

Integration of *in silico* and *in vitro* approaches to investigate genetic and functional drivers of frontotemporal lobar degeneration

Beatrice Costa

Supervisors:

Prof. Helene Plun-Favreau

Prof. John Hardy

Dr Raffaele Ferrari

Department of Neurodegenerative Disease
Queen Square Institute of Neurology
University College London (UCL)

A thesis submitted to University College London for the degree of Doctor of
Philosophy on the 08th of July 2022

Declaration

I, Beatrice Costa, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature: Beatrice Costa

Date: 08-07-2022

Statement of contribution

I have performed all procedures and experiments in this Thesis with the following contributions:

- **Chapter 3:** Dr Claudia Manzoni, School of Pharmacy, UCL, performed the statistical analyses on *C9ORF72* hexanucleotide repeat expansion prevalence and characterisation.
- **Chapter 5:** Jacqueline Casey, Queen Square Institute of Neurology, created the *GRN* KO iPSC neuronal and H4 *GRN* siRNA KD cell lines and performed western blotting on them.

Declaration form: referencing doctoral candidate's own published work in thesis

Please use this form to declare if parts of your thesis are already available in another format, e.g., if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

This form should be completed as many times as necessary. For instance, if a student had seven thesis chapters, two of which having material which had been published, they would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2):

a. Where was the work published? (e.g. journal name)

Neurology

b. Who published the work? (e.g. Elsevier/Oxford University Press)

Neurology

c. When was the work published?

17-09-20

d. Was the work subject to academic peer review? YES

e. Have you retained the copyright for the work? YES

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):

a. Where is the work intended to be published? (e.g. journal name)

b. List the manuscript's authors in the intended authorship order:

c. Stage of publication:

- Not yet submitted
- Submitted
- Undergoing revision after peer review
- In press

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):

B. Costa, M. Bernal-Quiros and J. Polke contributed to C9ORF72 expansions screening, data interpretation and drafting of the manuscript. D. A. Kia contributed to data interpretation and drafting the manuscript. R. Ferrari, C. Manzoni, V. Escott-Price contributed to project design, data analyses and interpretation, manuscript drafting. M. Aguilar, I. Alvarez, V. Alvarez, O. Andreassen, M. Anfossi, S. Bagnoli, L. Benussi, L. Bernardi, G. Binetti, D. Blackburn, M. Boada, B. Borroni, L. Bowns, G. Bråthen, A.C. Bruni, H.-H. Chiang, J. Clarimon, S. Colville, M.E. Conidi, T.E. Cope, C. Cruchaga, C. Cupidi, M.E. Di Battista, J. Diehl-Schmid, M. Diez-Fairen, O. Dols-Icardo, E. Durante, D. Flisar, F. Frangipane, D. Galimberti, M. Gallo, M. Gallucci, R. Ghidoni, C. Graff, J.H. Grafman, M. Grossman, J. Hardy, I. Hernández, G.J.T. Holloway, E.D. Huey, I. Illán-Gala, A. Karydas, B. Khoshnood, M.G. Kramberger, M. Kristiansen, P.A. Lewis, A. Lleó, G.K. Madhan, R. Maletta, A. Maver, M. Menendez-Gonzalez, G. Milan, B. Miller, M.O. Mol, P. Momeni, S. Moreno-Grau, C.M. Morris, B. Nacmias, C. Nilsson, V. Novelli, L. Öijerstedt, A. Padovani, S. Pal, Y. Panchbhaya, P. Pastor, B. Peterlin, I. Piaceri, S. Pickering-Brown, Y.A.L. Pijnenburg, A.A. Puca, I. Rainero, A. Rendina, A.M.T. Richardson, E. Rogaeva, B. Rogelj, S. Rollinson, G. Rossi, C. Rossmeier, J.B. Rowe, E. Rubino, A. Ruiz, R. Sanchez-Valle, S.B. Sando, A.F. Santillo, J. Saxon, E. Scarpini, M. Serpente, N. Smirne, S. Sorbi, E. Suh, F. Tagliavini, J.C. Thompson, J.Q. Trojanowski, V.M. Van Deerlin, J. Van der Zee, C. Van Broeckhoven, J. van Rooij, J.C. Van Swieten, A. Veronesi, E. Vitale, M.L. Waldö, C. Woodward and J. Yokoyama contributed to the critical review of manuscript for intellectual content, contribution of samples and demographics metadata.

4. In which chapter(s) of your thesis can this material be found?

Chapter 3

5. Candidate's e-signature:

Date: 24-06-22

6. Supervisor/senior author(s) e-signature:

Date: 24-06-22

Publications

- Zhang, D.*, Guelfi, S.*, Garcia-Ruiz, S., **Costa, B.**, Reynolds, R. H., D'Sa, K., Liu, W., Courtin, T., Peterson, A., Jaffe, A. E., Hardy, J., Botia, J. A., Collado-Torres, L., Ryten, M. (2020). Incomplete annotation has a disproportionate impact on our understanding of Mendelian and complex neurogenetic disorders. *Science Advances*. **6**, 24. DOI: 10.1126/sciadv.aay8299.
- **Costa, B.**, Manzoni, C., Bernal-Quiros, M., Aguilar, M., Alvarez, I., Alvarez, V., Andreassen, O., Anfossi, M., Bagnoli, S., Benussi, L., Bernardi, L., Binetti, G., Blackburn, D., Boada, M., Borroni, B., Bowns, L., Bråthen, G., Bruni, A. C., Chiang, H., Clarimon, J., Colville, S., Conidi, M. E., Cope, T. E., Cruchaga, C., Cupidi, C., Di Battista, M. E., Diehl-Schmid, J., Diez-Fairen, M., Dols-Icardo, O., Durante, E., Flisar, D., Frangipane, F., Galimberti, D., Gallo, M., Gallucci, M., Ghidoni, R., Graff, C., Grafman, J. H., Grossman, M., Hardy, J., Hernández, I., Holloway, G. J. T., Huey, E. D., Illán-Gala, I., International FTD-Genetics Consortium (IFGC), Karydas, A., Khoshnood, B., Kramberger, M. G., Kristiansen, M., Lewis, P. A., Lleó, A., Madhan, G. K., Maletta, R., Maver, A., Menendez-Gonzalez, M., Milan, G., Miller, B., Mol, M. O., Momeni, P., Moreno-Grau, S., Morris, C. M., Nacmias, B., Nilsson, C., Novelli, V., Öijerstedt, L., Padovani, A., Pal, S., Panchbhaya, Y., Pastor, P., Peterlin, B., Piaceri, I., Pickering-Brown, S., Pijnenburg, Y. A. L., Puca, A. A., Rainero, I., Rendina, A., Richardson, A., M., T., Rogaeva, E., Rogelj, B., Rollinson, S., Rossi, G., Rossmeier, C., Rowe, J. B., Rubino, E., Ruiz, A., Sanchez-Valle, R., Sando, S. B., Santillo, A. F., Saxon, J., Scarpini, E., Serpente, M., Smirne, N., Sorbi, S., Suh, E., Tagliavini, F., Thompson, J. C., Trojanowski, J. Q., Van Deerlin, V. M., Van der Zee, J., Van Broeckhoven, C., van Rooij, J., Van Swieten, J. C., Vitale, E., Waldö, M. L., Woodward, C., Yokoyama, J., Escott-Price, V., Polke, J. M., Ferrari, F. (2020). *C9ORF72*, AAO, and genetic ancestry discriminate behavioural from language variant FTLD syndromes. *Neurology*. **95** (24) e3288-e3302; DOI: 10.1212/WNL.0000000000010914.
- **Costa, B.**, Botia, J., Lewis, P. A., Hardy, J., Ferrari, R.*, Plun-Favreau, H.*, Manzoni, C*. Integration of *in silico* and *in vitro* approaches to investigate genetic and functional drivers of frontotemporal lobar degeneration. *Manuscript in preparation*.

- Manzoni, C., **Costa, B.**, International FTD-Genetics Consortium (IFGC), Escott-Price, V., Ferrari, R. An IFGC FTLD-GWAS study. *Manuscript in preparation.*

Presentations and posters

- **Poster presentation** at the Vesicle trafficking and pathways to neurodegeneration online conference, '*Integration of in silico and in vitro approaches to investigate genetic and functional drivers of frontotemporal lobar degeneration.*' (17th-19th April 2021)
- **Oral presentation** at the 2nd International *C9ORF72* meeting, '*Frontotemporal lobar degeneration: age at onset, C9ORF72 expansions and ancestry discriminate behavioural from language variant syndromes*' (30th-31st January 2020, London, UK).
- **Oral and poster presentation** at the 8th Winter Seminar on Dementia and Neurodegenerative Disorders, '*Frontotemporal lobar degeneration: age at onset, C9ORF72 expansions and ancestry discriminate behavioural from language variant syndromes*' (22nd-24th January 2020, Bressanone, Italy).
- **Poster presentation** at the 2018 Queen Square Symposium '*Frontotemporal dementia: in depth characterization of C9ORF72 expansion prevalence in FTD subtypes and bioinformatics analysis*' (29th May 2018, London, UK).

Acknowledgements

My deepest thanks go to all the individuals and organisations that contributed to the completion of this thesis and PhD. I would like to thank my supervisor, Prof. H el ene Plun-Favreau, for offering her continued support and feedback while allowing me the independence to grow as a researcher. I would also like to thank Dr Raffaele Ferrari, who gave me the opportunity to join this department and constantly supported me throughout both my BSc and PhD projects during my years at UCL, I would not have started this PhD without him. I have had the pleasure and privilege of working with Prof. John Hardy, whose formidable deep scientific knowledge and never-ending wit and smile have always been a source of inspiration. Words cannot express my gratitude to Dr Claudia Manzoni who, without being my supervisor, followed my work with sympathy and interest, and has been key to the success of this project: her knowledge and guidance regarding bioinformatics and network analysis shaped much of this thesis. I would also like to thank Jackie Casey for her contributions in wet-lab work.

Additionally, this endeavour would not have been possible without the generous support from the Alzheimer's Society, who financed my research.

I also want to say a huge thank you to all the postdocs and PhD students, past and present, who made the lab an enjoyable place to work. Special thanks go to Dr Benjamin O'Callaghan, a great scientist whose help and guidance in the lab saved me countless times and whose contagious enthusiasm pushed me to work harder. Sam, Katie, Amy, Capucine: I will always be grateful that our desks were allocated as they were, you made the lab an enjoyable place to work! And Andrea, Leslie, Sonia, Mark, Francois, Roberto, Jonah, Darija, Riccardo, Eva, Michael, Alberto and many more – I will miss you all.

Finally, I would like to thank my family and friends for their support through the highs and lows of the last three years. First and foremost, to my beautiful family: my beloved mamma e pap a, for always being there for me, when I need them and especially when I don't; my brother Chicco, with his genuine and contagious enthusiasm for all kinds of science, you remind me every day what a true scientist looks like; my sister Dida, always the funniest person in room who can make me laugh even in the darkest hours; your love and support have been a priceless gift during this PhD; my dear nonni, Laura, Paolo, Gabriella, Piera and Tonino, you are a constant source of inspiration; my favourite uncle and aunt Zigozago and Claudia, my biggest fans and supporters 'fin da small'; Kathi, Annina, Panna, Ilaria, Alice, my best friends, whose constant love and emotional support

through this PhD have been invaluable; David and Regina, my dear friends and colleagues, for the morning coffees in Holland Park and lunches at Zayed that kept me going, and for your endless patience in teaching and helping me with R coding and telling me about the wonders of bioinformatics, you made me a better scientist. And finally, but most importantly, to my querido Sebas. Your love, friendship and unwavering support and belief in me have consistently encouraged me to achieve this goal; your passion, enthusiasm and experience accompanied me through this journey and shaped my present and future as a scientist. Thanks for brightening the gloomiest day and for being a constant source of inspiration. I could not have undertaken nor finished it without you, so this PhD is yours as much as it is mine.

Abstract

Frontotemporal lobar degeneration (FTLD) is a common form of early-onset dementia presenting with a complex gene and pathological architecture. Impairment in several molecular pathways has been associated with FTLD pathology however translating genetic knowledge into functional understanding of impacted biological processes in complex disorders still represents a major challenge.

This thesis proposes a multi-omics network approach integrating genetics, transcriptomics and proteomics to investigate the common pathogenic mechanisms and their associated genes/proteins underlying Mendelian and sporadic forms of FTLD. Protein-protein interaction (PPINs) and gene co-expression (GCNs) networks were used to prioritise: i) disease-specific biological processes on the basis of known FTLD Mendelian and GWAS genes and ii) their associated genes/proteins, which were carried forward for hypothesis-driven functional validation using *in vitro* cellular models. A separate study specifically investigated *C9ORF72* repeat expansions (i.e., a common genetic signature of FTLD) as a potential genetic modifier of FTLD syndromes in relation to genetic ancestry and age at onset (AAO).

Results from these studies revealed that waste disposal, and autophagy/mitophagy dysfunction in particular, are among the most relevant biological processes impacted in FTLD. Further work indicated that *CDC37*, protein with no reported link to the mitochondria, has an important role in modulating PINK1-dependent mitophagy, as *CDC37* deficient cells exhibited an increase in both *PINK1* mRNA and protein expression as well as PINK1-dependent phosphorylation of ubiquitin, a common marker of mitophagy; additionally, *CDC37* deficiency was shown to increase transcriptional expression of *GRN*, a common FTLD Mendelian gene. Finally, an iPSC *GRN*^{-/-} model showed that both *CDC37* protein expression and mitophagy were downregulated upon *GRN* insufficiency, suggesting a shared mechanism between *GRN*, PINK1 and *CDC37* where *GRN* modulation might be regulating the trafficking and stability of PINK1 via *CDC37*. Additionally, *C9ORF72* analyses indicated a correlation between pathogenic *C9ORF72* expansions, AAO, principal component analysis (PCA)-based Central/Northern European ancestry, and a diagnosis of bvFTLD, implying complex genetic risk architectures differently underpinning the behavioural and language variant syndromes.

These findings contribute to our understanding of PINK1-dependent mitophagy regulation via *CDC37* and *GRN* and further emphasise the relevance of autophagy/mitophagy dysfunction in FTLD.

Impact statement

In an era where scientific research is the golden standard for novel discoveries and technological progress, one may wonder whether it is as reliable as we think. Several studies across many fields have brought to light that only less than half of published findings can be replicated reliably^{1,2}, leading many to consider whether there should be a paradigm shift in the way we do research, hence introducing the concept of triangulation¹(**Figure a**).



Figure a. Triangulation as a strategy to produce robust research. Different approaches are needed to address the same underlying question or more simply said, to prove a point using many distinct lines of evidence. Adapted from 'Robust research needs many lines of evidence' by M. R. Munafò, 2018, *Nature*. Copyright 2022 Springer Nature Limited. Adapted with permission.

Triangulation can be defined as the 'strategic use of multiple approaches to address a single question', implying that results that are consistent across different methodologies are less likely to be artefacts¹. This concept has been dramatically changing the way we investigate human diseases, because rather than viewing a disease as an independent entity, it recognises the interplay of multiple molecular processes that interact in a complex network and are reflected in the expressed pathophenotype^{3,4}. When translating this concept into the study of the brain, we realise that inter- and intracellular interconnectivity implies that the impact of a specific genetic abnormality is not restricted to the activity of the gene product that carries it but can spread along the links

of the network and alter the activity of gene products that otherwise carry no defects. Therefore, an understanding of a gene's network context is crucial in determining the phenotypic impact of defects that affect it.

In this study, I propose a holistic network approach in line with the concept of triangulation where I integrate genetics, transcriptomics and proteomics to identify candidate FTLN-risk biological processes and associated genes/proteins. The characterisation of different networks (i.e., PPIN and GCN) behaviour as a whole is an essential component of this project as it allows to look at the same problem through different lenses and thus to improve our understanding of disease complexity at the cellular and molecular levels. This approach sheds light on how genetic and molecular perturbations propagate through the system by identifying and characterising network modules and hubs connecting apparently different pathophenotypes that can be targeted for clinical intervention. Through this *in silico* study of a multifactorial, complex neurodegenerative disorder, I show that this pipeline can be applied to: i) define disease-specific biological processes on the basis of known Mendelian and GWAS genes; ii) identify genes/proteins involved in disease-specific processes. While these genes/proteins can be carried forward for functional validation, they can also be used to prioritise candidate genes in GWAS loci or be screened for rare variants within whole exome sequencing datasets, giving rise to several manifold downstream applications. Additionally, it is important to mention that this approach can be implemented in and facilitate the investigation of any complex disease, providing a resource that can be used by researchers in any field of biomedicine.

