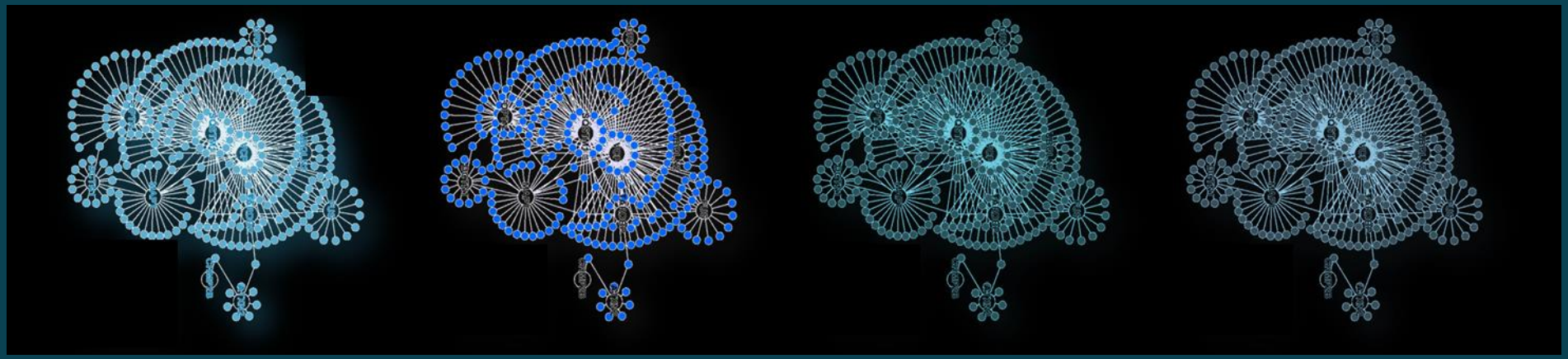
Using PPI networks to (i) define molecular underpinnings and (ii) prioritize genes in GWAS loci for Parkinson's Disease



Dr. Raffaele Ferrari on behalf of Dr. Claudia Manzoni

c.manzoni@reading.ac.uk

r.ferrari@ucl.ac.uk

University of Reading

2015

PeerJ. 2015 Feb 19;3:e778. doi: 10.7717/peerj.778. eCollection 2015.

Computational analysis of the LRRK2 interactome.

Manzoni C[1], Denny P[2], Lovering RC[2], Lewis PA[3].

2016

Brief Bioinform. 2016 Nov 22. pii: bbw114. [Epub ahead of print]

Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences.

Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, Ferrari R.

2016

Mol Neurodegener. 2016 Feb 24;11:21. doi: 10.1186/s13024-016-0085-4.

Frontotemporal dementia: insights into the biological underpinnings of disease through gene co-expression network analysis.

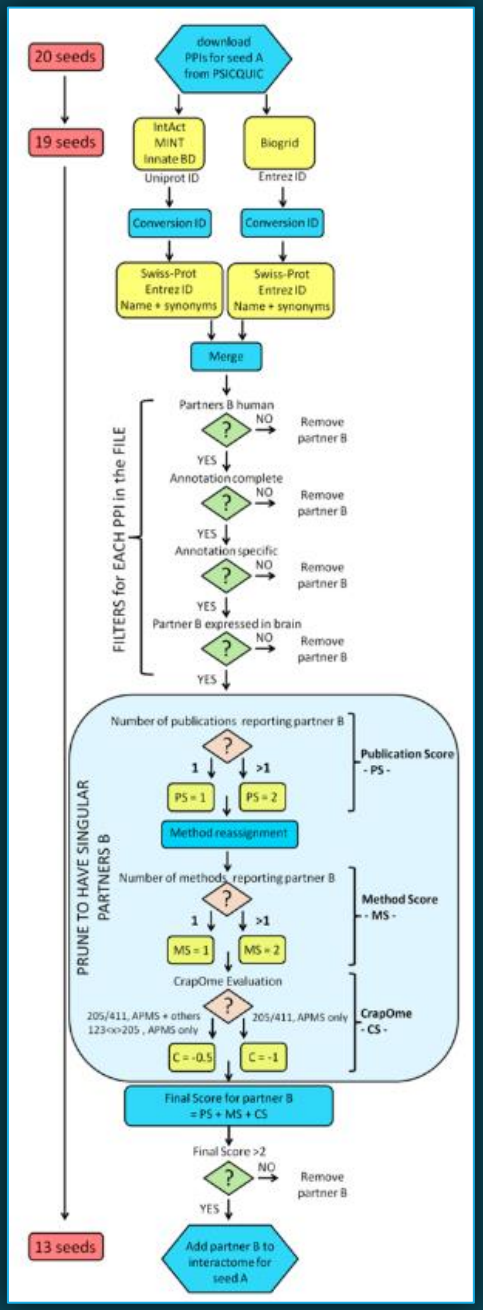
Ferrari R[1], Forabosco P[2], Vandrovcova J[3,4], Botía JA[5,6], Guelfi S[7,8], Warren JD[9]; UK Brain Expression Consortium (UKBEC), Momeni P[10], Weale ME[11], Ryten M[12,13], Hardy J[14].

