

Mechanism Based Stratification of Patient-Level Data

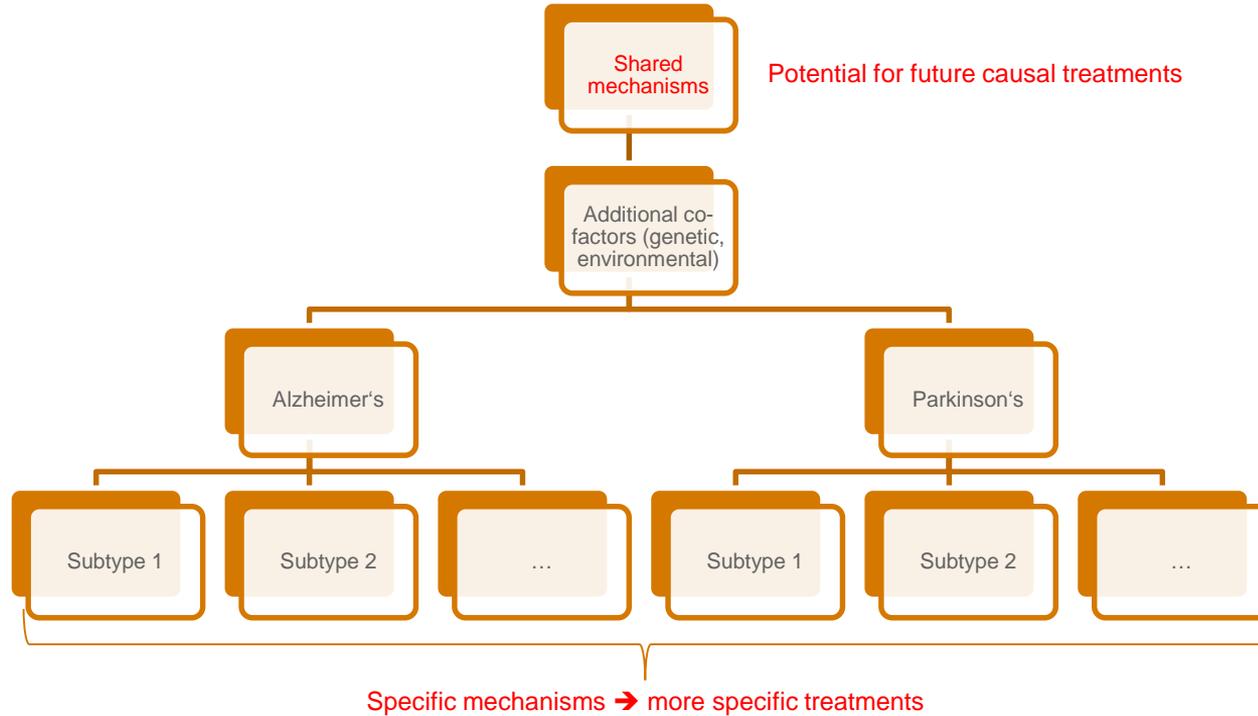
Holger Fröhlich

30.11.2018



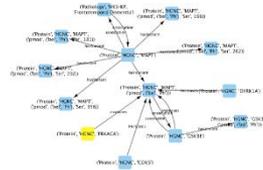
Guoqiong, living with epilepsy

A Hypothetical Molecular Mechanism Based Disease Taxonomy



Realizing a Mechanism Based Disease Taxonomy

1.) Mechanism based stratification within separate diseases

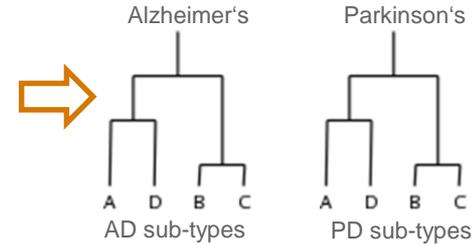


Short-listed NeuroMMSig mechanisms
(disease specific)

