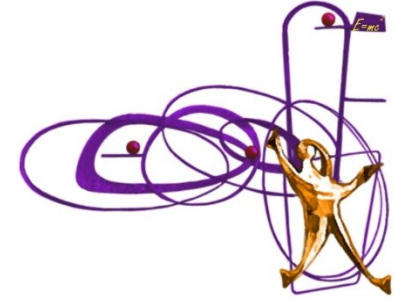


Dottorato di ricerca in Fisica (XXXI Ciclo)



UNIVERSITÀ
DEGLI STUDI DI BARI
ALDO MORO



Reti complesse per uno studio quantitativo delle malattie neurodegenerative

Progetto di ricerca di Eufemia Lella

Mi presento

Tesi di laurea - Aprile 2014:

«Caratterizzazione di un sistema di rivelazione gamma per l'analisi di materiali con attivazione neutronica»

Dicembre 2014 - Novembre 2015 (Dipartimento di Informatica):

Progetto di formazione e ricerca Edoc@Work 3.0: utilizzo di tecnologie Cloud per la produzione e fruizione di contenuti per la *smart education*

Stage di ricerca in collaborazione con INFN (tutor: Prof. Bellotti):

Complex network analysis su dati di benchmark e su reti di collaborazioni europee (RStudio su piattaforma Cloud² di BeCaS)

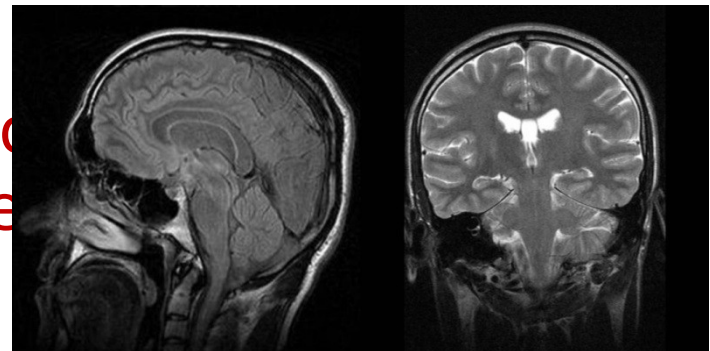
Contesto

Malattie neurodegenerative: Alzheimer, Parkinson, ...

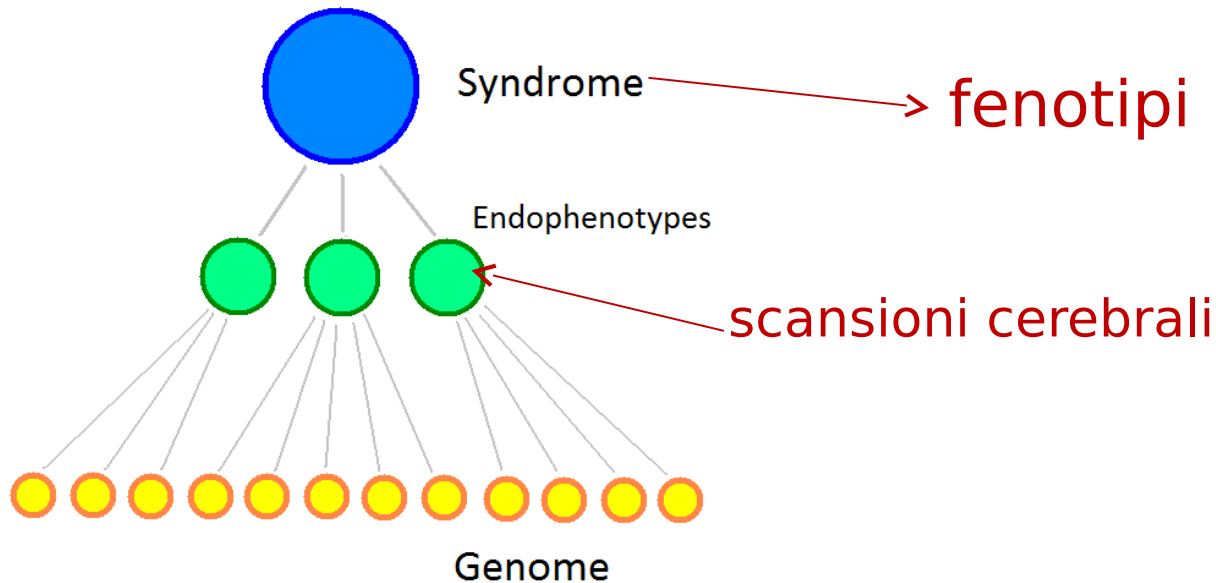
Tra i sintomi precursori: variazioni delle caratteristiche strutturali e morfologiche del cervello