This approach presents a robust pipeline to dissect the genetic and functional architecture underlying polygenic forms of disease. In doing so, it provides the basis for further genetic and hypothesis-driven functional studies to validate disease-risk pathways as well as identify targets for the future development of therapeutic interventions. On the basis of this approach, it is clear that to cure a specific disease it is of paramount importance to understand the cells' global organization — the 'think globally, act locally' paradigm of network medicine⁴.

Finally, in this study I provided proof of concept that machine learning can be leveraged to segregate FTLN clinical subtypes. Additionally, the comprehensive analysis of *C9ORF72* repeat expansion size has shown strong correlations with age at onset and

ancestry, revealing novel genetic modifiers of FTLD. These results warrant further characterisation of genetic, environmental and additional clinical variables to fine-tune models that further characterise FTLD genetic risk architecture and are able to predict disease outcome to complement diagnostic criteria.

Table of Contents

DECLARATION	3
STATEMENT OF CONTRIBUTION	4
DECLARATION FORM: REFERENCING DOCTORAL CANDIDATE’S OWN PUBLISHED WORK IN THESIS.....	5
PUBLICATIONS	7
PRESENTATIONS AND POSTERS	8
ACKNOWLEDGEMENTS.....	9
ABSTRACT.....	11
IMPACT STATEMENT	13
TABLE OF FIGURES	21
TABLE OF TABLES	24
ABBREVIATIONS.....	25
CHAPTER 1.....	31
1. INTRODUCTION	31
1.1. <i>Frontotemporal dementia</i>	31
1.1.1. Diagnostic criteria and clinical features	31
1.1.2. Neuropathology.....	35
1.1.3. Genetics	39
1.1.3.1. Familial FTLD-Mendelian genetics.....	39
1.1.3.2. Sporadic FTLD	42
1.1.4. Molecular mechanisms of FTLD pathogenesis	45
1.1.4.1. Protein quality control systems	48
1.1.4.2. Autophagy and proteostasis in neurodegenerative diseases	54
1.1.4.3. Mitochondria and mitochondrial quality control mechanisms.....	57
1.1.4.4. PINK1-dependent mitophagy.....	60
1.1.5. Modelling FTLD <i>in vitro</i>	62
1.2. <i>Emerging technologies powering the genetic and molecular dissection of complex diseases</i>	66

1.2.1. Using networks for biological inference	66
1.2.2. Network theory and topological properties.....	68
1.2.3. Modularity in biological networks	69
1.2.4. Gene co-expression networks.....	70
1.2.5. Protein-protein interaction networks	71
1.3. <i>Ongoing and future efforts to tackle FTLD: the rise of systems-level computational approaches.....</i>	72
1.4. <i>Significance of the research and objectives of this thesis.....</i>	75
CHAPTER 2.....	77
2. MATERIAL AND METHODS	77
2.1. <i>C9ORF72 repeat expansion screening</i>	<i>77</i>
2.1.1. Cohort, clinical phenotyping and patients consent	77
2.1.2. Genotyping, C9ORF72 repeat expansions and analysis cohorts	77
2.1.3. Statistical analyses	79
2.2. <i>Bioinformatics.....</i>	<i>81</i>
2.2.1. Weighted Gene Co-expression Network Analysis (WGCNA)	81
2.2.2. Weighted Protein – Protein Interaction Network analysis (WPPINA)	82
2.2.3. Functional enrichment analysis	83
2.2.4. Evaluation of gene expression in brain	83
2.2.5. Software.....	84
2.2.6. Re-evaluation of WGCNA networks.....	84
2.3. <i>Molecular biology.....</i>	<i>84</i>
2.3.1. Materials.....	84
2.3.2. Quantitative PCR.....	85
2.3.2.1. RNA extraction	85
2.3.2.2. Reverse transcription.....	85
2.3.2.3. Quantitative real-time PCR.....	85
2.3.3. Cell culture.....	86
2.3.3.1. Materials.....	86
2.3.3.2. Culture of immortalised cell lines.....	87
2.3.3.3. Differentiation of iPSC into cortical neurons.....	87
2.3.3.4. Characterisation of cortical neurons derived from PSCs.....	89

2.3.3.5. Freezing and thawing.....	89
2.3.3.6. Cell counting.....	89
2.3.3.7. siRNA-mediated gene silencing.....	89
2.3.3.8. Mitophagy induction.....	90
2.3.4. Protein biochemistry.....	90
2.3.4.1. Materials.....	90
2.3.4.2. Harvesting and lysing of cells.....	92
2.3.4.3. Mitochondrial fractionation for biochemistry.....	92
2.3.4.4. Protein electrophoresis and western blotting.....	92
2.3.5. Immunocytochemistry.....	93
2.3.5.1. Materials.....	93
2.3.5.2. Cell staining.....	93
2.3.5.3. Visualisation of protein relocation and image analysis.....	94
2.3.6. Live cell imaging: mt-Keima.....	95
2.3.6.1. Materials.....	95
2.3.6.2. Cells treatment and imaging.....	95
2.3.7. Statistical analysis.....	95
CHAPTER 3.....	96
3. RESULTS.....	96
3.1. <i>C9ORF72, AAO, and genetic ancestry discriminate behavioural from language variant FTLD syndromes.....</i>	96
3.1.1. Introduction.....	96
3.1.2. <i>C9ORF72</i> repeat expansions.....	96
3.2. <i>Results.....</i>	97
3.2.1. <i>C9ORF72</i> expansions frequency and syndromes.....	97
3.2.2. <i>C9ORF72</i> expansions (and repeat counts [rc]) and genetic ancestry.....	98
3.2.3. <i>C9ORF72</i> repeat expansions (and counts [rc]) and age at onset (AAO)....	102
3.2.4. Syndrome prediction.....	105
3.3. <i>Discussion and conclusions.....</i>	105
CHAPTER 4.....	109
4. RESULTS.....	109
4.1. <i>Stratification of FTLD-risk candidate genes using WPPINA and WGCNA.....</i>	109

4.1.1. Introduction.....	109
4.1.2. WPPINA and WGCNA	109
4.2. <i>Results</i>	111
4.2.1. Seeds selection.....	111
4.2.2. Construction of the gene co-expression networks (WGCNA)	112
4.2.3. Construction of the PPI FTLD network (WPPINA)	114
4.2.4. Functional enrichment and prioritisation of FTLD-risk genes/pathways	121
4.2.5. Re-evaluation of WGCNA to improve FTLD <i>in vitro</i> modelling.....	133
4.3. <i>Discussion</i>	137
CHAPTER 5.....	144
5. RESULTS	144
5.1. <i>Validation of FTLD-risk candidate genes in vitro</i>	144
5.1.1. Introduction.....	144
5.1.1.1. The prioritised genes/proteins: STUB1, CUL2 CDC37	146
5.1.1.2. Modelling PINK1-dependent mitophagy	148
5.1.2. Results.....	149
5.1.2.1. <i>CDC37</i> KD increases pUb (Ser65) accumulation in SH-SY5Y cell line and increases <i>PINK1</i> mRNA expression	149
5.1.2.2. <i>CDC37</i> KD does not increase pUb (Ser65) accumulation in H4 cell line 157	
5.1.2.3. <i>CDC37</i> KD increases mitochondrial clearance after 12h in mt-Keima but not in WB	158
5.1.2.4. <i>CDC37</i> KD increases <i>GRN</i> mRNA expression but decreases its protein expression in FTLD iPSC model.	165
5.1.3. Discussion.....	172
CHAPTER 6.....	178
6. DISCUSSION AND FUTURE DIRECTIONS.....	178
6.1. <i>General limitations</i>	178
6.2. <i>Future directions in FTLD research</i>	180
6.3. <i>Current therapeutic approaches to treat FTLD and the waste disposal pathway as a therapeutic target</i>	183
6.4. <i>Concluding remarks</i>	186

REFERENCES..... 187

Table of Figures

FIGURE A. TRIANGULATION AS A STRATEGY TO PRODUCE ROBUST RESEARCH.....	12
FIGURE 1.1. CLINICAL PRESENTATIONS WITHIN THE FRONTOTEMPORAL LOBAR DEGENERATION PATHOLOGICAL SPECTRUM.....	34
FIGURE 1.2. DIFFERENT CLASSIFICATIONS OF FRONTOTEMPORAL LOBAR DEGENERATION.....	38
FIGURE 1.3. LANDSCAPE OF THE GENES ASSOCIATED WITH THE FTLD-ALS SPECTRUM.....	42
FIGURE 1.4. SCHEMATIC OVERVIEW OF COMMON FAMILIAL AND RARE FTLD GENES AND GENETIC RISK FACTORS.....	43
FIGURE 1.5. SUMMARY OF ALL PUTATIVE IMPACTED MOLECULAR PROCESSES CONTRIBUTING TO FTLD PATHOGENESIS ON THE BASIS OF THE CURRENT GLOBAL GENETIC CHARACTERIZATION OF FTLD INCLUDING MENDELIAN, SPORADIC AND GWAS FTLD GENES.....	46
FIGURE 1.6. ILLUSTRATIVE REPRESENTATION OF PROTEIN QUALITY CONTROL MECHANISMS IN THE CELL.	50
FIGURE 1.7. THE AUTOPHAGY LYSOSOMAL PATHWAY.	52
FIGURE 1.8. GENETIC LINKS BETWEEN AUTOPHAGY AND NEURODEGENERATIVE DISEASE.....	55
FIGURE 1.9. MAJOR PATHWAYS OF MITOCHONDRIAL QUALITY CONTROL.....	58
FIGURE 1.10. PINK1-DEPENDENT AND MITOPHAGY.	62
FIGURE 1.11. MEASURING EXISTING OR PREDICTED RELATIONSHIPS AMONG GENES/PROTEINS.	68
FIGURE 2.1. STUDY COHORTS	78
FIGURE 2.2. ANALYSIS WORKFLOW.....	80
FIGURE 2.3. CORTICAL NEURONS DIFFERENTIATION PROTOCOL. DIV, DAYS <i>IN VITRO</i>	88
FIGURE 3.1. ASSOCIATION BETWEEN AAO AND: ANCESTRY; SYNDROME; EXPANSION LENGTH.	103
FIGURE 3.2. PATIENT SUBPOPULATIONS (bvFTLD AND PPA SYNDROMES) BASED ON C9ORF72 EXPANSIONS GENETIC SIGNATURES AND ANCESTRY.	106
FIGURE 4.1. HOLISTIC BIOINFORMATICS APPROACH EMPLOYED IN THIS CHAPTER.	110
FIGURE 4.2. WORKFLOW EXEMPLIFYING THE BIOINFORMATICS PIPELINE IMPLEMENTED IN THE PRESENT STUDY.	112
FIGURE 4.3. GENE CO-EXPRESSION PATTERNS FOR THE FTLD GENES (SEEDS) AND THEIR RELEVANCE WITHIN MODULES.....	113
FIGURE 4.4. ENTIRE-FTLD-PPIN.	115
FIGURE 4.5. INTER-INTERACTOMES DEGREE DISTRIBUTION.....	116
FIGURE 4.6. WORKFLOW TO GENERATE THE CORE-FTLD-PPIN.....	120
FIGURE 4.7. FUNCTIONAL ENRICHMENT ANALYSIS OF CORE-FTLD-PPIN.....	122

FIGURE 4.8. FUNCTIONAL ENRICHMENT ANALYSIS OF GCNs.....	124
FIGURE 4.9. FUNCTIONAL ENRICHMENT FOR WASTE DISPOSAL OF WPPINA AND WGCNA.....	127
FIGURE 4.10. TOPOLOGICAL FEATURES AND FUNCTIONAL PATTERNS OF THE ATG - MITO/UPS-FTLD-PPIN.....	129
FIGURE 4.11. WORKFLOW TO PRIORITISE FTLD-RISK CANDIDATE GENES AND VENN DIAGRAM SHOWING THE THREE PRIORITISED NODES (STUB1, CUL2 AND CDC37).....	131
FIGURE 4.12. DENDROGRAM COMPARING INTER-MODULAR RELATIONSHIPS AMONG GTEx V6 MODULES.....	135
FIGURE 4.13. DENDROGRAM COMPARING INTER-MODULAR RELATIONSHIPS AMONG UKBEC MODULES.....	136
FIGURE 5.1. SCHEMATIC REPRESENTATION HIGHLIGHTING THE INVOLVEMENT OF VARIOUS GENETIC RISK FACTORS ASSOCIATED TO ALS/FTD IN DEFECTIVE MECHANISMS OF NUTRIENT SENSING AND AUTOPHAGY.....	145
FIGURE 5.2. <i>CDC37</i> KD INCREASES PUb (SER65) ACCUMULATION IN POE SH-SY5Y CELLS (WCL).....	151
FIGURE 5.3. <i>CDC37</i> KD INCREASES PUb (SER65) ACCUMULATION IN POE SH-SY5Y CELLS (MITOCHONDRIA-ENRICHED FRACTIONS).....	153
FIGURE 5.4. <i>CDC37</i> KD DOES NOT SIGNIFICANTLY INCREASE PUb (SER65) ACCUMULATION IN POE SH-SY5Y CELLS (IF).....	155
FIGURE 5.5. <i>CDC37</i> KD INCREASES <i>PINK1</i> mRNA EXPRESSION IN POE SH-SY5Y WHILE TRANSCRIPTIONAL LEVELS OF <i>CDC37</i> REMAIN UNCHANGED UPON <i>PINK1</i> KD.....	156
FIGURE 5.6. <i>CDC37</i> KD DOES NOT SIGNIFICANTLY INCREASE PUb (SER65) ACCUMULATION IN H4 CELLS (WCL).....	157
FIGURE 5.7. <i>CDC37</i> KD INCREASES MITOCHONDRIAL CLEARANCE.....	159
FIGURE 5.8. <i>CDC37</i> KD DOES NOT SIGNIFICANTLY INCREASE PUb (SER65) ACCUMULATION IN POE SH-SY5Ys AT DIFFERENT TIMEPOINTS (WCL).....	161
FIGURE 5.9. <i>CDC37</i> KD DOES NOT SIGNIFICANTLY INCREASE PUb (SER65) ACCUMULATION IN POE SH-SY5Ys AT DIFFERENT TIMEPOINTS (IF).....	164
FIGURE 5.10. <i>CDC37</i> KD INCREASES <i>GRN</i> mRNA EXPRESSION IN POE SH-SY5Y.....	166
FIGURE 5.11. <i>GRN</i> KD INCREASES MITOCHONDRIAL CLEARANCE.....	167
FIGURE 5.12. MITOPHAGY AND <i>CDC37</i> EXPRESSION DECREASE IN <i>GRN</i> ^{-/-} iPSC-DERIVED NEURONS (WCL).....	170

FIGURE 5.13. CDC37 EXPRESSION IS NOT SIGNIFICANTLY DECREASED IN *GRN* siRNA KD H4s (WCL).