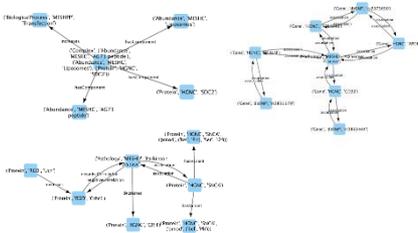


	AD1_002_N_0019	AD1_002_N_0020	AD1_002_N_0021	AD1_002_N_0022	AD1_002_N_0023
AD1_002_N_0017	0	0	1	0	0
AD1_002_N_0018	1	0	0	0	0
AD1_002_N_0019	0	0	0	0	0
AD1_002_N_0020	1	1	1	0	1
AD1_002_N_0021	0	0	0	0	0
AD1_002_N_0022	1	1	1	1	0
AD1_002_N_0023	0	1	1	1	0
AD1_002_N_0024	0	0	0	0	0
AD1_002_N_0025	0	1	1	1	0
AD1_002_N_0026	0	0	0	0	0
AD1_002_N_0027	0	0	0	1	0

omics data



2.) Joint AD/PD stratification by shared mechanisms

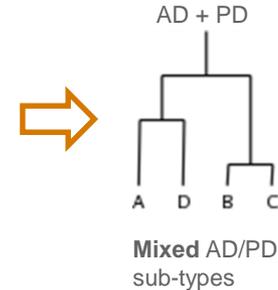


Shared molecular mechanisms



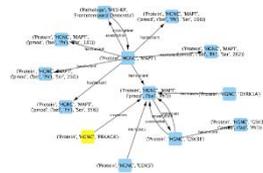
	AD1_002_N_0019	AD1_002_N_0020	AD1_002_N_0021	AD1_002_N_0022	AD1_002_N_0023
AD1_002_N_0017	0	0	1	0	0
AD1_002_N_0018	1	0	0	0	0
AD1_002_N_0019	0	0	0	0	0
AD1_002_N_0020	1	1	1	0	1
AD1_002_N_0021	0	0	0	0	0
AD1_002_N_0022	1	1	1	1	0
AD1_002_N_0023	0	1	1	1	0
AD1_002_N_0024	0	0	0	0	0
AD1_002_N_0025	0	1	1	1	0
AD1_002_N_0026	0	0	0	0	0
AD1_002_N_0027	0	0	0	1	0

omics data



Realizing a Mechanism Based Disease Taxonomy

1.) Mechanism based stratification within separate diseases



Short-listed NeuroMMSig mechanisms
(disease specific)

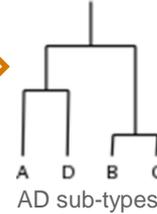


	AD1_002_X_0019	AD1_002_X_0020	AD1_002_X_0021	AD1_002_X_0022	AD1_002_X_0023
AD1_002_X_0017	0	0	1	0	0
AD1_002_X_0018	1	0	0	0	0
AD1_002_X_0019	0	0	0	0	0
AD1_002_X_0020	1	1	1	1	1
AD1_002_X_0021	1	1	1	1	1
AD1_002_X_0022	0	1	1	1	1
AD1_002_X_0023	0	1	1	1	1
AD1_002_X_0024	0	0	0	1	1
AD1_002_X_0025	0	0	0	0	0

omics data

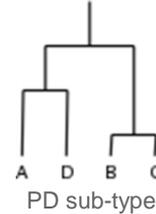


Alzheimer's



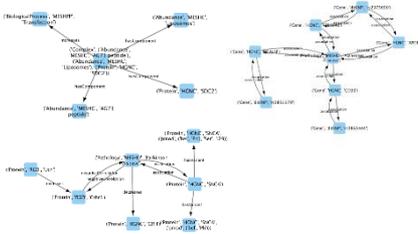
AD sub-types

Parkinson's



PD sub-types

2.) Joint AD/PD stratification by shared mechanisms



Shared molecular mechanisms

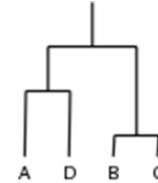


	AD1_002_X_0019	AD1_002_X_0020	AD1_002_X_0021	AD1_002_X_0022	AD1_002_X_0023
AD1_002_X_0017	0	0	1	0	0
AD1_002_X_0018	1	0	0	0	0
AD1_002_X_0019	0	0	0	0	0
AD1_002_X_0020	1	1	1	1	1
AD1_002_X_0021	1	1	1	1	1
AD1_002_X_0022	0	1	1	1	1
AD1_002_X_0023	0	1	1	1	1
AD1_002_X_0024	0	0	0	1	1
AD1_002_X_0025	0	0	0	0	0

omics data



AD + PD



Mixed AD/PD
sub-types

Which omics data to use?