2017

J Proteome Res. 2017 Feb 3;16(2):999-1013. doi: 10.1021/acs.jproteome.6b00934. Epub 2017 Jan 12.

Weighted Protein Interaction Network Analysis of Frontotemporal Dementia.

Ferrari R[1], Lovering RC[2], Hardy J[1], Lewis PA[1,3], Manzoni C[1,3].



TRANSPARENCY:
clear steps for filtering and thresholding

Filter for annotation **QC**

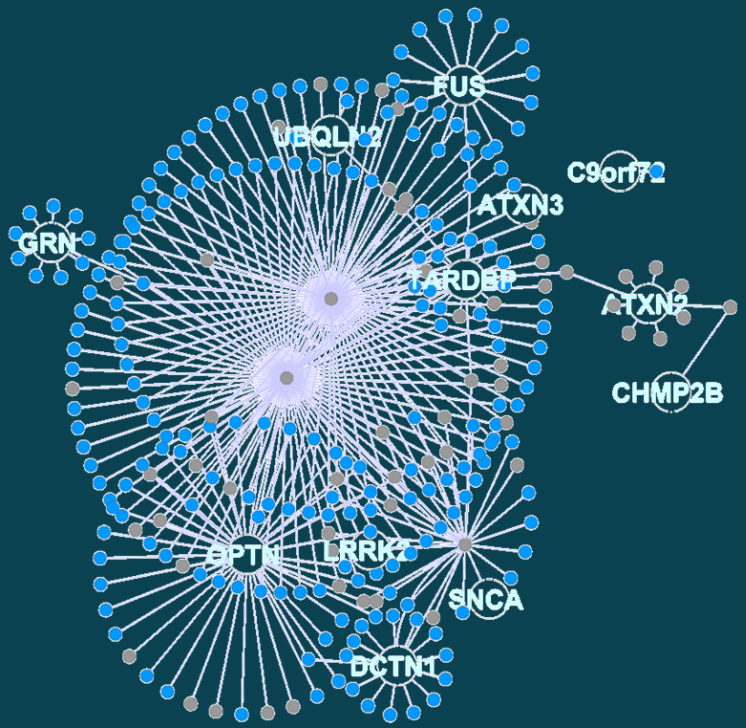
Filter for brain **expression**

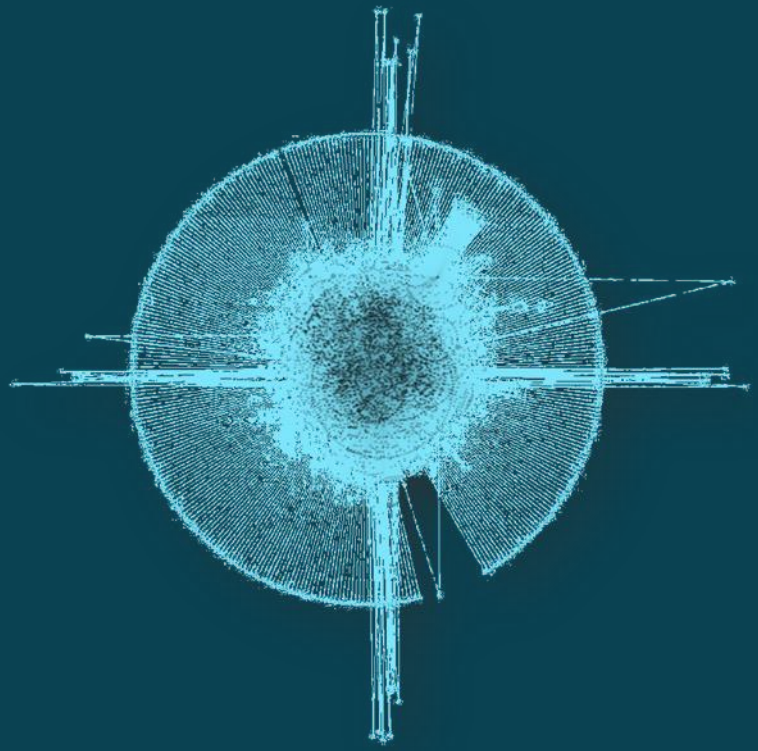
Weight based on # publication # techniques