..... 171

Table of Tables

TABLE 2.1. QUANTITATIVE PCR THERMAL CYCLING PROTOCOL (SYBR).	86
TABLE 2.2. CHARACTERISATION OF CELL LINES USED. iNDI, iPSC NEURODEGENERATIVE DISEASE INITIATIVE.	86
TABLE 2.3. LIST OF ANTIBODIES FOR IMMUNOBLOTTING.	91
TABLE 2.4. PRIMARY ANTIBODIES USED FOR IMMUNOFLUORESCENCE.	93
TABLE 2.5. EXCITATION AND EMISSION WAVELENGTH FOR IMMUNOCYTOCHEMISTRY.	94
TABLE 3.1. FREQUENCY OF EXPANSION CARRIERS IN THE ENTIRE COHORT AND BY SYNDROME.	97
TABLE 3.2. FREQUENCY OF EXPANSION CARRIERS IN THE 'NORDIC' AND 'MEDITERRANEAN' CLUSTERS.	99
TABLE 3.3. STRATIFIED FISHER'S EXACT TESTS COMPARING PREVALENCE OF PATHOGENIC EXPANSIONS ACROSS bvFTLD AND PPA AND THE 'NORDIC' AND 'MEDITERRANEAN' CLUSTERS.	101
TABLE 4.1. MENDELIAN (FAMILIAL) AND GWAS FTLD-ALS GENES USED AS SEEDS FOR DOWNSTREAM BIOINFORMATICS ANALYSES.	111
TABLE 4.2. INTER-INTERACTOME HUBS (IIHS) (N=26) INTERACTING WITH ≥ 12 SEEDS INTERACTOMES.	118
TABLE 4.3. ABSOLUTE VALUES AND ASSOCIATED PERCENTAGES OF GENE LOSSES IN GPROFILER FOR EACH MODULE.	125
TABLE 4.4. COMPARING CO-EXPRESSION PATTERNS BETWEEN PINK1 (MITO MARKER) AND THE PRIORITISED GENES IN FCTX.	133
TABLE 4.5. COMPARING CO-EXPRESSION PATTERNS BETWEEN THE FTLD MENDELIAN AND THE PRIORITISED GENES IN FCTX.	137

Abbreviations

AAO	Age At Onset
AAV	Adeno associated vectors
AD	Alzheimer's Disease
aFTLDDU	atypical Frontotemporal Lobar Degeneration with Ubiquitin-positive inclusions
AGD	Argyrophilic Grain Disease
ALP	Autophagy – Lysosome Pathway
ALS	Amyotrophic Lateral Sclerosis
ATP	Adenosine triphosphate
BIBD	Basophilic Inclusion Body Disease
bp	base-pair
BPAN	β -Propeller Protein-Associated Neurodegeneration
Braineac	Brain eQTL Almanac
bvFTD	Behavioural Frontotemporal Dementia
CBS	Cortico Basal Syndrome
CHCHD10	Coiled-coil-Helix-Coiled-coil-Helix Domain-containing protein 10
CHMP2B	Charged Multivesicular body Protein 2B
CMT2	Charcot-Marie-Tooth disease 2
CNS	Central Nervous System
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRi	CRISPR interference
CSF	Cerebrospinal fluid
C9ORF72	Chromosome 9 Open Reading Frame 72
DA	Dopaminergic neurons
dCAS9	dead Cas9
DCTN1	Dynactin 1
DDR	DNA Damage Response
DIV	Days <i>in vitro</i>
DLB	Dementia with Lewy Bodies
DMEM	Dulbecco's modified Eagle medium
DN	Dystrophic Neuritis
DPI	Data Processing Inequality

DRP1	Dynamin-Related Protein 1
DTT	Dithiothreitol
$\Delta\psi_m$	Mitochondrial membrane potential
ELISA	Enzyme-linked immunosorbent assay
eQTL	Expressed Quantitative Trait Loci
ER	Endoplasmatic Reticulum
ERAD	ER-associated degradation
FBS	Foetal bovine serum
FCCP	Carbonyl cyanide 4-(trifluoromethoxy)phenylhydrazone
FDR	False Rate Discovery
FET	FUS (fused in sarcoma), EWS (Ewing sarcoma) and TAF15 (TATA binding associated factor 15)
FTD	Frontotemporal Dementia
FTD-MND	Frontotemporal Dementia with Motor Neuron Disorder
FTLD	Frontotemporal Lobar Degeneration
FTLD-UPS	Frontotemporal Lobar Degeneration with inclusions positive for Ubiquitin Proteasome System markers
FUS	Fused in Sarcoma
GDNF	Glial cell line-Derived Neurotrophic Factor
GEP	Gene Expression Profile
GGT	Globular Glial Tauopathy
GO	Gene Ontology
GO-BP	Gene Ontology – Biological Process
GRCh	Genome Reference Consortium Human reference
GRN	Granulin precursor
GRNs	Granulins
GTE _x	Genotype-Tissue Expression
GWAS	Genome Wide Association Study
HD	Huntington's Disease
HEK293	Human Embryonic Kidney 293
HKs	Hexokinases
HLA-DRA	Major Histocompatibility Complex, Class II, DR Alpha
hLGDB	the Human Lysosome Gene Database

HPF	Helene Plun-Favreau
HSC70	Heat Shock Cognate 71
HSP	Hereditary Spastic Paraplegia
H4	Human neuroglioma
IB	Immunoblotting
IBM	Inclusion Body Myopathy
ICC	Immunocytochemistry
IF	Immunofluorescence
IFT74	Intraflagellar Transport 74
IFGC	International Frontotemporal dementia Genomics Consortium
IIH	Inter-Interactome Hub
iNDI	iPSC Neurodegenerative Disease Initiative
IPSCs	Induced-Pluripotent Stem Cells
KD	Knock-Down
KO	Knock-Out
LAMP2A	Lysosome-associated membrane glycoprotein 2
LOOCV	Leave-One-Out Cross Validation
LRRK2	Leucine-Rich Repeat Kinase 2
MAD	Mitochondrial-associated degradation
MAF	Minor Allele Frequency
MAPT	Microtubule Associated Protein Tau
MFN1	Mitofusin 1
MFN2	Mitofusin 2
MI	Mutual Information
mQTL	Methylation Quantitative Trait Loci
mRBP	mRNA-binding proteins
mtUPR	Mitochondrial unfolded protein response
MPP	Mitochondrial Processing Peptidase
MTS	Mitochondrial Targeting Sequence
NCBI	National Center for Biotechnology Information
nfvPPA	Non-fluent variant Primary Progressive Aphasia
NGN2	Neurogenin 2
NGS	Next Generation Sequencing

NIFID	Neuronal Intermediate Filament Inclusion Disease
O/A	Oligomycin/Antimycin A
OPA1	Optic Atrophy 1
OPTN	Optineurin
OR	Odds Ratio
ORF	Open Reading Frame
OXPPOS	Oxidative phosphorylation
PARL	Presenilin-Associated Rhomboid-Like protein
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Solution Tween
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PC12	Pheochromocytoma 12 cells
PD	Parkinson's Disease
PDB	Paget Disease of the Bone
PFA	Paraformaldehyde
PGRN	Progranulin
PiD	Pick's Disease
PINK1	Pten-induced Kinase 1
PNFA	Progressive Non-Fluent Aphasia
POE	Parkin Over-Expressing
PPA	Primary Progressive Aphasia
PPAOS	Primary progressive apraxia of speech
PSA	PsammalyseA
PSCs	Pluripotent stem cells
PSICQUIC	the Proteomics Standard Initiative Common QUery InterfaCe
PSP	Progressive Supranuclear Palsy
PTM	Post-Translational Modification
PVDF	Hydrophilic polyvinylidene fluoride
QC	Quality control
qPCR	quantitative Polymerase Chain Reaction
RAB38	Member RAS Oncogene Family
RBP	RNA-binding proteins

RC	Repeat Counts
RNA-Seq	RNA Sequencing
ROS	Reactive Oxygen Species
RRM-1	RNA Recognition Motif 1
rtvFTD	Right temporal variant Frontotemporal dementia
SCR	Scrambled control
scRNAseq	single-cell RNA sequencing
SD	Semantic Dementia
SGs	Stress granules
sgRNA	single guide RNA
SMAD	Small mother against decapentaplegic
snRNAseq	single-nucleus RNA sequencing
SNP	Single Nucleotide Polymorphism
SQSTM1	Sequestosome 1
StD	Standard Deviation
svPPA	Semantic variant Primary Progressive Aphasia
TARDBP	TAR DNA-Binding Protein
TBK1	TANK Binding Kinase 1
TBS	Tris-buffered saline
TDP-43	TAR DNA-Binding protein 43
TF	Transcription Factors
TIA1	TIA1 Cytotoxic Granule Associated RNA Binding Protein
TIM23	Translocase of inner membrane 23
TMEM106B	Transmembrane Protein 106B
TOM20	Translocase of outer membrane 20
TPM	Transcripts Per Million
TWAS	Transcriptome Wide Association Study
UBQLN2	Ubiquilin 2
UKBEC	UK Brain Expression Consortium
ULK	Unc-51-like kinase
UPR	Unfolded Protein Response
UPS	Ubiquitin Proteasome System
VCP	Valosin-Containing Protein

VDACs	Voltage-Dependent Anion Channels
WCL	Whole Cell Lysate
WES	Whole Exome Sequencing
WGCNA	Weighted Gene-Co-expression Network Analysis
WGS	Whole Genome Sequencing
WHO	World Health Organisation
WPPINA	Weighted Protein-Protein Interaction Network Analysis
WT	Wild-Type

Chapter 1

1. Introduction

1.1. Frontotemporal dementia

Frontotemporal dementia (FTD) is a clinically and pathologically heterogeneous group of incurable non-Alzheimer dementias characterised collectively by progressive atrophy of the frontal or temporal lobe, or both, which results in several cognitive deficits^{5,6}. Cases of FTD have first been described at the end of the 19th century by Arnold Pick, who lent his name to the historical designation of the entire FTD spectrum as Pick's disease⁷. Research over the past three decades however has shed light on the clinical, genetic and pathological complexity of these diseases as well as its unique selective brain degeneration patterns.

Based on North American and European epidemiological studies, FTD is the second most common form of dementia after Alzheimer's disease (AD) representing 10-20% of all dementia cases. FTD occurs in about three to 15 per 100 000 individuals with its onset typically beginning in the sixth decade of life, although several cases have been reported as early as the third or as late as the ninth decade^{5,8,9}. Although FTD is substantially less common than AD, this disease group is of disproportionate importance as a cause of early-onset dementia and, as a result, of the socioeconomic and human costs that it entails⁵. With an ageing population worldwide, this represents a substantial social and economic problem and novel therapeutic strategies are urgently needed. A great deal has already been learnt about the pathology and pathogenesis of the disease, but there are still gaps in our knowledge and these remain important areas for current and future research.

1.1.1. Diagnostic criteria and clinical features

FTD is characterised by several clinical alterations which include gradual changes in personality, behaviour, and speech output¹⁰. To date, researchers have been able to develop multiple methodologies to classify FTD disorders using biological and molecular analytical tools. However, diagnosis is still obtained clinically with the help of consensus criteria, due to the absence of definitive biomarkers¹¹. The main aim of establishing diagnostic criteria is to achieve uniformity in disease definition and to develop a sensitive

standard for early disease screening and management¹¹. The consensus criteria published by Neary et al. in 1998 identify two major prototypical syndromes produced by FTD, defined on the basis of leading features at presentation: these are the behavioural variant (bvFTD) and the language variants (primary progressive aphasia [PPA], which include semantic dementia (SD or semantic variant primary progressive aphasia [svPPA]) and progressive non-fluent aphasia (PNFA or non-fluent variant primary progressive aphasia [nfvPPA])^{5,12,13}. All the three variants share an insidious onset and inexorably progressive but variable decline, and each one is associated with topographically distinct cerebral involvement¹⁴. A list of the main clinical features of FTD have been summarised below.

Behavioural variant FTD (bvFTD). This syndrome is characterised by marked change in personality and social conduct, often associated with executive dysfunction and language impairment as secondary features. Patients may as well be affected by loss of volition, inertia, emotional blunting and loss of insight that can later develop into minimal speech output or even mutism¹⁴. Although behavioural changes and cognitive deficits represent bvFTD predominant features, small proportions of patients have shown to also have visual hallucinations and prominent parkinsonism.

Semantic dementia (SD). Semantic dementia is diagnosed when the dominant presenting features are loss of word and object knowledge and comprehension deficits, while speech output remains fluent and effortless, albeit with a tendency for repetition¹. Comprehension impairment can be associated with varying degrees of prosopagnosia (loss of facial recognition). It is noteworthy that behavioural changes are also observed in SD but tend to be somewhat distinct from the ones seen in bvFTD. Josephs¹² further divides SD into two separate variants based on which temporal lobe shows more severe atrophy: the right temporal variant, rather than the left one, holds behavioural changes as a typical presenting feature.

Progressive non-fluent aphasia (PNFA). Non-fluent speech production, agrammatism and telegraphic speech are the major characteristics defining the third syndrome, known as progressive non-fluent aphasia¹. Phonologic errors and difficulties in word retrieval, reading and writing are also part of the clinical picture. Whilst behavioural and personality changes typical of bvFTD are less common in PNFA, the development of extrapyramidal features is very often observed, sometimes leading to changes in diagnosis.

Over the past two decades, many case reports have recognised extrapyramidal features as prominent part of disease diagnostics; extensive overlap among the three above mentioned syndromes and progressive supranuclear palsy (PSP), corticobasal syndrome (CBS), and motor neuron disease (MND) has in fact been noted. In some reported cases of bvFTD, features of PSP are observed, suggesting that bvFTD and PSP could be in the same spectrum. Features of PSP are akinesia and rigidity, vertical supranuclear gaze palsy, and early falls. Milder changes in personality and behaviour can also occur in PSP, while apathy can be a salient feature, hence the overlap with bvFTD.

Akinesia and rigidity also typically characterise CBS, albeit in an asymmetric manner. A combination of non-fluent speech, limb apraxia, and behavioural and personality changes characterises CBS, hence the overlap with the FTD syndromes.

In parallel, the disease can also exhibit clinical features typical of parkinsonian and movement disorders, including MND. Some FTD cases may include features of MND, such as difficulty in speaking and swallowing, spasticity, weakness, hyperactive reflexes etc.; whereas MND is rare in SD, PNFA, PSP and CBS, it has been frequently associated with bvFTD, hence the term FTD-MND. FTD-MND displays rapid progression, with death occurring after approximately 2 years since onset. Behavioural and cognitive deficits are usually observed earlier than MND clinical features, however a small subset of patients with MND has or can later develop features suggestive of FTD. A summary of the main overlapping clinical features of FTLD are shown in **Figure 1.1**.

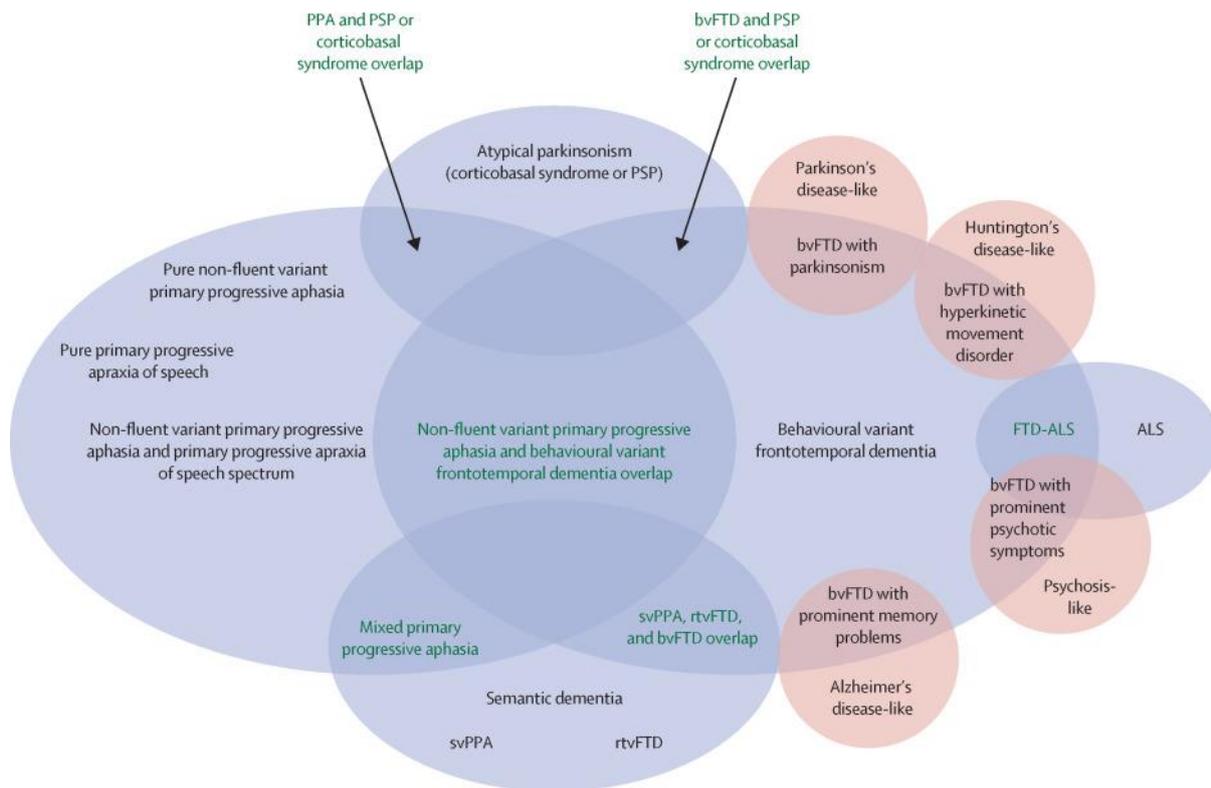


Figure 1.1. Clinical presentations within the frontotemporal lobar degeneration pathological spectrum. The core syndromes within the FTLT pathological spectrum are represented in blue, while syndromes depicted in red represent clinical presentations typically outside of the FTLT spectrum that have been found in people with pathologically confirmed FTLT. Note that the size of each circle is not intended to reflect the approximate prevalence of each syndrome. PPAOS, primary progressive apraxia of speech; rtvFTD, right temporal variant frontotemporal dementia. Adapted from 'Advances and controversies in frontotemporal dementia: diagnosis, biomarkers, and therapeutic considerations' by B. F. Boeve, 2022, *The Lancet Neurology*, Volume 21, Issue 3. Copyright 2022 by Elsevier Inc. Adapted with permission.