Neuroimaging: immagini del cervello (**MRI** o **f-MRI**)

Diagnosi precoce (es. previsione) e identificazione del trattamento farmacologico più adatto



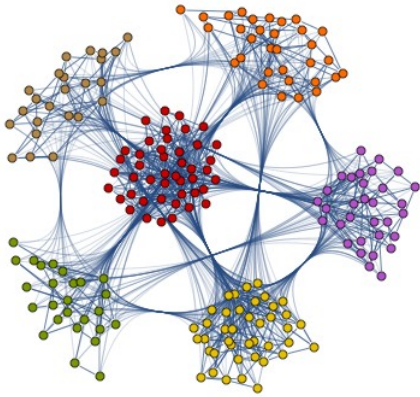
Problema



Tecniche statistiche classiche sono insufficienti a catturare l'intrinseca complessità delle interazioni tra geni e le associazioni fra endofenotipi e genotipi

La descrizione di tali associazioni, ad oggi, è un **problema aperto**

Reti complesse

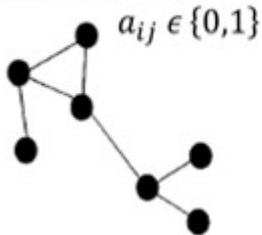


Gran numero di elementi interagenti tra loro
 Esempi: reti di trasporti, reti sociali, reti di telecomunicazioni, reti biologiche

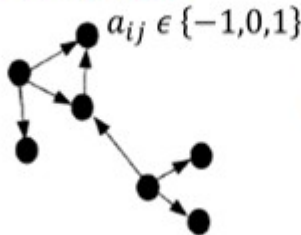
Esempio di trattazione matematica: **teoria dei grafi**

Grafo $G(N,E)$ N = insieme di n nodi E = insieme di archi

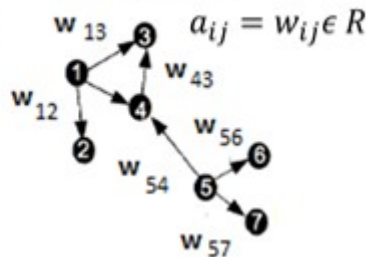
UNDIRECTED



DIRECTED



WEIGHTED



Rappresentazione compatta
Matrice di adiacenza A
 elementi a_{ij} con $i, j = (1, \dots$
 generici nodi della rete

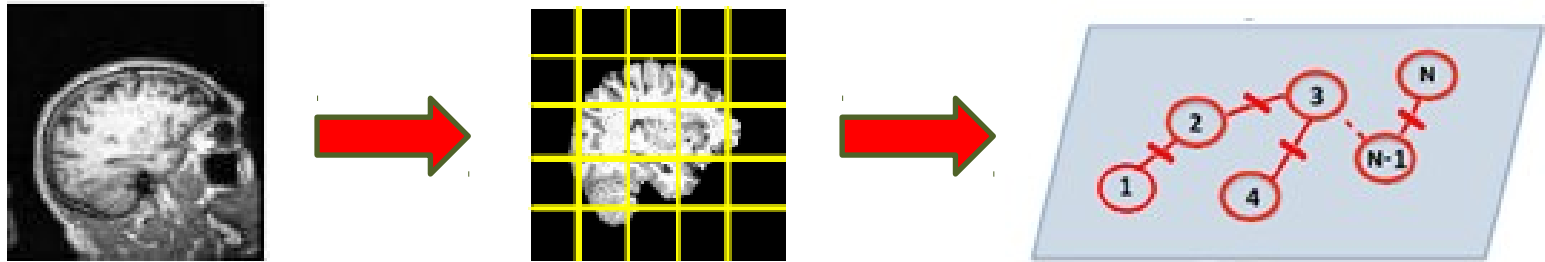
Indicatori statistici: *degree centrality*, *betweenness centrality*, **connectivity**, etc...