Threshold to control for **reproducibility**

Only proven and confirmed PPI

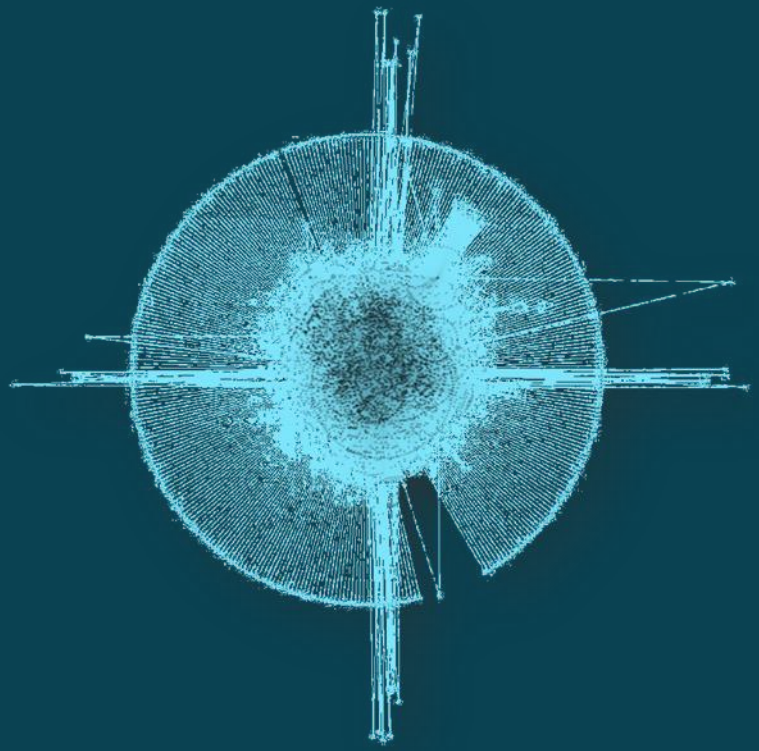
LAYERS: to dilute the seed centrality bias

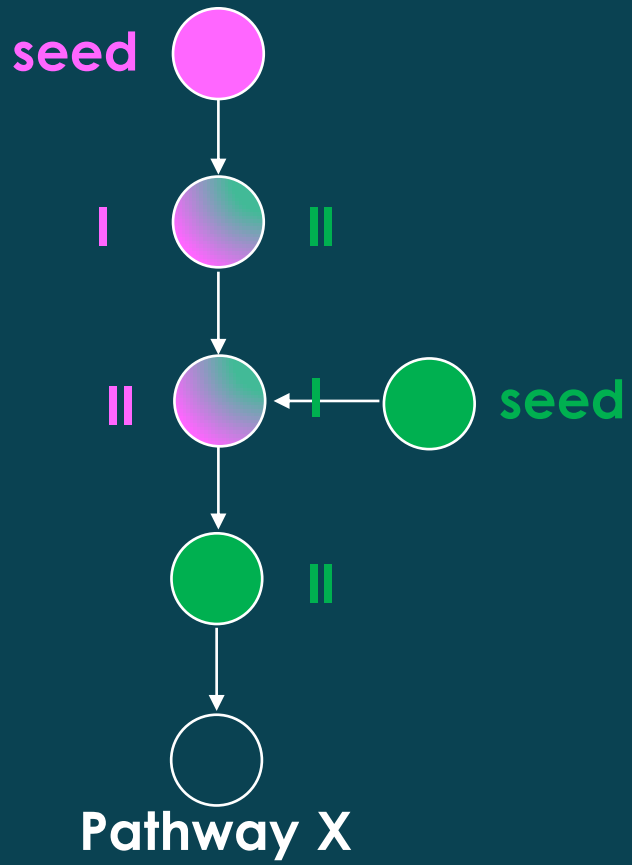




Building of the second layer to correct for the seed centrality bias: we are not just going bigger, we increase network **complexity – interconnectivity – density**

Hunting for pathways





Different Mendelian genes are causative for the same disease

Assumption:

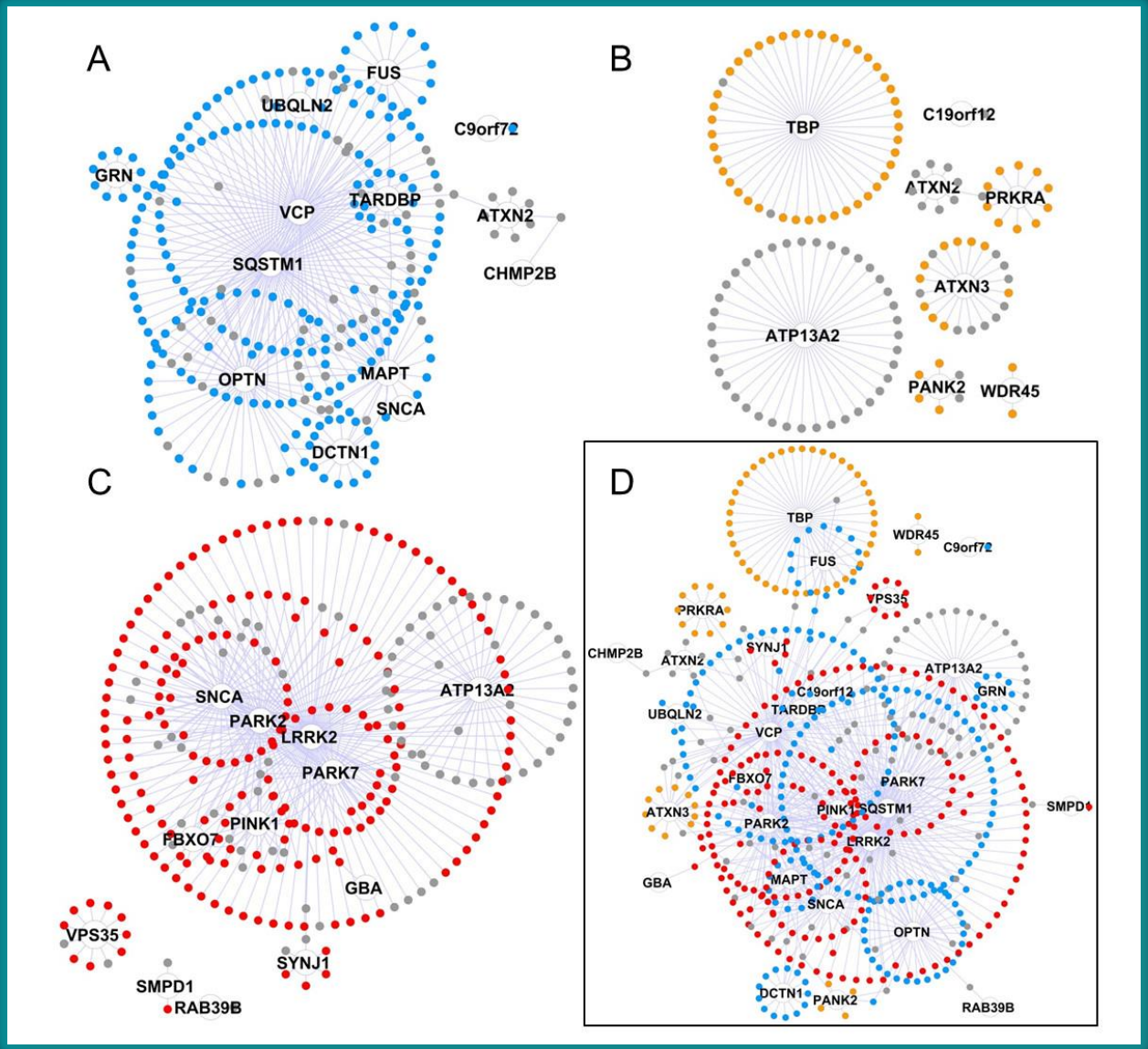
There should be similarities in terms of shared pathways.