	AD					PD					
	ADNI	ROSMAP	IDIBAPS	INSIGHT	UKB	PPMI	DIGPD	ICEBERG	AETIONOMY PD	KCL	Tübingen
<i>N (disease cases)</i>	486		537		220	362			733		232
SNPs	■					■					
genome-wide DNA methyl	■					■					
proteomics	■					■					
CHIPseq	■					■					
RNAseq	■					■					
inflammation markers	■					■					

Initially only ADNI + PPMI available

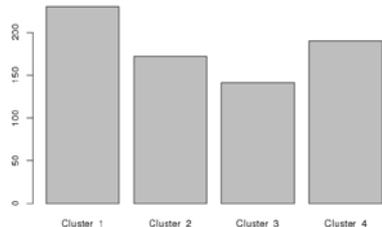
- Choose ADNI + PPMI as discovery cohorts for joint AD/PD stratification. Others for validation purposes.

Largely rely on genotype based stratification

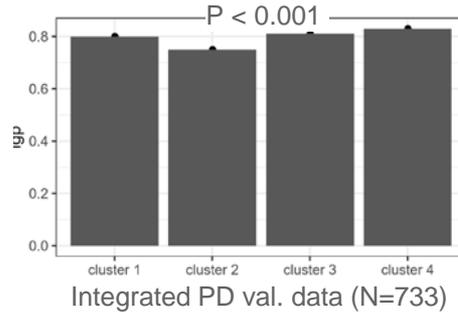
- Data availability
- Significant genetic disease component (Arenson et al., Journal of Genetics, 2018)

Other data used for biological contextualization

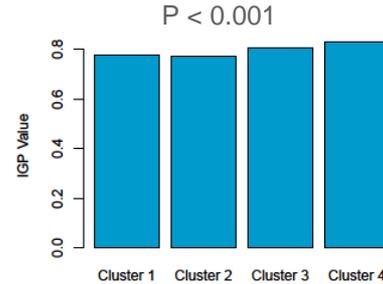
Clusters are confirmed in Validation Data



Classifier predicted clusters

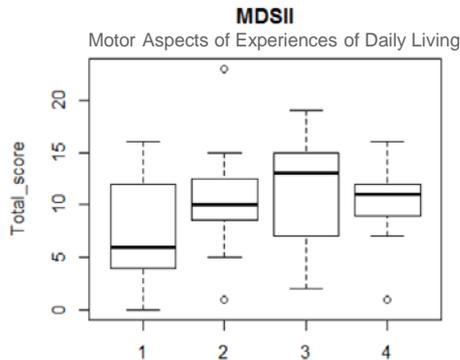


Integrated PD val. data (N=733)



Integrated AD val. data (N=237)

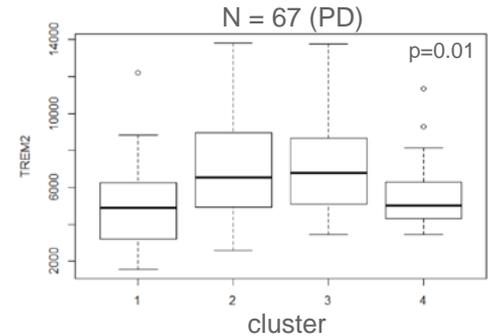
IGP = proportion of samples in a cluster, whose nearest neighbors are also in the same cluster. (Kapp & Tibshirani, 2007)



Downstream analysis

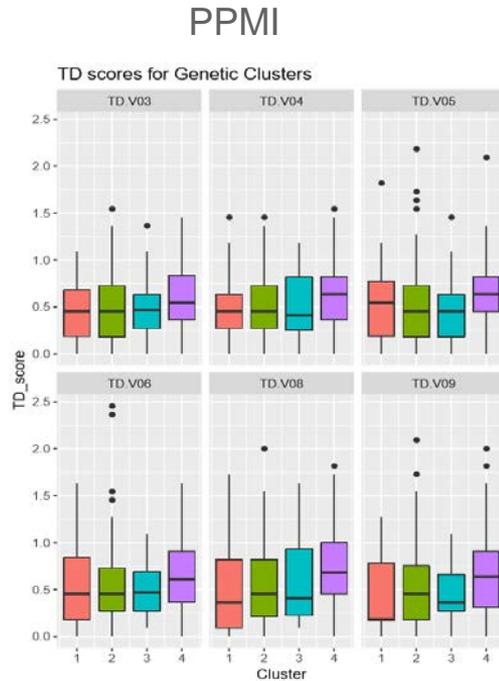
Differential inflammation markers:

- TREM2 (PD)
- MIF (PD)
- TYRO3 (PD)