informazioni sulla centralità dei nodi e sulle associazioni tra di essi

Da immagini a reti complesse

Analisi delle scansioni cerebrali espresse sotto forma di **reti complesse**

Sotto-volumi 3D (*patches*) della scansione = nodi di una rete complessa



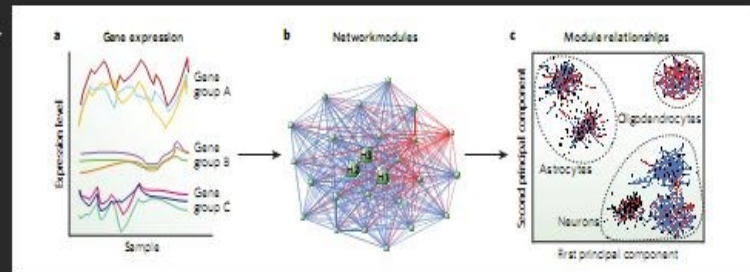
Sintesi delle caratteristiche strutturali e morfologiche attraverso indicatori statistici

Weighted Gene Co-expression Network Analysis (WGCNA)

From LISTS of genes...

Gene ID	Gene Name	Function
1445537_AT	RNBI DNA 30036251 GENE	10090 Rat transcript
1442258_W	DIAMINOCACETATE METHYL TRANSFERASE	10090 Rat transcript
1439172_AT	CYTOCHROME P450 FAMILY 4, SUBFAMILY F, POLYPEPTIDE 14	10090 Rat transcript
1442940_W	PROTEIN KINASE, SHIP-ACTIVATED, ALPHA 2, CATALYTIC SUBUNIT	10090 Rat transcript
1432342_AT	RNBI DNA 40213035 GENE	10090 Rat transcript
1448914_AT	7-DEHYDROCOLESTEROL REDUCTASE	10090 Rat transcript
1439803_AT	RNBI DNA 31382761 GENE	10090 Rat transcript
1441935_AT	RNBI DNA 27683013 GENE	10090 Rat transcript
1441992_AT	CYTOCHROME P450 FAMILY 2, SUBFAMILY C, POLYPEPTIDE 16	10090 Rat transcript
1437532_W	RNBI DNA 28399522 GENE	10090 Rat transcript
1448891_AT	COPOLYMERIZATION SYNTHASE (PHOSPHATASE CYTOSOLIC) TRANSFERASE 2	10090 Rat transcript
1438638_W	EXPRESSED SEQUENCE ANNOT 206	10090 Rat transcript
1438635_AT	EXPRESSED SEQUENCE ANNOT 206	10090 Rat transcript
1441892_AT	REGULATORY LINE NUCLEIC ACID TRANSFERASE	10090 Rat transcript
1432443_A_AT	AMP YRNA REPEAT AND P/B P/D/DOMAIN CONTAINING 2	10090 Rat transcript
1442752_AT	ACETYL COENZYME A CARBOXYLASE BETA	10090 Rat transcript
1443255_AT	RNBI DNA 28399591 GENE	10090 Rat transcript
1429622_W	SIMILAR TO GLUTATHIONE S-TRANSFERASE MU 1 (GST CLASS MU 1) (GLUTATHIONE S-TRANSFERASE)	10090 Rat transcript
1434844_AT	NUCLEAR RECEPTOR SUBFAMILY 1, GROUP 1, MEMBER 1	10090 Rat transcript
1442479_AT	ACETYL COA SYNTHETASE	10090 Rat transcript
1421207_A_AT	ORPHAN NUCLEOTIDE TRANSACTINATOR WITH GALLIUM-INDUCIBLE CARRIER TERMINAL DOMAIN 2	10090 Rat transcript
1439207_AT	RNBI DNA 40089181 GENE	10090 Rat transcript
1439205_A_AT	PHOSPHATIDYLTRANSFER PROTEIN, CYTOSOLIC 1	10090 Rat transcript
1448920_AT	CASPASE 2	10090 Rat transcript