These similarities can be inferred by looking at the interactomes

FTD

PS

PD



spinocerebellar ataxia (SCA)

neurodegeneration with brain iron accumulation (NBIA)

spastic paraplegia (SP)

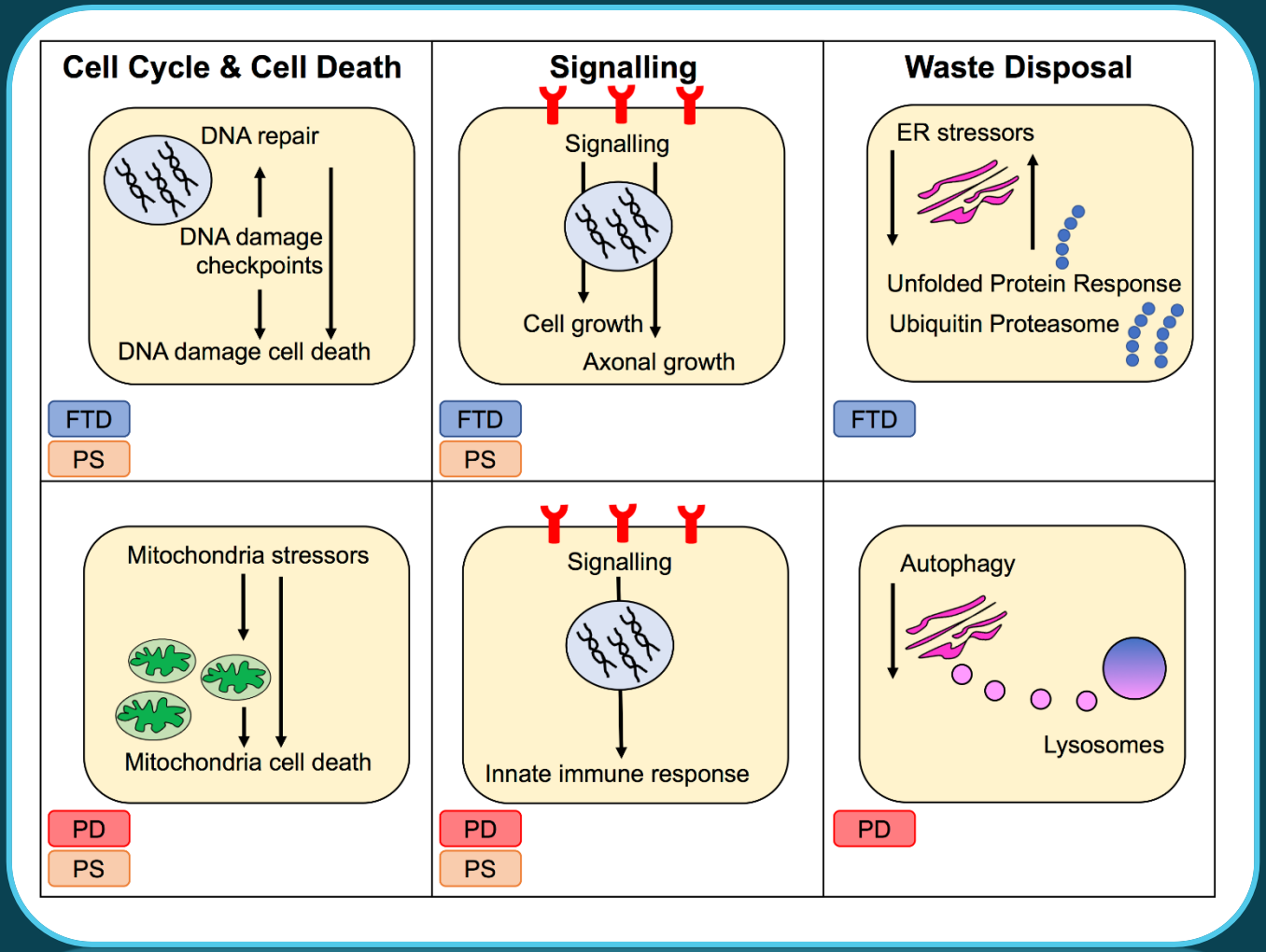
dystonia presenting with parkinsonism (DS)

Functional enrichment for:

- Single interactomes, i.e. 1st + 2nd layers
- Defined by inter-interactome hubs (IIH)

> syndrome/phenotype-specific functional enrichment

Using PPI networks to (i) define molecular underpinnings …… for Parkinson's Disease



*Ferrari et al. manuscript under revision

Focus on PD only now

Assumption:

Functions inferred from Mendelian networks are conserved in sporadic PD and can be used to evaluate (a posteriori) genetic data.