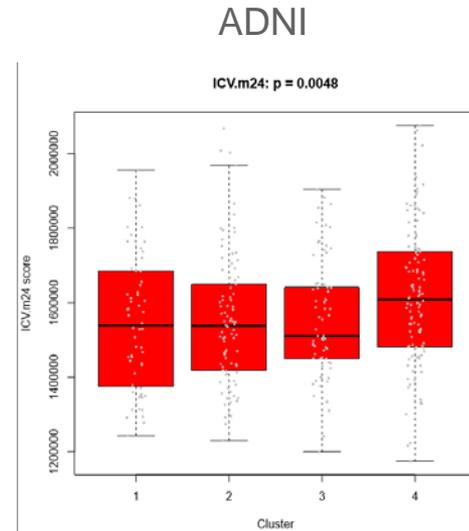


Clusters express differences in neuroinflammation and motor symptoms in PD

Association to Clinical Variables in PPMI and ADNI



$P < 0.01$ at all
time points



Association of brain volume shrinkage and
neuroinflammation: Datta et al., Brain 2017

Association of clusters to tremor severity in PD and inter-cranial brain volumes in AD.

Summary of Key Achievements

Novel Data Science methods confirm relevance of knowledge derived disease mechanisms

- predicting disease onset: AD risk model (Khanna et al., Sci Rep, 2018)
- mechanism based patient stratification

Stratification has been validated in integrated AD and PD cohorts

- Downstream analysis ongoing

Identified patient sub-groups differ in

- **neuroinflammation**
- motor symptoms (PD)
- inter-cranial brain volumes (AD)

A first step towards a mechanism based taxonomy of neurodegenerative diseases.

Acknowledgements

▪ Fraunhofer SCAI

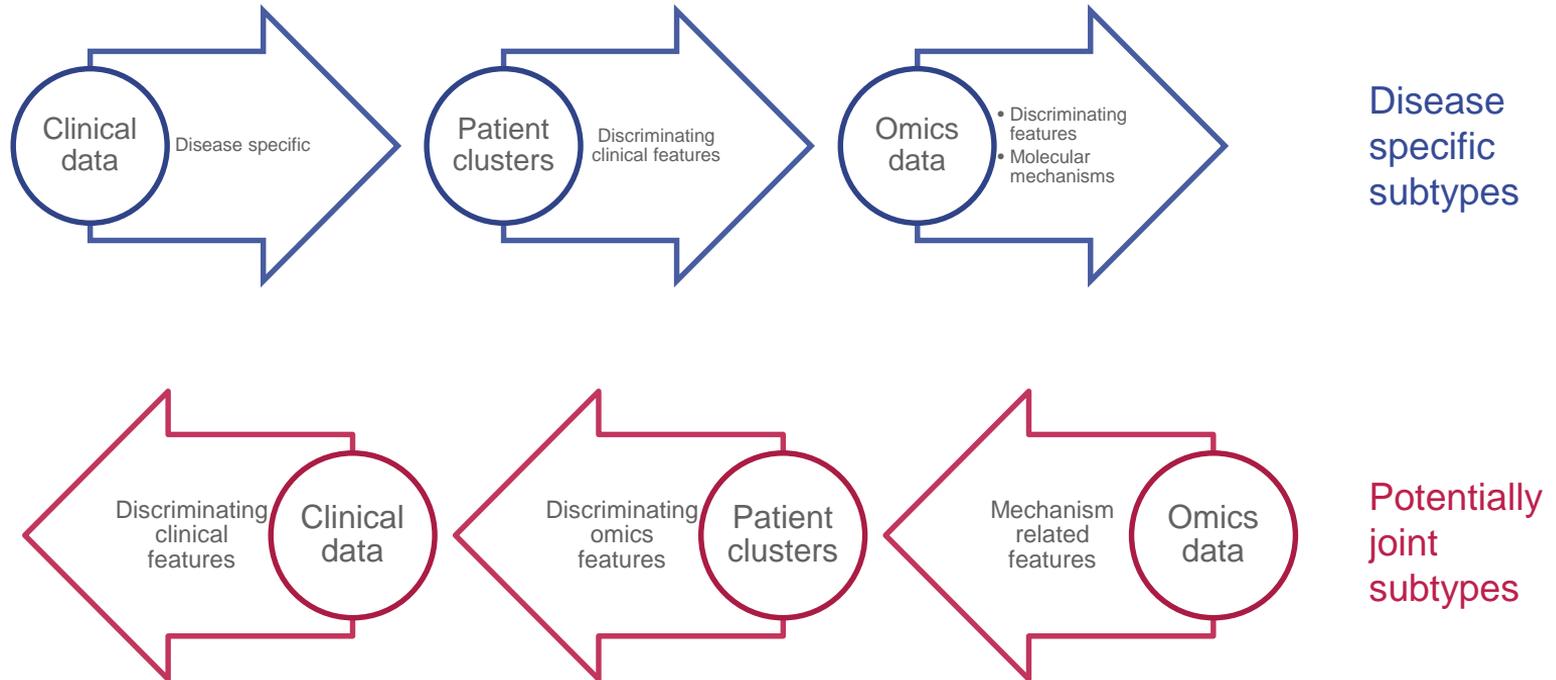
- Asif Emon
- Daniel Dominguez-Fernandez
- Shashank Khanna
- Anandhi Iyappan
- Colin Birkenbihl
- Reagon Kharki
- Martin Hofmann-Apitius