1437938_A_AT	CATEPSIN C	10090 Rat transcript
1445959_W	SELENIN BINDING PROTEIN 1	10090 Rat transcript
1441786_L_AT	SELENIN BINDING PROTEIN 1	10090 Rat transcript
1432173_AT	RNBI DNA 57343269 GENE	10090 Rat transcript
1441892_AT	COAGULATION FACTOR F	10090 Rat transcript
1434946_AT	STROMAL ANTIWEN 1	10090 Rat transcript
1440115_W	SEMA DOMAIN IMMUNOCOLLEGEN DOMAIN 20, TRANSMEMBRANE DOMAIN 19E AND SHORT CYTOSOLIC TAIL	10090 Rat transcript
1441761_W	TU TRANSGLUTININ-LINKAGE FACTOR, MICRODOMAIN	10090 Rat transcript
1439122_A_AT	DNA METHYL TRANSFERASE (CYTOSINE 5)	10090 Rat transcript
1448914_AT	RNBI DNA 10038919 GENE	10090 Rat transcript
1432098_AT	DNA SEQUENCE 109 13, BRAT 200 REF. EXPRESSED	10090 Rat transcript
1432268_A_AT	ZINC FINGER AND P/B DOMAIN CONTAINING 20	10090 Rat transcript
1432257_A_AT	GLUTATHIONE S-TRANSFERASE MU 1A	10090 Rat transcript
1448911_AT	RNBI DNA 14300380 GENE	10090 Rat transcript
1430676_A_AT	POLY (U) POLYMERASE ALPHA	10090 Rat transcript
1429192_AT	CALCIUM AND MAGNESIUM-DEPENDENT PROTEIN KINASE D	10090 Rat transcript
1441702_AT	OLIGOMERIN C ACETYL TRANSFERASE (D-AMINO-SEROTONIN-5-HYDROXYASE COENZYME A LIASIS)	10090 Rat transcript
1438937_AT	ESTROGEN RECEPTOR 1 (ALPHA)	10090 Rat transcript
1448944_A_AT	MAD HELIX LOOP 1 (PROSOMAL)	10090 Rat transcript
1437982_W	MYOZIN BA	10090 Rat transcript
1429482_A_AT	URENIN PHOSPHOLIPASE 2	10090 Rat transcript
1445148_AT	URENIN PHOSPHOLIPASE 2	10090 Rat transcript
1421746_A_AT	SOLUBLE CARRIER FAMILY 1 (DEURIC) TRANSFERRED TRANSPORTER, LAURINE, MEMBER 5	10090 Rat transcript
1421745_AT	SOLUBLE CARRIER FAMILY 1 (DEURIC) TRANSFERRED TRANSPORTER, LAURINE, MEMBER 5	10090 Rat transcript
1448920_A_AT	RNBI DNA 85334510 GENE	10090 Rat transcript
1432936_AT	GLUTATHIONE S-TRANSFERASE, ALPHA 1 (14)	10090 Rat transcript
1445537_W	CYTOCHROME P450 FAMILY 4, SUBFAMILY F, POLYPEPTIDE 14	10090 Rat transcript
1445535_W	CARBONIC DEHYDRATASE	10090 Rat transcript
1448927_AT	CARBONIC DEHYDRATASE	10090 Rat transcript
1445764_W	COMPLEMENT COMPONENT 2 (HRRW125)	10090 Rat transcript
1439172_W	POLYMERASE ALPHA 2 FACTOR	10090 Rat transcript
1438330_AT	MYELOBLASTOSIS OR MIB2-LIKE B-LYMPHOMA TRANSLOCATION TO 4 HOMOLOG (PROSOMAL)	10090 Rat transcript