Probable FTD diagnosis would need to meet the consensus criteria, but also includes imaging suggesting of specific atrophy topography on magnetic resonance imaging or computed tomography measurements¹¹. Alternatively, positron emission tomography or single-photon emission computed tomography imaging demonstrating hypoperfusion or hypometabolism in specific brain regions^{5,11}. Histopathological and genetic evidence of FTD are eventually assessed in order to diagnose definite FTD pathology^{5,11}.

FTD shows, to some extent, overlap with the diagnostics of numerous neurodegenerative diseases, such as Alzheimer's disease (AD) and Parkinson's disease (PD). This has led to the suggestion that multiple neurodegenerative disorders share a number of features and molecular underpinnings¹⁶. For example and of note, significant evidence suggests positive genetic correlation between amyotrophic lateral sclerosis (ALS) and FTD¹⁷. Over the past decade it has been established that the TAR DNA binding protein 43 (TDP-43) is

the major pathological hallmark for the vast majority of both ALS and FTD cases^{17,18}. This inference strongly supports the significant overlap of genetic and molecular features characterizing these diseases.

1.1.2. Neuropathology

Degeneration of the frontal and temporal cerebral lobes is a relatively consistent feature and the term frontotemporal lobar degeneration (FTLD) is routinely used to describe the pathological conditions that commonly present in FTD cases leading to a wide range of clinical symptoms^{8,9,19,20}. As a result, the abbreviation FTLD will be used throughout the rest of this thesis.

Pathologically, the brain of FTLD patients shows marked shrinkage and spongiform morphology due to progressive neuronal loss during the course of the disease. Within the three main syndromes of FTLD there are distinct patterns of atrophy: bvFTLD cases show atrophy affecting the bilateral frontal lobes, specifically the medial frontal lobes and the anterior temporal lobes. In SD, asymmetric atrophy is detected at the bilateral level in the middle, inferior and medial anterior temporal lobe, whereas PNFA mainly displays atrophy of the left posterior frontal and insular regions²¹. Degeneration may occasionally involve the basal ganglia as well as the substantia nigra.

The advent of immunohistochemistry made evident that many different pathologies underlie the FTLD syndromes and associated FTLD-related disorders, such as PSP, CBS, and MND. It is in fact now well established in the field that more than 15 different pathologies can underlie FTLD and related disorders¹². FTLD pathology is mainly characterized by i) aberrant accumulation of protein aggregates and ii) dystrophic neurites (DN) in the affected area²¹. Kurz & Pernecky¹⁸ suggest that protein misfolding, cleavage and phosphorylation play a crucial role in determining the development of protein aggregates. Intracellular deposits of these proteins often define the major histopathological signatures of common neurodegenerative disorders associated with cognitive deterioration.

Different protein inclusions characterize the various FTLD subtypes, which can be subdivided into broad categories on the basis of the molecular defects that is most characteristic: the majority of FTLD cases shows almost equal distribution of microtubule-associated protein tau (MAPT) tauopathy (FTLD-tau) (~40% of cases) and

ubiquitin/TDP-43 (FTLD-TDP) proteinopathy (~50% of cases), while fused in sarcoma (FUS) positive (FTLD-FUS) and p62 (FTLD-UPS) pathology represent 10% and 1-2% of cases, respectively²⁰. Such protein inclusions mostly accumulate in the cytoplasm of neurons and glial cells located mainly in the frontal and temporal cortex, but also in the hippocampus and sub-cortical nuclei^{21,22}.

Tauopathies are a group of diseases which show tissue deposition of the abnormally aggregated protein tau¹². FTLD cases presenting tauopathies show as a predominant pathological feature hyperphosphorylated tau inclusions, which accumulate in the cytoplasm of neurons and glial cells located, mainly, in the frontal and temporal cortex, the hippocampus and subcortical nuclei, with a minor prevalence in midbrain, brainstem, cerebellum and spinal cord^{14,21}.

Tau is a microtubule-associated protein that stabilizes microtubule assembly by binding to tubulin. Tau has eight different isoforms that are generated by alternative splicing of exons 2, 3, and 10 from a single gene that in humans is designated *MAPT*, located on chromosome 17. Each tauopathy presents distinct morphological characteristics of neuronal and glial lesions, varies in the anatomic distribution of the lesions, and may differ by the molecular pathology.

Although a significant number of FTLD and related disorders (i.e., PiD, CBD and PSP) are characterized by the accumulation of the protein tau, the majority of FTLD cases are TDP-43 proteinopathies¹².

The TAR DNA-binding protein 43 (TDP-43) was identified in 2006 by researchers in Japan and USA as one of the mostly ubiquitinated proteins in FTLD, FTLD-MND and ALS. TDP-43 is a highly conserved nuclear protein that plays a key role in transcriptional regulation as well as in alternative splicing control¹². Recent studies¹² suggest that FTLD-TDP could be further sub-classified into three variants based on the morphology, distribution, and ratio of TDP-43-positive neuronal cytoplasmic inclusions to dystrophic neurites; these have been described as FTLD-TDP types I, II, and III and they have shown to be associated with the FTLD syndromes. FTLD-TDP-typical inclusions (tau-negative, ubiquitin-positive) are often seen in patients carrying mutations in the progranulin (*GRN*), valosin-containing protein (*VCP*) and Chromosome 9 Open Reading Frame 72 (*C9ORF72*) genes²³. Interestingly, a recent study by Jiang et al.²⁴ found that the pathological amyloid fibrils in FTLD-TDP might be composed of TMEM106B and not TDP-43, thus potentially refocusing future pathogenic studies of FTLD-TDP.

Alternatively, neuronal inclusions immunoreactive to ubiquitin are characteristic of the FTLD-FUS form. FUS is the last major subgroup of FTLD pathology, and is a proteinopathy associated with the co-aggregation of FET (FUS [fused in sarcoma], EWS [Ewing sarcoma] and TAF15 [TATA binding associated factor 15]) proteins into ubiquitinated inclusions found in sporadic FTLD cases¹¹. FUS proteinopathy has been suggested to be involved in the shared neuropathologic mechanisms underlying FTLD disorders and ALS¹¹. Lastly, an ultimate FTLD subtype characterized by ubiquitin-immunoreactive inclusions has been recognized: FTLD-UPS (frontotemporal lobar degeneration with inclusions positive for ubiquitin proteasome system markers) is commonly associated with a mutation in the charged multivesicular body protein 2B (*CHMP2B*) gene that causes disruption of endosomal traffic. A decline in the levels of CHMP2B results in a decrease of hippocampal dendritic branching and diminished excitatory synapse activity, thus significantly decreased synaptic plasticity and causing cognitive and neurological deficits¹¹. However, it is noteworthy to mention that *CHMP2B* mutations are relatively rare¹¹. A summary of the different genetic, molecular and clinical classifications of FTLD are listed in **Figure 1.2**.

	FTLD-tau	FTLD-TDP	FTLD-FUS	FTLD-UPS
Prevalence in FTLD population	~45%	~50-55%	~6-8%	<1%
Clinical presentation	bvFTD, PSP, CBS, PiD	bvFTD, FTLD-MND, SD, ALS	bvFTD, FTLD-MND, ALS	bvFTD, FTLD-MND
Associated genetic mutations	<i>MAPT</i>	<i>GRN, VCP, C9ORF72</i>	?	<i>CHMP2B</i>
Deposited proteins	tau	TDP-43/ TMEM106B?	FUS, EWS, TAF15	Ubiquitinated proteins

Figure 1.2. Different classifications of frontotemporal lobar degeneration.

1.1.3. Genetics

Neurodegenerative disorders are generally characterised by a complex genetic architecture, and often a minority of familial cases is outnumbered by a large population of sporadic cases⁹.

Mendelian genes have been classically isolated in familial studies through genetic linkage analysis and, more recently, whole-exome sequencing (WES) of large well-phenotyped pedigrees^{9,25}. On the other hand, sporadic forms of disease are generally investigated through case/control association studies, such as genome-wide association studies (GWASs). GWASs are able to identify genetic variants associated with complex human diseases and traits by assessing the significant differences in allele frequencies of common genetic markers across large cohorts of patients and population-matched controls⁹. While in the past decade GWASs have been able to highlight hundreds of risk variants associated with neurodegenerative disorders, most variants identified so far confer relatively small increments in risk, and explain only a small portion of familial clustering, leading many to question how to explain the remaining ‘missing’ heritability²⁶.

1.1.3.1. Familial FTLD-Mendelian genetics

Familial forms account for a minority of global FTLD cases, ranging from 30 to 40% overall.

Several post-mortem studies demonstrated that clinically and neuropathologically diagnosed FTLD can result from different underlying molecular pathologies (e.g., tau or TDP-43 pathology), indicating that multiple pathogenic pathways might result in converging and/or overlapping clinical phenotypes²⁷.

Corresponding with this complex pathological architecture of FTLD, genes in manifold pathways have recently been implicated in its molecular pathogenesis²⁷. However, the full spectrum of neurodegenerative disease genes and the relative proportions in which each contributes to the complex genetic architecture of FTLD have not yet been systematically explored.

Mutations in a handful of genes, that are usually referred to as Mendelian FTLD genes^{9,14,28,29}, are known to be classically associated with familial FTLD. The three most common ones are the microtubule-associated protein tau (*MAPT*), progranulin (*GRN*) and Chromosome 9 Open Reading Frame 72 (*C9ORF72*)^{10,11,30}.

MAPT encodes different isoforms of tau proteins, which are essential for cell shape maintenance, molecules trafficking and signal transduction²¹. *MAPT* mutations (≥ 44 pathogenic mutations⁹) are in fact involved in the formation of hyperphosphorylated tau in cortical and subcortical grey and white matter and they occur on average at a frequency of 6-11% of all FTLD subjects¹¹. According to Young et al.¹¹, they are inherited in an autosomal-dominant manner, they usually have high penetrance, and may have a shorter duration of illness relative to other mutations.

GRN (≥ 70 pathogenic mutations³¹) is involved in cell cycle progression, autophagy-lysosome pathway regulation and inflammation processes, and accounts for 5-20% of the cases in the FTLD population^{11,32}. Just like *MAPT*, *GRN* mutations are inherited in an autosomal-dominant manner, but interestingly they have relatively low penetrance until the age of 70¹¹.

Mutations in both genes have been associated with asymmetrical frontal, temporal and inferior parietal lobe atrophy and mainly with TDP-43 positive inclusions. The overall frequency of *MAPT* and *GRN* mutations in FTLD accounts for only 20% of the cases²¹.

The non-coding hexanucleotide repeat expansion in *C9ORF72* exhibits an autosomal-dominant pattern of inheritance and has been suggested to be the most common genetic cause of chromosome 9-linked FTLD and ALS^{11,33,34} as well as, although with lower prevalence, in AD, PD, ataxia and CBS cases^{19,21,35-43}.

Furthermore, a number of additional genes explain a minority of familial FTLD cases. Truncation mutations in the charged multivesicular body protein 2B (*CHMP2B*) gene are extremely rare but have been linked to a large Danish family and a Belgian patient presenting with FTLD^{44,45}.

FTLD-ALS pathology has also been associated with very rare mutations in other genes including sequestosome 1 (*SQSTM1*)⁴⁶, ubiquilin 2 (*UBQLN2*)⁴⁷, optineurin (*OPTN*)⁴⁸, coiled-coil-helix-coiled-coil-helix domain containing 10 (*CHCHD10*)⁴⁹, TANK binding kinase 1 (*TBK1*)^{48,50}, intraflagellar transport 74 (*IFT74*)⁵¹, dynactin-1 (*DCTN1*)⁵² and, although still awaiting to be confirmed by further studies, the TIA1 cytotoxic granule-associated RNA-binding protein (*TIA1*)⁵³. To date, none of these mutations have been associated with a unique pathological signature²⁰. Families carrying mutations in these genes have been found to present with variable FTLD and/or ALS features within the same pedigree, suggesting that mutations in these genes may predispose to both

diseases^{54,55}.

Mutations in a number of genes have been shown to co-occur with various phenotypes: mutations in the valosin-containing protein (*VCP*) have been identified in few cases carrying a combination of conditions such as inclusion body myopathy (IBM) with Paget disease of the bone (PDB) and FTD, and in FTLN-ALS cases, while mutations in *SQSTM1* have recently been described in families whose affected members presented with ataxia, dysarthria, dystonia, vertical gaze palsy and cognitive decline^{56,57}. Finally, although pathogenic genetic variability in the transactive response DNA-binding protein (*TARDBP*) and the fused in sarcoma (*FUS*) genes is usually associated with clinically pure ALS, there have been very few reports of cases of clinical FTD (\pm ALS) harbouring mutations in these genes^{20,58-60}. This is surprising in light of the fact that the protein products of both *TARDBP* and *FUS* have been shown to be among the main pathological hallmarks contributing to FTLN pathology (FTLN-TDP and FTLN-FUS respectively)²⁰.

To summarise, Mendelian FTLN genes can be classified on the basis of their disease specificity: *MAPT*, *GRN* and *CHMP2B* can be considered the main 'pure' FTLN genes, as mutations in these genes are the most common and they have been mainly or exclusively identified in FTLN cases⁶¹; *C9ORF72*, *VCP*, *TARDBP*, *FUS*, *SQSTM1*, *UBQLN2*, *IFT74*, *OPTN*, *CHCHD10*, *TBK1*, and *TIA1* seem to encompass ALS and/or some heterogeneous array of extra-phenotypic features, as a result they can be considered as 'spectrum-FTLN genes'⁹. A summary of the major genes associated with the FTLN-ALS continuum with their discovery timeline is presented in **Figure 1.3**.

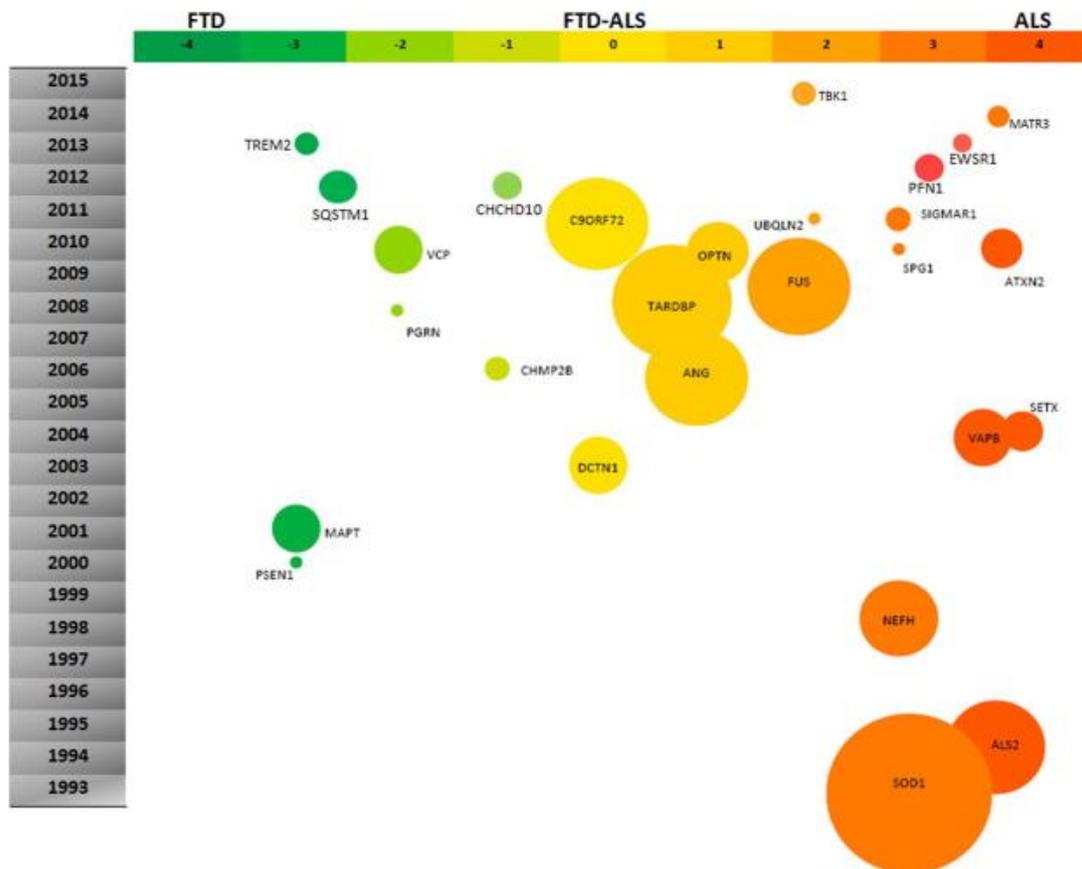


Figure 1.3. Landscape of the genes associated with the FTLD-ALS spectrum. It has now been established that ALS and FTD are linked clinically, pathologically, and mechanistically, and the diseases are now properly recognized as representatives of a continuum of a broad neurodegenerative disorder, with each presenting in a spectrum of overlapping clinical symptoms. It has been estimated that approximately 15% of FTD patients meet ALS criteria and perhaps as much as 15% of ALS affected individuals can also meet FTD criteria. ALS-FTLD genes are plotted to show clinical phenotype (x-axis), year of discovery (y-axis) and level of research on each gene (balloons' size). Adapted from 'Delineating the relationship between amyotrophic lateral sclerosis and frontotemporal dementia: Sequence and structure-based predictions' by V. Kumar, 2016, *Biochimica et Biophysica Acta*, Volume 1862, Issue 9, Pages 1742-1754. Copyright 2022 by Elsevier Inc. Adapted with permission.

