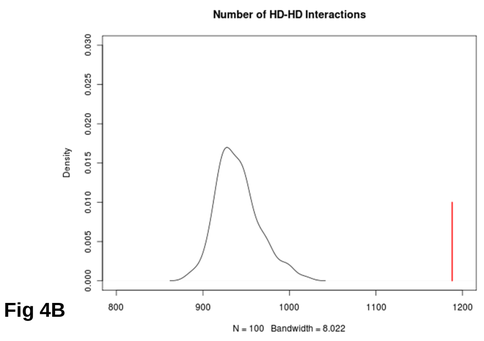



Fig 4B

**Figure 4: Comparison of Interactions: A:** Comparison of total number of interactions between HD for the PPI interactome and randomized networks(for each background model, 100 randomized networks were produced) **B**. Distribution of networks. Black : Distribution for scale free random network. Red: Observed number of interactions between HD genes.

Fig 5A

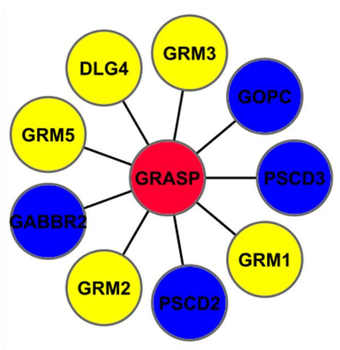



Fig 5B

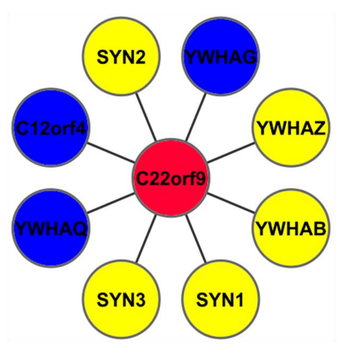



**Figure 5: Novel potential targets and their interacting partners: A:** GRASP and their interacting partners **B**. C22orf9 and their interacting proteins. Red node = Novel potential targets, yellow node = interacting HD proteins and blue node = other interacting proteins.

## Conclusion and Future Directions

- **We performed GO enrichment analysis and identified numerous biological processes associated with HD. Similarly, a network of associated pathways presenting a global view of the molecular mechanisms underlying HD.**
- **Identification of E2F1, EGR-1 and ETF as transcription factors linked to differential expression in HD brains.**
- **Through our analysis we identified GRASP and C22orf9 as novel potential targets of HD therapeutics. Further experimental validations are indicated for a better understanding the role of these genes in HD.**

## Acknowledgement

We would like to thank the **CHDI Foundation** for financial support of this research project.

Nature Precedings - doi:10.1038/npre.2012.7090.1 Posted 2 Apr 2012