...To networks



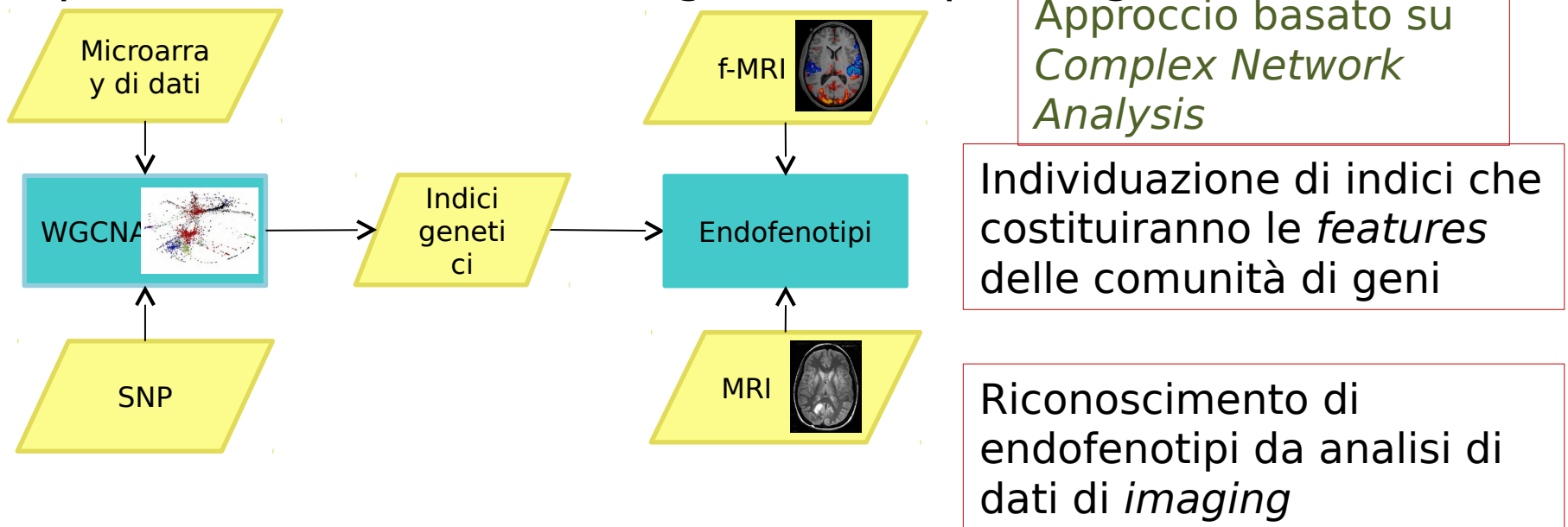
WEIGHTED GENE CO-EXPRESSION NETWORK ANALYSIS (WGCNA)

Il profilo di espressione può essere organizzato in networks secondo correlazioni di espressione

nodì: espressioni geniche
archi: $\rho > \theta$

Approccio al problema

Obiettivo: studio delle associazioni tra endofenotipi (marcatori delle variazioni strutturali e morfologiche del cervello nelle scansioni cerebrali) e genotipi al fine di predire l'eventuale insorgere della patologia



Classificazione sani/malati: algoritmi di **Data Mining** e **Machine Learning** per l'elaborazione di grandi moli di dati (**big data**) e la costruzione di un modello predittivo

Fasi del progetto

Primo anno:

- Studio dello stato dell'arte
- Acquisizione degli strumenti di analisi statistica e computazionale (R, MATLAB)
- Casi di studio
- Acquisizione dei dati d'interesse

Secondo e terzo anno:

- Studio di previsione da dati di *imaging*:
 - Elaborazione dei dati (da immagini a reti complesse)
 - Misura delle caratteristiche strutturali e morfologiche
 - Classificazione supervisionata e predizioni
- Associare l'analisi genetica allo studio morfologico (*Imaging genetics*)
- Raffinamento del modello predittivo
- Stesura tesi finale

Alcuni riferimenti

1. Barabási A.L., Gulbahce N., Loscalzo J. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, 12, pp. 56-68, 2011.
2. Bellotti R., Pascazio S. Editorial: Advanced physical methods in brain research. *The European Physical Journal Plus*, 127, p. 145, 2012.
3. Boccaletti S., Latora V., Moreno Y., Chavez M., Hwang D.U. Complex Networks: Structure and dynamics. *Physics Report*, 424 (4-5), pp. 175-308, 2006.
4. Frisoni G.B. Structural imaging in the clinical diagnosis of alzheimer's disease: problems and tools. *Journal of Neurology, Neurosurgery & Psychiatry*, 70(6), pp. 711-718. 2001.
5. Horvath S., Dong J. Geometric Interpretation of Gene Coexpression Network Analysis. *PLOS Computational Biology*, 4(8), 2008.