1.1.3.2. Sporadic FTLD

Although sporadic forms of FTLD account for the majority (60-70%) of global FTLD cases⁹, the genetics of idiopathic FTLD is still relatively unclear. Besides the reports of a few *MAPT*, *GRN* and *C9ORF72* mutations^{62,63}, the genetic architecture of sporadic FTLD primarily refers to genetic risk markers with small effect size likely modulated by multiple genetic and/or environmental modifiers, as shown in **Figure 1.4**^{9,26,62}.

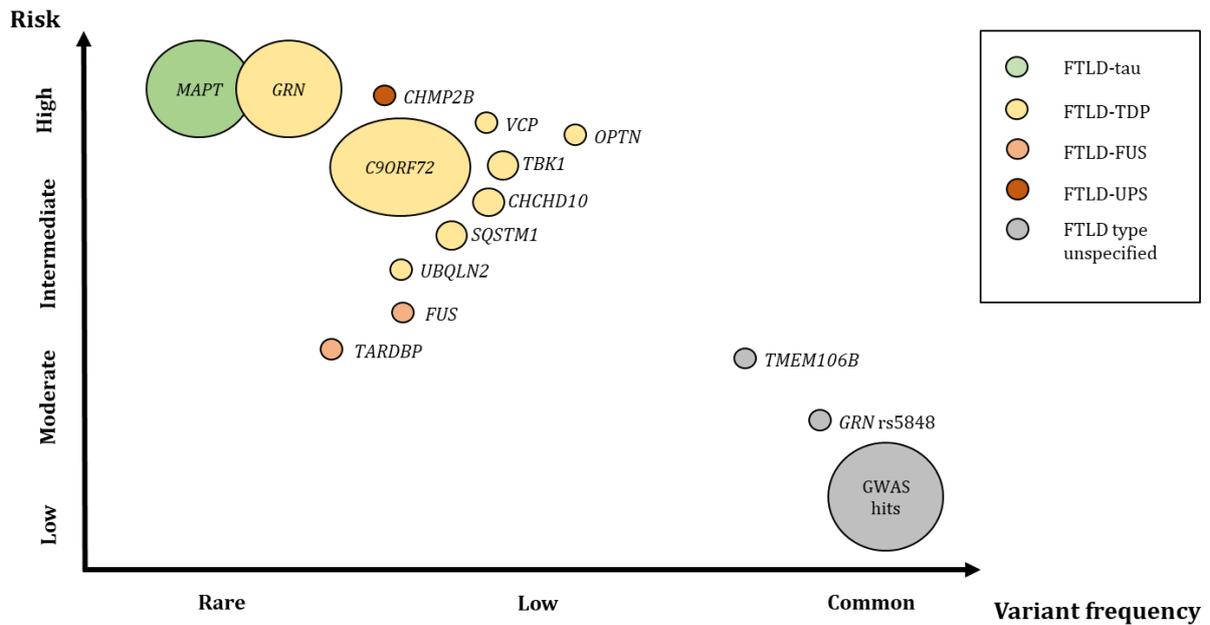


Figure 1.4. Schematic overview of common familial and rare FTLD genes and genetic risk factors. The risk attributed to a variant in a gene associated to FTLD (y-axis) is represented in function of its variant frequencies in the general population (x-axis). The sizes of the circles correspond approximately to the proportion of FTLD patients harbouring a variant in the corresponding gene. Each gene has been associated with certain types of pathology: FTLD-tau, FTLD-TDP, FTLD-FUS and FTLD-UPS, whereas in some instances the type of pathology is unspecified.

To date, the most informative studies exploring genetic risk and/or modifying factors in sporadic FTLD have been GWASs. These studies were designed to either address actual sporadic cohorts, as in the case of the International clinical FTLD GWAS⁶⁴, the Italian clinical FTLD-GWAS⁶⁵ and the Dutch clinical FTLD-GWAS⁶⁶, or more selective cohorts presenting with a specific pathology (i.e., TDP-43 pathology⁶⁷) or a specific genetic variant (i.e., *GRN* mutation carriers⁶⁸).

Van Deerlin et al.⁶⁷ published a GWAS study in 2010 that identified a number of single nucleotide polymorphisms (SNPs) encompassing the transmembrane protein 106B (*TMEM106B*) gene as risk factors for the FTLD-TDP subtype and observed that their risk alleles appeared to be particularly enriched in *GRN* mutation carriers, thus suggesting *TMEM106B* as a *GRN* modifier. Of note, *TMEM106B* association with FTLD has been confirmed in subsequent studies^{67,69–71} and functional analyses suggested that *TMEM106B* is involved in endolysosomal pathways and modulates PGRN protein levels⁷².

In 2014, Ferrari et al.⁶⁴ published a clinical international FTLD GWAS study. The analysis revealed two novel susceptibility loci for clinical FTLD: i) the *RAB38 – CTSC* locus for the bvFTD subtype, which was suggested to be involved in the lysosome-phagosome

pathways; ii) the *BTNL2* and *HLA-DRA – DRB5* locus in global FTLD, which was shown to regulate immune system processes^{9,64}.

A second clinical FTLD-GWAS was published only a year later by Ferrari et al.⁶⁵, this time involving a multicentre Italian FTLD cohort. The study defined two suggestive (i.e., close to significance, yet not genome-wide significant) loci encompassing the *CEP131* and *ENTHD2* genes suggesting that neuronal development, differentiation and maturation processes might be the biological mechanisms underlying the association and thus driving FTLD pathogenesis in the Italian population^{64,65}.

Another FTLD-GWAS was published in 2018 by Pottier et al.⁶⁸, who performed the analysis on a FTLD cohort selected for the specific feature of carrying loss of function *GRN* mutations. Interestingly, the results seem to replicate the findings by Van Deerlin et al.⁶⁷: genome-wide significance was reached by a lead SNP encompassing *TMEM106B*, reinforcing the hypothesis that transcriptional regulation of such gene might underlie FTLD pathology by modulating PGRN protein levels. In addition, another outstanding marker regulating *GFRA2* expression was identified and it was shown that *GFRA2* and PGRN interact at the cellular level and seem to be both part of the glial cell line-derived neurotrophic factor (GDNF) pathway, which is known to promote neuronal survival⁷³. Altogether, this study confirms *TMEM106B* and brings forward *GFRA2* as novel potential modifiers in *GRN* mutation carriers.

The most recent FTLD-GWAS was published in 2021 by Reus et al.⁶⁶, who collected a multicentre Dutch FTLD cohort. The study identified two risk SNPs tagging a *C9ORF72* haplotype which is carried by ~4% of the population and has been shown by several studies to greatly increase the risk for a pathological *C9ORF72* repeat expansion, which has been associated with the FTLD-ALS continuum^{34,74,75}. These findings imply that these novel variants increase susceptibility to *C9ORF72* pathological repeat expansions.

Besides the typical GWASs, more recently evidence originating from a study⁷⁶ investigating epigenetics combined with GWAS data suggested that the *HLA* locus might play a key role in regulating the expression of proinflammatory players in the brain cortex, impacting FTLD patients' age at onset (AAO). This study is one of many that further support that multiple risk factors and/or modifiers might in fact significantly contribute to disease endophenotypes or disease-specific features⁹.

Overall, the associations reported in the different FTLD-GWASs hardly replicated across each other, with the exception of the *TMEM106B* locus which is shown in two

independent studies (the FTLN-TDP and FTLN-GRN GWAS). Given the specific characteristics of the cohorts analysed, one featuring genetic variability in the *GRN* gene and the other TDP-43 pathology, it would appear that *TMEM106B* is probably an exclusive modifier of *GRN* mutation carriers and cases presenting TDP-43 pathology. The clinical GWASs indicated different loci and were not cross-supportive, suggesting that differences in population substructure might be determining such genetic associations, and therefore underlie this lack of replication together with sample size and statistical power.

1.1.4. Molecular mechanisms of FTLN pathogenesis

Impairment in several pathways has been associated with FTLN pathology however translating genetic knowledge into functional understanding of impacted biological processes in complex disorders still represents a major challenge⁷⁷. To date, the study and functional characterisation of the protein products of Mendelian, sporadic and GWAS genes has pointed to a number of susceptibility processes that seem to be conserved across familial and sporadic forms of disease and thus have been informative on the potentially associated impacted biological processes. Among others, *waste disposal* and *immune signalling* have been commonly prioritised, while *DNA damage response* is emerging as a novel intriguing candidate pathway⁹ (**Figure 1.5**).

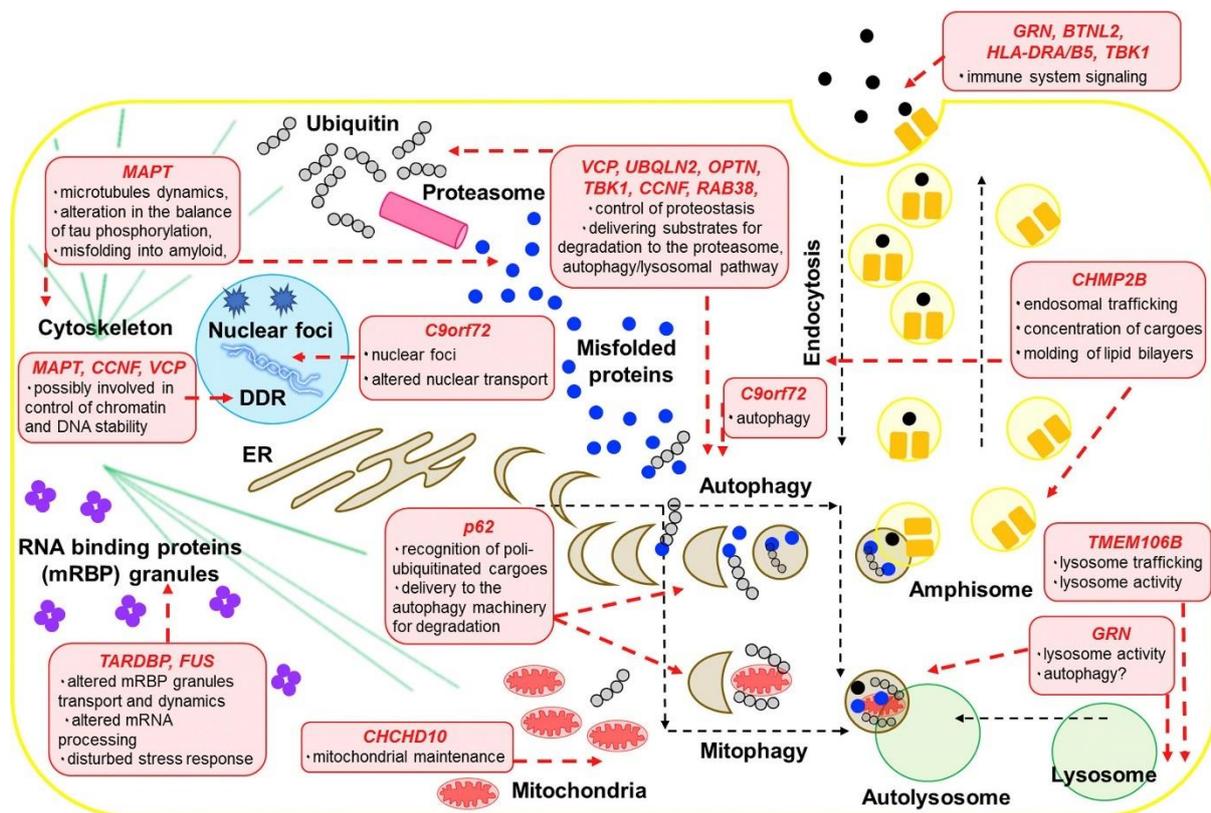


Figure 1.5. Summary of all putative impacted molecular processes contributing to FTLD pathogenesis on the basis of the current global genetic characterization of FTLD including Mendelian, sporadic and GWAS FTLD genes. DDR, DNA damage response; ER, endoplasmic reticulum. Adapted from 'Genetics and molecular mechanisms of frontotemporal lobar degeneration: an update and future avenues' by R. Ferrari, 2019, *Neurobiology of Aging*, Volume 78, Pages 98-110. Copyright 2022 Elsevier, Inc. Adapted with permission.

Several genes, namely *CHMP2B*, *C9ORF72*, *GRN*, *VCP*, *UBQLN2*, *OPTN*, *SQSTM1*, *TBK1*, *CCNF*, *TMEM106* and *RAB38* have been shown to underpin cellular *waste disposal* (i.e., the process of breaking down damaged cellular components and debris), which encompasses a number of specialised subprocesses such as the endolysosomal pathway, macroautophagy and the unfolded protein response (UPR)^{9,78-81}. Although alterations of the *waste disposal* process have been often associated to various NDs, the underlying detailed molecular mechanisms are still unknown⁷⁸. However, it is widely accepted that one of the major causes for neuronal damage and death is the toxic accumulation of misfolded proteins within the subcellular environment driven by dysfunctional *waste disposal*, due to mutations in one or more of these genes. Briefly, *GRN* encodes for a glycoprotein product (progranulin [PGRN]) that is taken up by the cell from the extracellular space and subsequently cleaved into seven units of granulins (GRNs) within the endolysosomal pathway⁸²⁻⁸⁴. The function of PGRN and GRNs is still not completely

clear, but several links have been made with modulation of inflammatory response and lysosomal homeostasis^{82,85}. The protein product of *CHMP2B*, another FTLD Mendelian gene, has been shown to be involved in endosomal trafficking while *SQSTM1*, whose protein product is known as p62, plays a key role in recognising polyubiquitinated cargoes for delivery to the autophagy machinery and subsequent degradation^{23,86}. Both VCP and UBQLN2 are responsible for ubiquitin-mediated proteostasis control, mostly delivering substrates tagged for degradation to the proteasome^{10,14}. OPTN and TBK1 act in synergy to recognise protein aggregates in a ubiquitin-independent fashion and deliver them for autophagy – lysosome-mediated degradation⁸⁷⁻⁸⁹. CCNF mediates proteasomal targeting of specific substrates during the G2 phase of the cell cycle⁹. Interestingly, only more recently C9ORF72 has been linked to autophagy and proteostasis, while most studies have indicated that pathological expansions might reduce mRNA expression or generate toxic nuclear foci⁶². CHCHD10 has been associated with maintenance of mitochondrial quality control and mitophagy^{49,90}. Finally, the transmembrane protein TMEM106B is involved in the maintenance of endolysosomal trafficking, while RAB38 regulates vesicle trafficking^{9,82,91}.