▪ UCB

- Ashley Heinson
- Ping Wu
- Johann de Jong
- Ashar Ahmad
- Phil Scordis

Backup

Options for Identifying a Mechanism Based Disease Subtypes



Joint AD/PD Clustering Approach for Genotype Data (Discovery)

Intersection of AD/PD OpenBEL graphs

- 15 shared cause-effect-relationship sub-graphs / „mechanisms“ (27 genes)

Map SNP Data

- ADNI (arrays), PPMI (WGS) with filters
- >1% MAF
- >90% genotyped
- Gene mapping of 148 common SNPs:
- Ensembl (proximity)
- Brain cis-eQTL mapping (GTex)

Aggregate: mechanism impact score

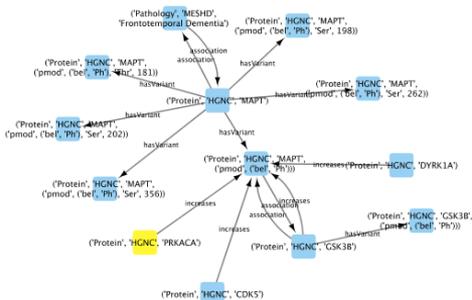
- none
- PCA
- ssGSEA
- Autoencoders

Cluster

- Sparse consensus NMF
- Internal evaluation:
 - silhouette index
 - PAC index

Comparison against healthy controls

- Classifier normal vs disease
- 10 x repeated 10-fold nested cross-validation



Validation with external data (example: ROSMAP)

Logistic regression classifier (discovery cohort)

- Cross-validated AUC (over-optimistic)

Assign ROSMAP patients to clusters

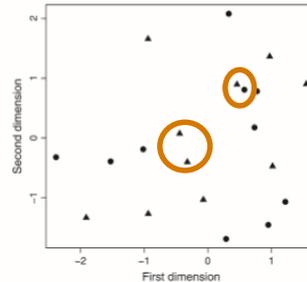
- Encode patients using 148 SNPs
 - ssGSEA
 - Autoencoder (trained with ADNI + PPMI)

Evaluation

- IGP measure (Kapp & Tibshirani, 2007)

Contextualization

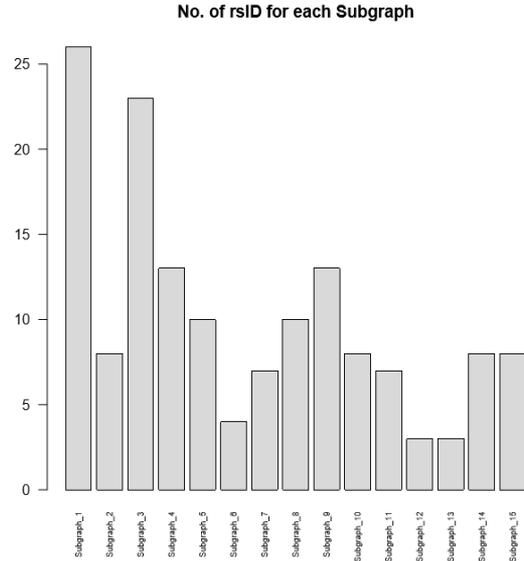
- Differential protein expression
- Differential gene expression
- Correlation with CpGs
- Correlation with histone marks (CHIPseq)



Mechanism Impact Score

SNP → gene mapping:

- proximity (Ensembl): +/- 50Kbp
- eQTLs (GTEx database)



	SNP1	SNP2	SNP3	...
Patient 1	0	0	2	
Patient 2	2	1	0	
Patient 3	1	2	0	
...				

	SNP1	SNP2	SNP3	...
Patient 1	0	0	2	
Patient 2	2	1	0	
Patient 3	1	2	0	
...				