GWAS top SNPS:

Take the Bonferroni significant SNPs

Create the associated LD blocks ($r^2 \geq 0.5, 0.6, 0.7, 0.8$)

List all the ORFs in the LD blocks

Match the ORFs with the protein in the network that are responsible for the common, PD shared functions

Using PPI networks to … and (ii) prioritize genes in GWAS loci for Parkinson's Disease

part A

SNP	Chr	Position	Prev. Proposed Gene(s) based on proximity	Newly Proposed Gene(s) LD 0.5	Newly Proposed Gene(s) LD 0.6	Newly Proposed Gene(s) LD 0.7	Newly Proposed Gene(s) LD 0.8	Expression
rs35749011	**1**	**1:155135036**	**GBA/SYT11**				**GBA**	**A**
							SYT11	**A, N, O**
							ZBTB7B	A, M
							MUC1	A, M
							CLK2	A, N
							DAP3	N, M
							SCAMP3	A, N, O
							ADAM15	E
							RUSC1	N
					SHC1			A, M, E
				ARHGEF2				O
rs823118	**1**	**1:205723572**	**RAB7L1/NUCKS1**				**RAB7L1**	**A, M**
rs10797576	1	1:232664611	SIPA1L2					
rs6430538	2	2:135539967	ACMSD/TMEM163					
rs1474055	2	2:169110394	STK39					
rs115185635	3	3:87520857	KRT8P25/APOOP2					
rs12637471	3	3:182762437	MCCC1	DCUN1D1				A, N
rs34311866	4	4:951947	TMEM175/GAK/DGKQ					
rs11724635	4	4:15737101	BST1					
rs6812193	4	4:77198986	FAM47E/SCARB2					
rs356182	**4**	**4:90626111**	**SNCA**			**SNCA**		**N, O**
rs9275326	6	6:32666660	HLA-DQB1					
rs199347	7	7:23293746	GPNMB			IGF2BP3		N, O
rs117896735	**10**	**10:121536327**	**INPP5F**				**INPP5F**	**N**
				BAG3				A, M
rs3793947	11	11:83544472	**DLG2**				**DLG2**	**N, O**
rs329648	11	11:133765367	MIR4697					
rs76904798	**12**	**12:40614434**	**LRRK2**				**LRRK2**	**M, O**
rs11060180	12	12:123303586	CCDC62					
rs11158026	**14**	**14:55348869**	**GCH1**				**GCH1**	**M**
rs1555399	14	14:67984370	TMEM229B					
rs2414739	15	15:61994134	VPS13C					
rs14235	16	16:31121793	BCKDK/STX1B				KAT8	A, N
							STX4	M
rs17649553	**17**	**17:43994648**	**MAPT**				**MAPT**	**A, N, O**
							NSF	N
rs12456492	18	18:40673380	RIT2					
rs62120679	19	19:2363319	SPPL2B					
rs8118008	20	20:3168166	DDRGK1					

part B

rs34016896	3	3:160992864	NMD3		PPM1L			A, N
rs591323	8	8:16697091	FGF20					
rs60298754	8	8:89373041	MMP16					
rs7077361	10	10:15561543	ITGA8					
rs11868035	**17**	**17:17715101**	**SREBF/RAI1**				**RAI1**	**N**
				TOM1L2				A, N
rs2823357	21	21:16914905	USP25					

Ten/27 candidate genes (5 of which are Also Mendelian) – *GBA*, *SYT11*, *SNCA*, *RAB7L1*(*RAB29*), *INPP5F*, *DLG2*, *LRRK2*, *GCH1*, *MAPT*, and *RAI1* – previously selected by proximity were confirmed by our functional approach

The remaining 17/27 candidate genes – *ZBTB7B*, *MUC1*, *CLK2*, *DAP3*, *SCAMP3*, *ADAM15*, *RUSC1*, *SHC1*, *ARHGEF2*, *KAT8*, *STX4*, *IGF2BP3*, *PPM1L*, *DCUN1D1*, *NSF*, *BAG3* and *TOM1L2* – are to be regarded as novel PD candidate genes

SUMMARY:

This pipeline:

- efficiently serves in defining **disease-specific biological processes** on the basis of **known Mendelian genes** and providing a **list of proteins** involved in **disease-specific processes** that can be used to **prioritize candidate genes in GWAS loci**;

- helps expanding on the **genetic and functional architecture** underlying **idiopathic forms of disease**, and;

- provides the basis for:
  - **further genetic** (e.g. such list of proteins/genes might be screened for rare variants within NGS data sets), and;
  - **hypothesis-driven functional studies** to validate **risk pathways** as well as identify **targets** for the development of therapies in the future

Thank you for your attention!

University of Reading

Patrick Lewis

Claudia Manzoni

Institute of Neurology UCL

John Hardy

Raffaele Ferrari

Nick Wood

Demis Kia

Institute of Cardiovascular Sciences UCL

Ruth Lovering