GRN, *TBK1*, *BTNL2* and *HLA-DRA* genes appear to support signalling pathways involved in *immune response*⁹. Interestingly, two genes strongly implicated in *waste disposal*-associated processes have also showed to be involved in immune system signalling: in fact, upregulation of PGRN has been shown by several studies to have an anti-inflammatory effect and to regulate innate immunity gene expression in microglial cells, while TBK1 plays several roles in the innate immune system response, suggesting a potentially ubiquitous role for both genes in FTLD pathogenesis that warrants additional focused studies²³. The ubiquitously expressed membrane protein BTLN2 is involved in repressing T-cell proliferation, while the transmembrane receptor HLA-DRA impacts modulation and regulation of immune responses in the brain, especially in microglia^{92,93}.

Despite TDP-43 and FUS being major pathological hallmarks of FTLD, their associated genes (i.e., *TARDBP* and *FUS* respectively) display extremely rare genetic variability. Both appear to be functionally involved in *RNA metabolism* and *regulation of gene expression* as well as in the regulation of mRBP (mRNA-binding proteins) granules (i.e., stress granules) relocalisation within the cell, which associated to neuronal health^{94,95}.

Other processes that might be impacted on the basis of genetic variability affecting both Mendelian and GWAS genes are *neuronal physiology* and *development*. Mutations in the *MAPT* gene lead to alterations in tau protein splicing, which results in: i) increased free cytosolic population of dysfunctional tau isoforms, which are very prone to aggregation; ii) unstable microtubules, which are essential structural cellular components that rely on tau-binding for stabilisation⁹⁶. Furthermore, the recent GWAS-associated locus *GFRA2* supports additional processes related to neuronal development and protection⁶⁸.

Finally, recent bioinformatics work raised the importance of *DNA damage response (DDR)* as a key molecular mechanism underlying FTLN pathology⁹⁷. The *DDR* is the processes through which DNA integrity is maintained, through mechanisms of damage recognition, repair and tolerance⁹⁸. Several studies have shown that mutations in *MAPT* result in alteration of the cellular cycle, chromatin damage and impaired DNA repair, which could possibly be linked to brain cell death and neurodegeneration observed in FTLN⁹. Interestingly, in addition to its role in the *waste disposal*-related processes (see section **1.1.4.2**), *CCNF* appears to control genome stability through cell-cycle check points⁹⁸. Similarly *VCP*, which is known for its crucial role in proteostasis, has been shown to be part of a complex of proteins recruited to the DNA double-strand breaks for repair⁸³.

Clearly, these avenues need to be further explored, yet it is relevant to note that many Mendelian, FTLN-ALS and GWAS FTLN genes point toward different yet convergent components of the broader *waste disposal* pathway, indicating that both dominant mutations (high penetrance) as well as common markers with small effect size (low penetrance) support a common process involved in disease pathogenesis. Formation and deposition of insoluble protein deposits is another common principle not only for FTLN but also for most NDs, where dysfunctional *protein quality control* and generally *waste disposal* mechanisms have been shown to underpin disease pathology (e.g., PD), and thus reinforcing their relevance in disease pathology.

1.1.4.1. Protein quality control systems

Protein homeostasis, also known as proteostasis, refers to maintenance of a stable and functional proteome to allow a myriad of cellular functions⁹⁹. Biogenesis, folding, trafficking and degradation are precisely coordinated and controlled so that proteins can

maintain their specific three-dimensional structural conformation, abundance and subcellular localisation within the cell¹⁰⁰. In particular, protein folding represents a key component of proteostasis as it determines protein final conformation and subsequent functionality. As a result protein folding is tightly regulated by a complex network of molecular chaperones that are at constant surveillance of proteins and assist in folding efficiently^{99,100}. Chaperones aim to prevent protein aggregation during the folding process by recognising exposed hydrophobic amino acid patches of unfolded or partially folded proteins that should normally be buried within their interior. When misfolded proteins cannot be refolded to their native state, some chaperones can act to target them for degradation or sequestration from the cellular environment, thus protecting the rest of the proteome from faulty interactions⁹⁹. Stably folded proteins can also unfold and aggregate under stress conditions, such as reactive oxygen species (ROS)-associated stress or elevated temperatures¹⁰¹. In any of these instances, misfolded proteins often acquire aberrant or even toxic functions, which frequently result in pathogenic phenotypes such as those of age-related neurodegenerative diseases including AD, PD, ALS, HD and FTLD^{102,103}.

In order to respond to imbalances in proteostasis and adjust their proteome status in response to metabolic and environmental changes, cells have a wide range of protein quality control and degradation strategies in place. These strategies have been identified as multiple distinct pathways including: i) molecular chaperones; ii) the ubiquitin-proteasome system (UPS); iii) ER and Golgi apparatus; iv) ER-associated degradation (ERAD); v) stress granules (SGs) and vi) autophagy (**Figure 1.6**)¹⁰⁰.

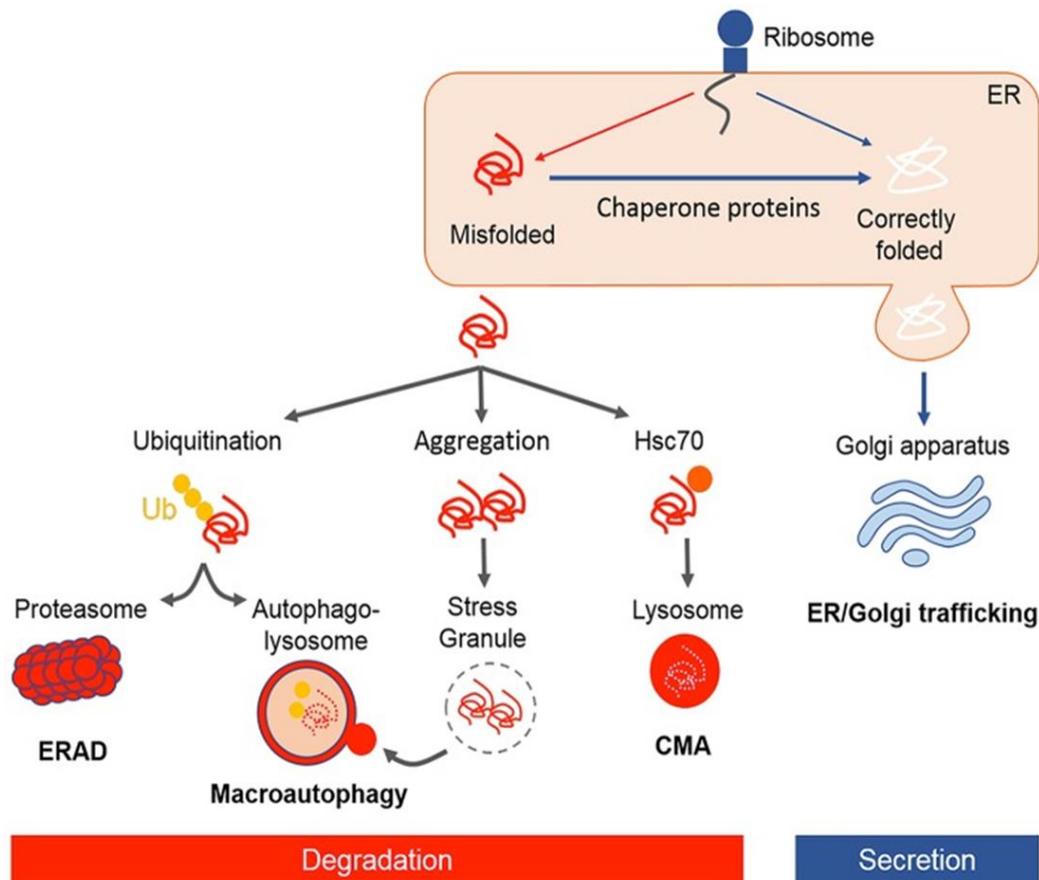


Figure 1.6. Illustrative representation of protein quality control mechanisms in the cell. Following translation, newly synthesized nascent polypeptides are constantly at risk of misfolding and aggregation. Chaperones facilitate folding of proteins or refolding misfolded proteins. Approximately one-third of newly folded proteins transit through the endoplasmic reticulum (ER)–Golgi pathway for post-translational modification and secretion. Proteins which are not correctly folded are recognized by ER-associated degradation (ERAD), targeted for ubiquitin–proteasome degradation, autophagy, or a smaller proportion are degraded by chaperone mediated autophagy (CMA). In case of protein aggregation, stress granules (SGs) form transiently and are cleared through macroautophagy. Adapted from ‘Protein Quality Control and the Amyotrophic Lateral Sclerosis/Frontotemporal Dementia Continuum’ by H. Shahheydari, 2017, *Front. Mol. Neurosci.* Copyright 2017 Shahheydari, Ragagnin, Walker, Toth, Vidal, Jagaraj, Perri, Konopka, Sultana and Atkin. Adapted with permission.

Molecular chaperones can be classified into several families based on their specific biological function and localisation within the cell. Among the most common ones is the heat shock proteins (Hsps) family, which protects intracellular proteins from denaturing stress conditions, apoptosis and inflammatory damage, especially during hyperthermic, hypoxic and oxidative stress conditions^{103–105}. Members of the Hsp family (e.g., Hsp70 and Hsp90) provide a natural protective system for the cell against excessive damage by preventing additional protein aggregation. By contrast other chaperones, such as the protein disulfide isomerases, can facilitate protein folding by catalysing disulfide bond formation between different protein domains^{100,106,107}.

The ubiquitin proteasome system (UPS) is the primary route for degradation of short-lived proteins and always involves a three-step enzymatic cascade of reactions that activate (E1 ubiquitin-activating enzyme), conjugate (E2 ubiquitin-conjugating enzyme) and ligate (E3 ubiquitin-ligating enzyme) a polyubiquitin chain to the target substrate^{100,105}. In turn, the polyubiquitin chain signal targets the substrate to the proteasome (i.e., a highly sophisticated protease complex designed to carry out hydrolysis of selected client proteins¹⁰⁵), followed by proteolysis by the 26S proteasome subunit. During degradation, ubiquitin moieties are removed and recycled while the target protein is cleaved into small peptides while passing through the proteasome.

The ER – Golgi compartments play a central role in cellular protein quality control, as they are the sites where post-translational modifications of newly synthesised proteins take place, leading to protein activation and maturation. While most proteins are transferred to Golgi and then exocytosed when properly folded and functional, misfolded/unfolded proteins accumulate in the ER saturating its folding capacity. This perturbation of ER homeostasis activates the unfolding protein response (UPR), which attempts to relieve ER stress by reducing protein synthesis and increasing translation of ER chaperones, and thus augmenting its protein folding capacity. When these measures are not enough, defective proteins are transported back to the cytosol, degraded via the ERAD pathway and eventually by the UPS¹⁰⁰.

The endoplasmic reticulum-associated degradation (ERAD) pathway is a critical stress response mechanism that gets activated when misfolded proteins accumulate within the ER obstructing its crucial homeostatic functions. In this scenario, the ER recognises, selects and retrotranslocates misfolded proteins into the cytosol and directs them towards UPS-mediated degradation¹⁰⁰.

Stress granules (SGs) are dynamic, membrane-less organelles that can rapidly and reversibly assemble and disassemble, assuring a quick response to stress conditions within the cell. SGs accumulate under stress conditions and are able to alleviate translational burden on cells by sequestering mRNA molecules, RNA-binding proteins, translation initiation factors and small ribosomal subunits^{53,100}. Once they re-establish cellular homeostasis by regulating protein expression, SGs are cleared through autophagy.

Autophagy is a lysosomal degradation process used to recycle obsolete cytoplasmic content, such as damaged organelles and proteins, that encompasses three different subprocesses, namely macroautophagy, chaperone mediated autophagy (CMA) and microautophagy¹⁰⁰ (**Figure 1.7**).

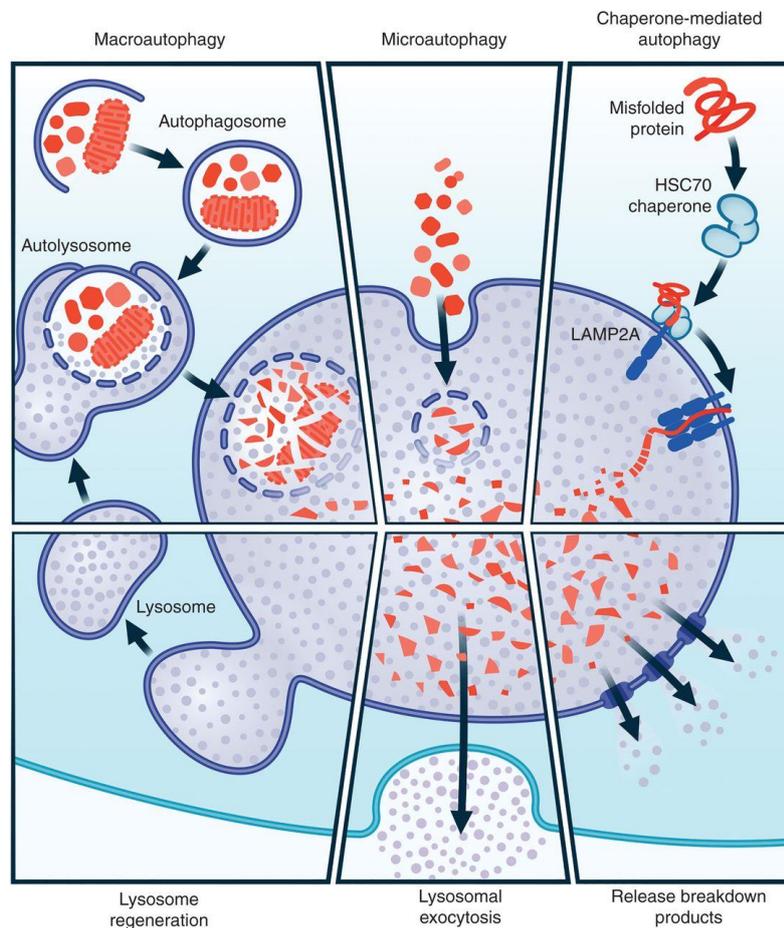


Figure 1.7. The autophagy lysosomal pathway. The lysosome is the hub of a network of pathways that feed cargo into its lumen for degradation. These include pathways depicted here to deliver intracellular cargo such as macroautophagy, chaperone-mediated autophagy and microautophagy, as well as others that deliver extracellular cargo to the lysosome including endocytosis and micropinocytosis. Cargo delivered to the lysosome can undergo degradation into molecular building blocks that can return to the cytoplasm to be catabolized further to supply cellular energy needs or to be reused in the synthesis of new macromolecules. Contents of lysosomes can also be extruded extracellularly by a Ca²⁺-dependent exocytic process. LAMP2A, Lysosome-associated membrane glycoprotein 2; HSC70, heat shock cognate 71. Adapted from ‘The Autophagy Lysosomal Pathway and Neurodegeneration’ by S. Finkbeiner, 2020, *Cold Spring Harb. Perspect. Biol.* Copyright 2020 Cold Spring Harbor Laboratory Press. Adapted with permission.

Macroautophagy, the most common form of autophagy, involves a multi-step process⁷⁹. Briefly, it begins with the *de novo* formation of an isolation membrane, named phagophore, following inhibition of the master regulator of macroautophagy, mTOR, and

the resulting activation of its downstream target, the Unc-51-like kinase (ULK) complex (ULK1, ULK2, Atg13, FIP200 and Atg101)¹⁰⁸. Activation of the ULK complex induces elongation of the phagophore around a region of the cytoplasm encompassing the misfolded protein/damaged organelle and closure of the membrane lipid bilayer to form a double membrane autophagosome. Cargo engulfment by the autophagosome is mediated by p62/SQSTM1, which interacts with the adaptor protein light chain 3 II (LC3-II, an active form of LC3), on the surface of the autophagosome. Finally, the autophagosome fuses with the lysosome, a membrane-bound organelle characterised by an acidified milieu (pH~4.5), resulting in the formation of the autophagolysosome, which leads to degradation of its content by lysosomal hydrolases^{109,110}.

CMA is another form of autophagy that targets soluble misfolded proteins directly to the lumen of the lysosome for digestion. Protein substrates of the CMA always carry a polypeptide motif (i.e., KFERQ) on their surface that is recognised by the molecular chaperone Hsc70, which translocates them to the lysosome¹¹¹. Interestingly, the KFERQ motif is found in ~30% of the proteome including several proteins involved in NDs, such as α -synuclein, suggesting that CMA might be critical to accelerate the clearance of some disease-relevant proteins and presumably reduce their propensity to accumulate and aggregate⁷⁹.