Mechanism 1

Mechanism N

Aggregate scores of member SNPs:

- PCA
- Autoencoder networks
- ssGSEA

↓ Patient specific mechanism impact profile by

	Mechanism 1	Mechanism 2	Mechanism 3	...
Patient 1				
Patient 2				
Patient 3				
...				

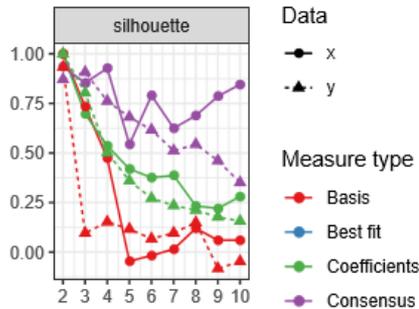
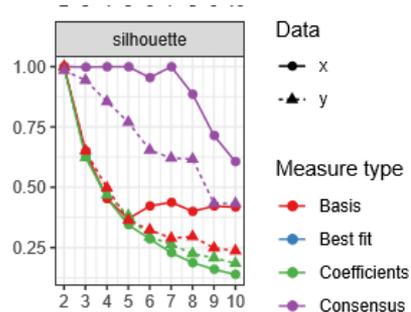
No robust clustering with raw genotype and PCA based aggregation

Raw genotype + sparse NMF

Clusters	AD	PD	Total	SI
Cluster 1	190	134	324	1
Cluster 2	296	224	520	1

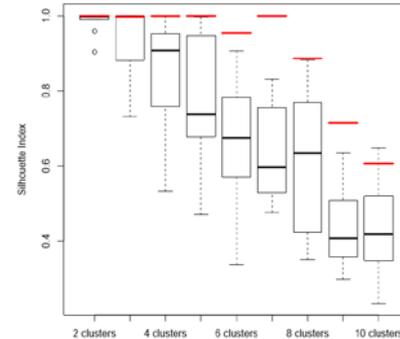
PCA + sparse NMF

Clusters	AD	PD	Total	SI
Cluster 1	102		102	1
Cluster 2	384	358	742	0,95
				0,975

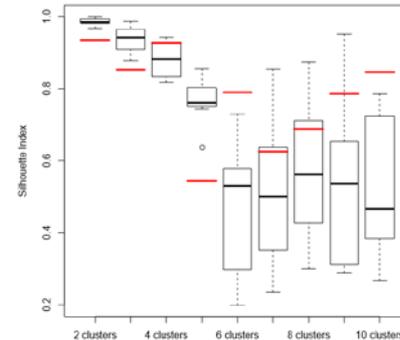


Dashed = 10 x randomly shuffled data

BoxPlots of original and random rawSNPs AD-PD clusters SI



BoxPlots of original and random PCA AD-PD clusters SI

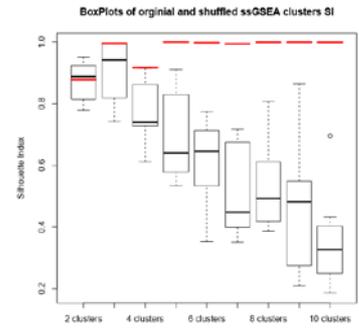
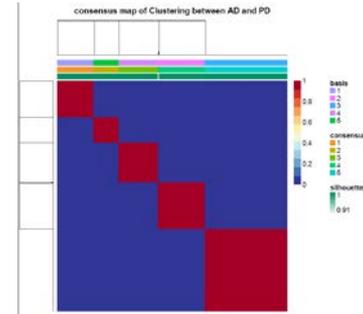
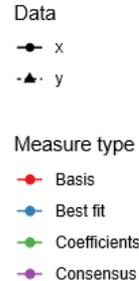
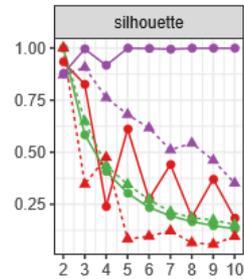


Clusters are probably random artifacts.