The third form of autophagy is microautophagy, which involves the recruitment of cytosolic components in proximity with the lysosome and their invagination by the lysosomal membrane^{79,112}.

Recent research in the field has brought to light the dual nature of the autophagy pathway: both macro- and microautophagy can be either selective, meaning substrate-specific, or non-selective^{79,100}. While non-selective autophagy targets any type of cargo for degradation, selective autophagy is characterised by the removal of specific organelles or cellular components; in this scenario, several subtypes exist: among others, aggrephagy for instance refers to the removal of insoluble protein aggregates; mitophagy is dedicated to the selective removal of damaged mitochondria (refer to section **1.1.4.3**).

Finally, additional pathway to the lysosome, including micropinocytosis, phagocytosis and endocytosis can deliver extracellular material to the lysosome for degradation. These pathways have been reported in the context of neurodegenerative disease as important

routes by which extracellular aggregation-prone proteins enter neurons and propagate proteinopathy between neighbouring cells⁷⁹.

1.1.4.2. Autophagy and proteostasis in neurodegenerative diseases

The abnormal intraneuronal accumulation of misfolded aggregated proteins and/or faulty intracellular components represent a common thread cutting across multiple neurodegenerative disorders, including AD, PD, HT, ALS and FTLD^{78,79}. Given the critical role of macroautophagy in the clearance of aggregated proteins and organelles, it stands to reason that autophagic dysfunction might be a common mechanism underlying several NDs pathologies. In this scenario, the autophagy pathway represents an intuitive therapeutic target, but most importantly it also presents with a remarkable genetic association with many NDs, where mutations in several different genes result in alterations at different steps of the autophagic process, from early steps of autophagosome formation through autolysosome formation^{45,113-115} (**Figure 1.8**).

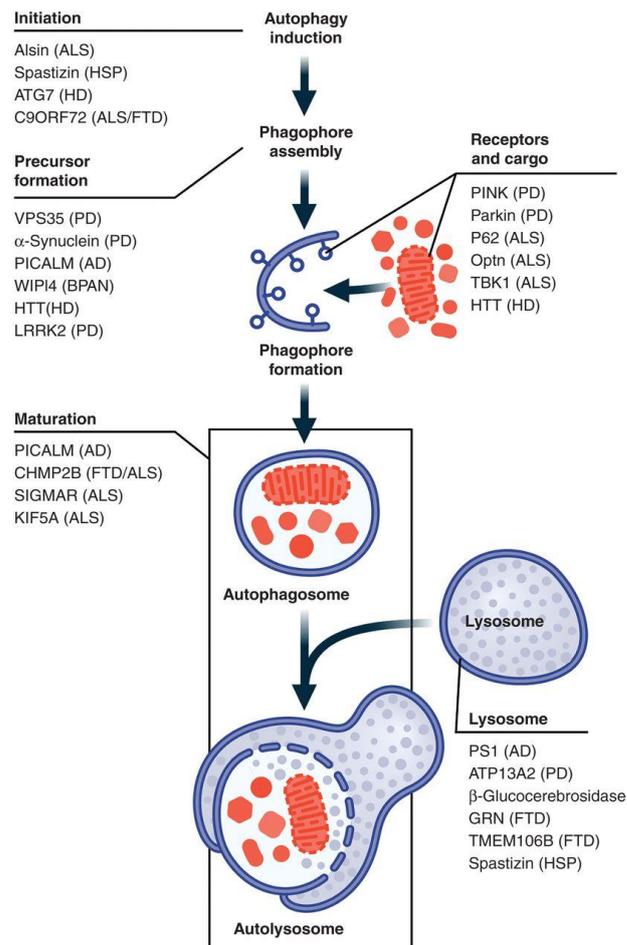


Figure 1.8. Genetic links between autophagy and neurodegenerative disease. Mutations in many genes that encode proteins that play a role in the autophagy lysosomal pathway (ALP) lead to neurodegenerative disease syndromes in humans, indicating that impairment of the ALP can be sufficient to produce neurodegenerative disease. Some examples are shown in this figure. AD, Alzheimer's disease; ALS/FTD, amyotrophic lateral sclerosis/frontotemporal dementia; BPAN, β -propeller protein-associated neurodegeneration; CMT2, Charcot-Marie-Tooth disease 2; HD, Huntington's disease; HSP, hereditary spastic paraplegia; PD, Parkinson's disease. An increasing number of genes associated with neurodegenerative diseases have now been implicated in autophagy function. These genes act at a number of different steps throughout the autophagic process, from early steps of autophagosome formation through autolysosome formation. Their proposed sites of action are indicated, along with the neurodegenerative disease with which they are associated. Adapted from 'The Autophagy Lysosomal Pathway and Neurodegeneration' by S. Finkbeiner, 2020, *Cold Spring Harb. Perspect. Biol.* Copyright 2020 Cold Spring Harbor Laboratory Press. Adapted with permission.

Autophagy is a complex pathway that presents with multiple steps and modes of regulation, which makes identification of potentially minor perturbations within the pathway very difficult⁷⁸. The mechanisms by which disease-associated genetic alteration in the autophagy pathway cause neurodegeneration have not been fully unravelled yet and are currently under investigation. Several studies have shown that mutations in these genes are often sufficient to cause NDs, either by disrupting autophagic clearance or by causing overproduction of misfolded proteins⁷⁹. It is important to consider that the

mismatch between the production and autophagic clearance of misfolded proteins might be the key to interpret disease phenotypes, meaning that many forms of disease may be caused by an overproduction of misfolded proteins and/or by reduced clearance⁷⁹. By contrast, compelling evidence supporting a role for dysfunctional autophagy as a causative factor in neurodegenerative disease has been produced from studies of mitophagy in PD: the importance of mitophagy in metabolising damaged mitochondria, and its associated genetics are now well-established, further underscoring that deficits in autophagy, and mitophagy specifically, may represent the most important risk factors for NDs^{114,116,117}.

Notably, in the past two decades the genetics of autophagy were characterised in non-neuronal models, raising the question on whether autophagy might differ in neurons¹¹⁸. Due to their unique longilineal morphology and post-mitotic nature, neurons must manage the degradation of cargo at distant sites (i.e., autophagosomes are biosynthesised at axon terminals and then trafficked to the cell body) and they have lost one important mechanism for clearance of long-lived proteins, namely cell division^{119,120}. Furthermore, several studies have showed that other brain cell-types (i.e., astrocytes) could contribute to induce starvation-dependent activation of mTOR-dependent autophagy, adding further layers of complexity to the study and characterisation of neuronal autophagy and its association with neurodegenerative disease^{121,122}.

More broadly, it is important to consider that a link between autophagy and ageing has been demonstrated: to date, ageing represents the major risk factor for NDs and has been shown to be often associated to significant downregulation of autophagy in the brain¹²³⁻¹²⁵. Conversely, several studies showed that autophagy induction and upregulation increase longevity and quality of life, conferring a substantial improvement in the quality of function of brain tissue and thus further supporting the autophagy pathway as an intriguing therapeutic target in NDs^{123,124}. Additionally, given the late-onset nature of the majority of NDs, it is possible that small alterations in proteins turnover could have cumulative effects that manifest later in life, as suggested by Menzies et al.⁷⁸, further reinforcing the contribution of autophagy dysfunction and ageing to disease progression.

To summarise, there is increasing evidence for the physiological importance of autophagy in neuronal health, raising the possibility that autophagy dysfunction may play a role in neurodegenerative diseases. The critical role of proteostasis and autophagic

clearance in maintaining a functional proteome and in disposing of aggregated proteins and other damaged organelles respectively, implies that autophagy dysfunction is likely a common mechanism in NDs and warrants further investigation for the design of targeted therapeutic strategies^{79,108,116,126}. Additionally, mitophagy dysfunction specifically has been shown to have an established role as a major pathological mechanism in NDs: deficits in bioenergetics, excessive ROS and cytoplasmic calcium, dysregulations of cellular redox balances are only few of the conditions associated to mitochondrial damage and accumulation¹²⁷.

1.1.4.3. Mitochondria and mitochondrial quality control mechanisms

Aside from the degradation of cytosolic aggregates, the autophagic pathway is also involved in the turnover of entire organelles such as mitochondria. Mitochondria are double-membraned organelles, with an outer and inner mitochondrial membrane (OMM and IMM respectively) separated by an intermembrane space, and a central matrix enclosed by the IMM¹²⁷. Long regarded as individual, 'bean-shaped' organelles, they are now understood as dynamic interconnected networks (i.e., mitochondrial network) linked to other organelles and shifting from innumerable punctate organelles to cell-wide tubular networks to regulate energy production and generally mitochondrial functions¹²⁸. While the shape of the mitochondrial network is governed by a complex fission/fusion machinery, its size is regulated by *de novo* mitochondrial biogenesis and macroautophagy^{127,128}.

Mitochondria are highly multifunctional organelles also known as the 'powerhouses' of the cell, as they are the cellular site of oxidative phosphorylation (OXPHOS), thus generating the bulk of adenosine triphosphate (ATP) required for cell activities¹²⁹. Besides cellular bioenergetics, they are responsible for several other biosynthetic reactions, including regulation of ROS levels and calcium homeostasis, and biosynthesis of macromolecules including lipids, amino acids and nucleotides^{127,128,130}. These essential processes generate reactive intermediates and oxidising agents as by-products, which in turn damage mitochondrial proteins and lipids^{128,131,132}. In order to maintain mitochondrial integrity and mitigate mitochondrial damage, cells have evolved a number

of elaborate mitochondrial quality control mechanisms. These cellular defence mechanisms against mitochondrial damage are organised in several lines of defence as follows: i) mitochondrial-associated degradation (MAD); ii) mitochondrial dynamics and quality control (QC); iii) the ubiquitin-proteasome system (UPS); iv) mitochondria-derived vesicles (MDVs); v) mitochondrial unfolded protein response (mtUPR); vi) mitophagy^{87,128,131–133} (**Figure 1.9**).

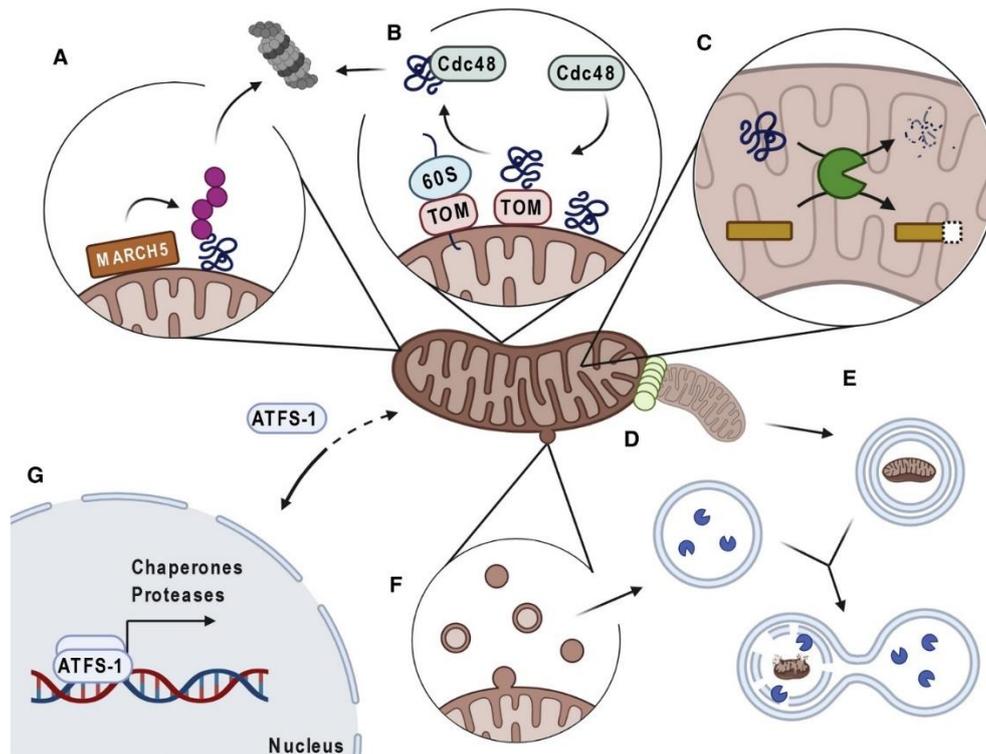


Figure 1.9. Major pathways of mitochondrial quality control. A) The UPS: mitochondrial proteins are conjugated with polyubiquitin chains (in purple) by E3 ubiquitin ligases like MARCH5 and are targeted for degradation by the proteasome; B) MAD involves the extraction, refolding and reinsertion of misfolded precursor proteins, such as in the case of cytosolic ribosomes (60S) localised on the OMM and of mitochondrial translocases with/without the assistance of AAA+ ATPases (e.g., Cdc48); if unsuccessful, such proteins are targeted for proteasomal degradation; C) Mitochondrial proteases carry out both quality control and regulatory functions; D) Mitochondrial morphology is regulated by fission and fusion events (green circles); E) Mitophagy: mitochondrial fragments destined for degradation are first engulfed in autophagosomes that eventually fuse with lysosomes; F) MDVs are small vesicles formed from the OMM, and at times IMM, and contain select mitochondrial proteins that are destined for peroxisomes or lysosomes; G) The mtUPR induces chaperone and protease gene expression in response to mitochondrial stress through activation of transcription factors such as ATFS-1. Adapted from 'Quality control of the mitochondrion' by M. Ng et al., 2021, *Developmental Cell*. Volume 56, Issue 7, Pages 881-905. Copyright 2022 Elsevier B.V. Adapted with permission.

The first line of defence consists in the *selective degradation of non-assembled and misfolded mitochondrial proteins*, also known as MAD⁸⁷. Key components of this process are the mitochondrial proteases, also known as mitoproteases, that exclusively localise to the mitochondria and that regulate substrate proteolysis, mitophagy, apoptosis, lipid

biosynthesis, OXPHOS and protein import. The best studied are the ATP-dependent AAA+ proteases (e.g., presenilin-associated rhomboid-like protein [PARL], VCP/p97, CDC48 etc.), which recognise and degrade through ATP hydrolysis damaged mitochondrial membrane proteins into smaller peptides to be exported from the organelle or to be further digested into amino acids by oligopeptidases¹³³. Degradation of cytosolic and outer membrane mitochondrial proteins is also carried out by the cytosolic UPS, where the E3 ubiquitin ligase MARCH5 mediates ubiquitination of misfolded proteins on the OMM to initiate their proteasomal degradation^{131,133}.

Additional players of mitochondrial QC are MDVs, small vesicles originating from either the OMM or IMM that selectively incorporate mitochondrial cargo and deliver it to different destinations (e.g., multivesicular bodies or lysosomes) for degradation¹³³. Although MDVs occur at baseline housekeeping levels, their biogenesis can be rapidly upregulated in response to stress, hypoxia and mitochondrial insults to carry out their mito-protective role¹³³.

Upon accumulation of misfolded proteins, mitochondria can induce nuclear transcriptional responses aimed at decreasing the mitochondrial protein burden. Specifically, the *mtUPR* is involved in the upregulation of nuclear transcribed mitochondrial chaperones and proteases, that can then carry out their protective functions under mitochondrial stress conditions¹³³.

The dynamic nature of the mitochondrial network provides additional protection against mitochondrial damage. Two distinct protein machineries mediate cycles of *fission* and *fusion* of mitochondrial membranes and determine the shape of the mitochondrial network: Mitofusin 1 and 2 (MFN1 and MFN2) and OPA1 (Optic Atrophy 1, a mitochondrial dynamin-like GTPase) regulate mitochondrial membrane fusion in the outer and inner membranes respectively, while DRP1 (Dynamin-related Protein 1) controls fission events. While fusion allows content mixing between intact and damaged mitochondria, thereby replacing damaged material and contributing to the integrity of the intracellular mitochondrial population, fission allows sequestration of catastrophically and irreversibly damaged mitochondria and their subsequent degradation via mitophagy (see section below)^{87,131}.

Extensive mitochondrial damage and fragmentation results in the activation of another line of response, the *mitophagy* pathway, which sequesters damaged mitochondria from the intact mitochondrial network. Mitophagy can be defined as ‘the specific and selective, targeted removal of excess or damaged mitochondria from the cell via autophagy’^{134,135} and is characterised by the engulfment of dysfunctional or depolarised mitochondria by autophagosomes and their degradation through the lysosome. How damaged mitochondria are specifically selected is not clear yet, however increasing evidence suggests that mitochondrial dysfunction could trigger mitophagy by itself via mitochondria-derived ROS, which act as signalling molecules of dysfunctional mitochondrial homeostasis^{131,136}. The mitochondrial membrane potential ($\Delta\psi_m$) also works as an indicator of mitochondrial health: depolarisation of the mitochondrial membrane indicates altered mitochondrial function (e.g., halted ATP production) and, if persisting, mitochondrial damage, which as a result triggers mitophagy initiation¹²⁹.