Robust clustering with ssGSEA and autoencoder based SNP aggregation

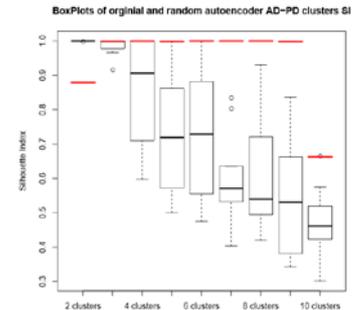
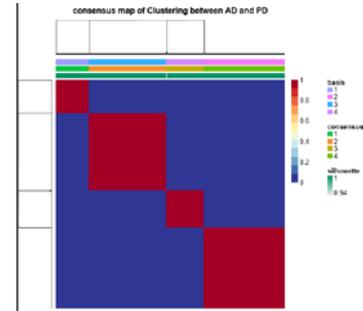
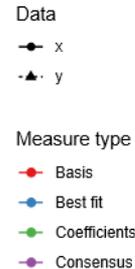
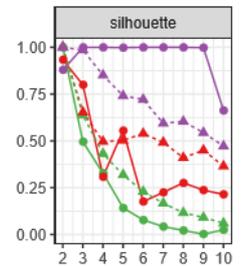
ssGSEA + sparse NMF

seed_ica_rank_2:10_nrun_50				
Clusters	AD	PD	Total	SI
Cluster 1	65	27	92	1
Cluster 2	80	65	145	1
Cluster 3	81	53	134	1
Cluster 4	88	83	171	1
Cluster 5	172	130	302	1



autoencoder + sparse NMF

autoencoder & sparse NMF				
Clusters	AD	PD	Total	SI
Cluster 1	95	27	122	1
Cluster 2	119	167	286	1
Cluster 3	106	32	138	1
Cluster 4	166	132	298	1

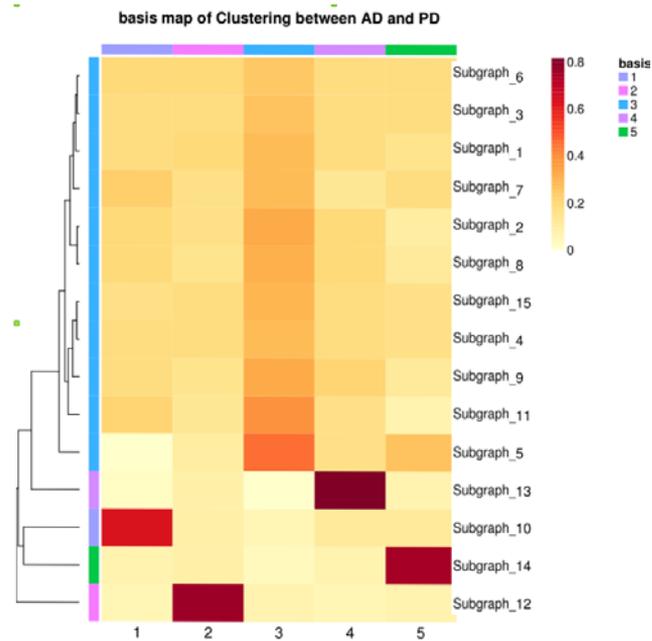


No. of cluster confirmed with PAC index analysis (Senbabaoglu et al., Sci Rep 2014).

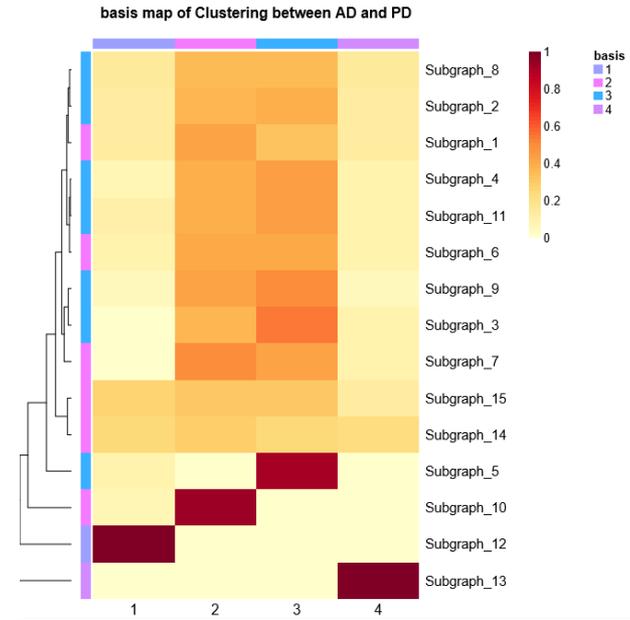
Clusters are highly stable and clearly better discriminated than random data

Association of Mechanisms to different Clusters

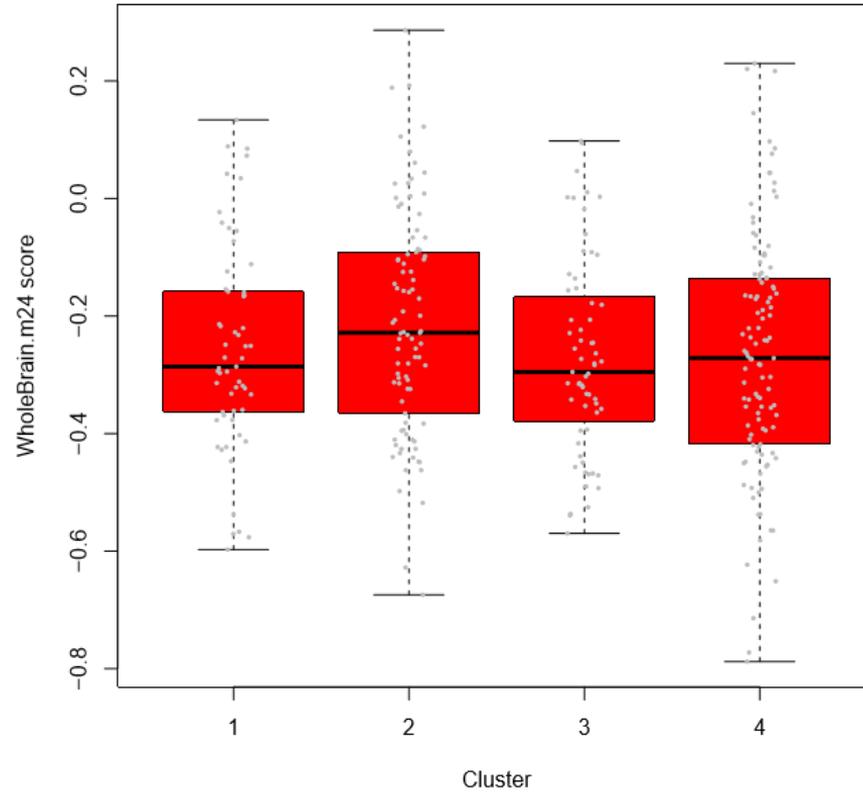
ssGSEA + sparse NMF



Autoencoder + sparse NMF



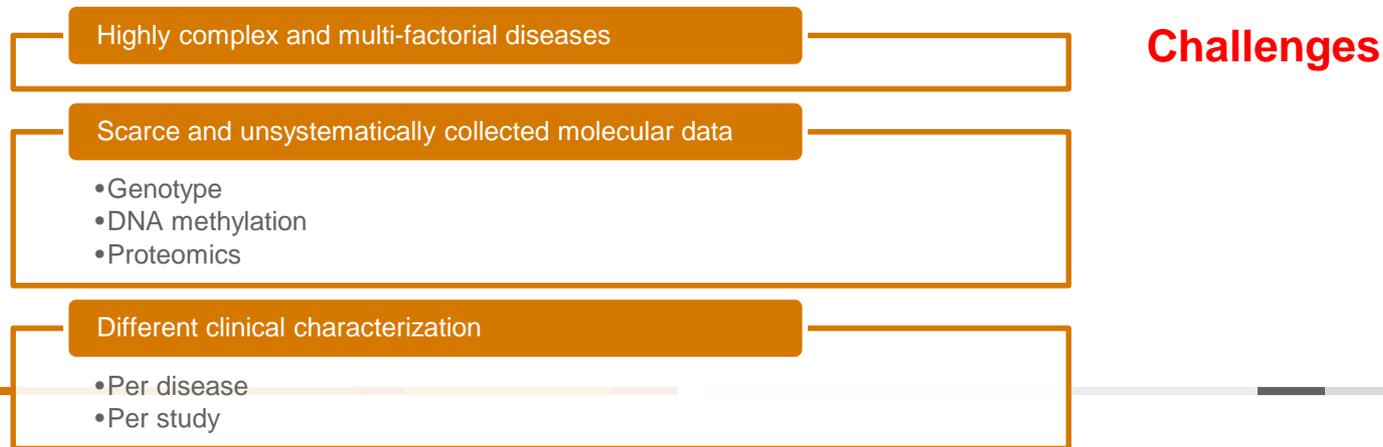
WholeBrain.m24: $p = 0.059$



AETIONOMY: Vision & Key Challenges

Developing a “mechanism-based” taxonomy of Alzheimer’s and Parkinson’s Disease

- **Current classification of neuro-degenerative diseases is purely phenotype based**
- **Vision: a molecular mechanism based classification**
 - Potentially new ways of treating patients
- **Project goal: first proof of principle**



Scientific Approach in AETIONOMY