In mammals, mitophagy can generally be classified into two main functionally distinct groups on the basis of the requirement for the kinase PINK1 and the E3 ubiquitin ligase Parkin, referred to as PINK1-dependent mitophagy and PINK1-independent mitophagy¹³³. The two pathways use different sets of autophagy receptors and their activation is dependent on different factors: while PINK1-dependent mitophagy can only be initiated by loss of the $\Delta\psi_m$, initiation of PINK1-independent mitophagy does not require mitochondrial membrane depolarisation. Despite these clear differences, cross-talk between the two pathways has been reported several times^{133,137}. Overall, mitophagy is crucial to maintain a healthy mitochondrial population and thus a high metabolic activity, which is particularly important for post-mitotic cells such as neurons and cardiomyocytes that have high energy demands and cannot dispose of dysfunctional mitochondria by cell division⁸⁷.

1.1.4.4. PINK1-dependent mitophagy

In the last decade, several lines of research reported that the decline in mitochondrial function observed in PD patients may stem from the genetic deregulation of mitochondrial quality control mechanism¹²⁸. In this scenario, a number of genes associated with autosomal recessive and early onset PD have been implicated in the

mitophagic degradation of dysfunctional, depolarized mitochondria: *PINK1* (Pten-induced kinase 1), a mitochondrially targeted serine/threonine kinase, and *PRKN* (Parkin), an E3 ubiquitin ligase^{117,138,139}. *PINK1* is the sensor of mitochondrial damage and triggers a signalling cascade that activates mitophagy¹³⁴. Briefly, under basal conditions, the N-terminal mitochondrial targeting sequence (MTS) directs *PINK1* to mitochondria, where it translocates across the outer OMM to the IMM through the TOM20/TIM23 complex. Following cleavage of the MTS by the mitochondrial processing peptidase (MPP) in the matrix, *PINK1* is cleaved within its transmembrane domain by the IMM-bound PARL mitoprotease, its resulting 52 kDa fragment is released into the cytosol and subsequently degraded by the proteasome, leading to undetectable basal levels of *PINK1* in the cell. Under mitochondrial stress conditions, such as membrane depolarisation, mitochondrial complex dysfunction and proteotoxicity, $\Delta\Psi_m$ -dependent mitochondrial import of *PINK1* is compromised, leading to its rapid accumulation and stabilisation outside the OMM. Subsequent *PINK1* homodimerisation and auto-/transphosphorylation on the OMM promotes the functional kinase activity of *PINK1* (i.e., active state) and the subsequent phosphorylation on serine 65 (Ser65) of ubiquitin (pUb) molecules anchored to some OMM proteins at low abundance. This *PINK1*-dependent phosphorylation of pre-existing ubiquitin at the OMM is required for the recruitment of the E3 ubiquitin ligase Parkin from the cytosol to depolarised mitochondria, where Parkin is activated by direct phosphorylation by *PINK1* (Ser65) on its ubiquitin-like domain. Upon transition into an active conformation, Parkin ubiquitinates several substrates at the mitochondrial surface assembling K63-, K48-, K11-, AND K6-linked ubiquitin chains, that in turn are further phosphorylated by *PINK1* in a ubiquitin-driven feed-forward mechanism, thereby amplifying the damage detection signal from *PINK1*. Once recruited to damaged mitochondria, Parkin carries out ubiquitination and proteasomal degradation of several OMM, notably MFNs (1 and 2), VDACs (voltage-dependent anion channels; 1,2 and 3), HKs (Hexokinases; 1 and 2), TOM20, Miro proteins etc. The accumulation of ubiquitinated proteins at the OMM ultimately triggers the recruitment of autophagy receptors (e.g., OPTN, NDP52, ATG5) and ubiquitin-binding adaptor proteins (e.g., p62/SQSTM1) to the mitochondrial surface, which bridge ubiquitinated cargo proteins with autophagosomes via the LC3 or GABARAP family members located on the outer surface of the autophagosomal membranes. OPTN and NBD52 recruited to ubiquitinated proteins at the OMM activate TANK-binding kinase 1

(TBK1) that in turn phosphorylates OPTN which as a result enhances its binding to polyubiquitin chains and further promotes TBK1 activation, thereby establishing a second feed-forward mechanism. Finally, after the rupture of the OMM follows the assembly of the phagophore, engulfment of cargo and, as a result, degradation via canonical macroautophagy (**Figure 1.10**)^{87,128,134,140}.

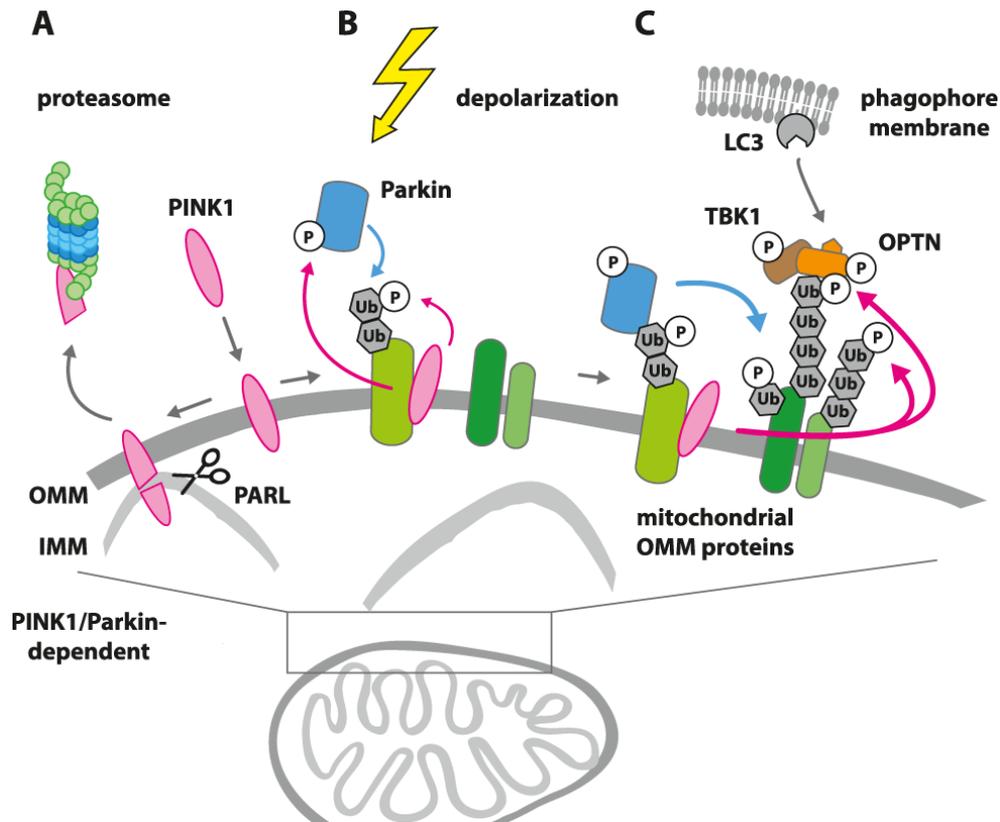


Figure 1.10. PINK1-dependent and mitophagy. (A) Under basal conditions, PINK1 is imported by the TOM20/TIM23 pathway and cleaved by PARL at the IMM. The resulting 52 kDa PINK1 fragment is then degraded in the cytoplasm by the proteasome. (B) Upon mitochondrial depolarization, PINK1 is stabilized at the OMM and phosphorylates ubiquitin molecules constitutively attached to OMM proteins. Parkin is recruited to the OMM by binding to phospho-ubiquitin and is activated by PINK1-dependent phosphorylation within its UBL domain. (C) Activated Parkin ubiquitinates several OMM proteins (green objects) that in turn are phosphorylated by PINK1, triggering a feed-forward activation loop. The autophagy machinery is recruited by autophagy receptors, such as OPTN and NDP52, which bind to LC3 via their LIR domains. A second feed-forward loop is established by activated TBK1 that phosphorylates OPTN and thereby enhances binding of OPTN to ubiquitinated cargo and activation of TBK1 by OPTN. Adapted from ‘PINK1 and Parkin: team players in stress-induced mitophagy’ by V. Bader, 2020, *Biological Chemistry*. Copyright Walter de Gruyter GmbH 2022. Adapted with permission.

1.1.5. Modelling FTL D *in vitro*

Currently, there are no disease-modifying treatments for FTL D. This may be due, in part,

to a lack of disease models that accurately recapitulate the genetic and molecular landscape of the disease¹⁴¹. In fact, the extreme complexity of FTLD pathology as well as the heterogeneity of affected cell-types significantly hamper progress towards disease-modifying therapies¹⁴¹.

Several culture systems have been developed to investigate the pathogenesis of NDs and to identify candidate drug targets, each carrying its own advantages and limitations. Traditional *in vitro* culture systems involve 2D monolayers often based on immortalised cell lines, such as human embryonic kidney 293 (HEK293) cells, human neuroglioma (H4) cells, pheochromocytoma (PC12) cells derived from the rat adrenal medulla or the human neuroblastoma cell line SH-SY5Y¹⁴². Although such models are widely used in neurodegeneration research because they can differentiate into neuron-like cells, their major limitation lies in the lack of a standardised protocol to maintain them in culture, which leads to variable cell growth and inconsistent experimental outcomes¹⁴².

Primary cultures of neurons are also commonly used models, as they have the potential to overcome many of the difficulties inherent to immortalised cell lines, but isolation and culture of primary neurons from human post-mortem brains is challenging. Therefore, primary neurons are usually obtained from embryonic murine brain tissue, due to their easy and rapid differentiation and formation of neurites and synapses in culture. Notably, primary neurons in this type of system are often contaminated with glial cells, such as microglia, which represents a fundamental advantage when investigating NDs: in fact, mixed cultures better reproduce the shared environment neurons belong to, allowing for more accurate characterisation of neuronal function and survival in a more similar physiological context¹⁴². Although *in vitro* immortalised cell lines and primary cultures have provided important insights into the mechanisms of numerous diseases, they are relatively simple systems that do not accurately replicate cellular organisation nor recapitulate disease-specific pathological features, leaving room for more advanced systems that are more closely related to human pathophysiology.

Notably, the development of induced pluripotent stem cell (iPSC) technology in 2006 by Takahashi et al.,^{143,144} has revolutionised *in vitro* modelling of complex traits and especially of NDs, for which cultures of human neurons are not available^{141,145}. Briefly, differentiated somatic cells, such as fibroblasts or peripheral blood mononuclear cells, can be taken from an individual presenting with a genotype and/or phenotype of interest, and are subsequently reprogrammed to a pluripotent state by inducing exogenous

overexpression of the pluripotency-associated transcription factors *Oct4*, *Klf4*, *Sox2* and *c-Myc*^{143,146}. The resulting iPSCs retain the ability for unlimited self-renewal and the potential to differentiate into cells of all three germ layers: they can therefore be differentiated into disease-relevant cell types, including multiple subclasses of functional neurons, astrocytes, oligodendrocytes and microglia¹⁴⁶. This way, researchers are now able to generate reproducible disease models containing the patient's precise genome in the cell type that selectively degenerates in disease. These human iPSC-neurons have the advantage of endogenous expression of the mutant gene of interest, in the cell type specifically affected by disease, thus enabling the development of patient-specific *in vitro* models of familial and sporadic FTLN¹⁴⁵. This approach has been successfully used in an ever-increasing number of studies to model the most common genetic causes of pure FTLN and of the FTLN-ALS continuum, including mutations in *MAPT*, *TARDBP*, *GRN* and *C9ORF72*. Nonetheless, several limitations apply: i) on a technical level, iPSC culture has considerably high costs and is very time-inefficient, limiting the number of patient and control lines that can be used in a single study; ii) although iPSC neurons do recapitulate the genetics of disease and present with some key FTLN pathological markers, these models do not faithfully recapitulate the full pathologies and phenotypes observed in FTLN patients; iii) iPSCs present with different degrees of inter and intra-patient variability due to cellular and epigenetic heterogeneity; iv) iPSC-neurons show a foetal rather than adult nature, which means that investigating age-related disorders (i.e., FTLN) is very challenging; v) iPSCs can be differentiated into a single cell type of interest at a time, which limits biological accuracy and relevance^{141,146}.

While some of these limitations still remain unsolved, many are currently being addressed, making iPSCs an extremely valuable, physiologically relevant model to investigate the molecular mechanisms underpinning FTLN and other NDs. Notably, to date FTLN research has largely focused on characterising pathological mechanisms in neurons, notwithstanding the emerging evidence supporting a role in pathogenesis for different neuronal subtypes as well as glial cells¹⁴⁷. Fortunately, recent progress in iPSC research has led to protocols that efficiently derive microglia from human iPSCs¹⁴⁸, which will have major implications in future studies investigating the role of neuroinflammation in FTLN¹⁴⁶.

While iPSC models still represent the 'golden standard' for modelling complex traits, increased optimisation of cell culture techniques is bringing forward 3D cell cultures,

such as organoids, as a valuable tool to model neurodegeneration. The fundamental advantage of employing organoids is that they reproduce the complex structural architecture and microenvironment of the brain much more accurately than simpler 2D models, by recapitulating different cell types together and thus replicating cellular cross-talk in disease. Although extremely advantageous, organoids also present with a major limitation: the lack of tissue maturity and vascularisation limits their usefulness. An alternative *ex vivo* modelling system closer to *in vivo* animal disease models is organotypic slice cultures, which consists of cutting and culturing of tissue-specific brain slices that represent the main areas affected by disease^{142,149}. Despite the significant advantages carried by organotypic models, which include a well-controlled genetic background and an accurate replication of physiological processes, they lack in reproducibility; in fact, dissection procedures require high precision, and variability between slices as well as mechanical damage during slice preparation are difficult to avoid. Additionally, these models originate from animals and thus present with significant differences in neural anatomy, physiology, gene expression patterns and drug metabolism compared to humans¹⁴².

In conclusion, current experimental models of FTLD recapitulate different aspects of the diverse phenotypes observed in patients, with different models recapitulating some disease aspects, but no model faithfully recapitulating all disease aspects. This does not come as a surprise given the genetic and molecular complexity of the disease, but it should be noted that many of the current models being pursued are driven by genetic mutations either observed in only a small minority of patients, or that are known to have divergent disease mechanisms, thus not representing the majority of cases or the most common overlapping pathologies¹⁴⁵. In this scenario, it could be argued that it might be slightly unrealistic to expect a 'perfect' model (i.e., a perfect replica of human FTLD pathology in the majority of cases), as it is likely that all models will continue to fail in this aspect. These diseases are complex and multigenic, indicating that experimental models may need to be targeted to different disease aspects. This would allow information to be gathered from a variety of different yet relevant models, each of which has the capacity to capture a specific aspect of the disease, and together can provide a more complete understanding of these complex and multi-layered diseases.

1.2. Emerging technologies powering the genetic and molecular dissection of complex diseases

The last decade was characterised by tremendous advances in the technologies and informatics tools for generating and processing large biological datasets (omics data), resulting in a critical shift in the study of biomedical sciences¹⁵⁰. High throughput technologies have now generated an extensive catalogue of genes and genetic variations associated with complex disorders and in recent years we have been witnessing the rise of interdisciplinary data integration strategies to support a better understanding of biological systems¹⁵¹⁻¹⁵³. Whilst the study of genomics, transcriptomics and proteomics, combined with informatics and biostatistics, is progressing at an unprecedented pace, they are still, for the most part, evaluated separately with distinct approaches generating sectorial rather than integrated information^{150,154}.

Genome wide association studies (GWAS), DNA microarray, exome and genome sequencing technologies have helped identifying large numbers of genetic loci and variants that significantly associate with increased risk of developing diseases. However, translating genetic knowledge into the understanding of the molecular mechanisms underlying complex traits has to date proved to be a major challenge. Indeed, while genetic association has greatly aided shedding light on wide variety of neurological disorders, such knowledge is still insufficient to fully explain disease pathogenesis⁹⁷. At present, there is no straightforward translation of genetic knowledge into its corresponding functional landscape of biochemistry and cell biology, constituting a substantial gap in our understanding of the molecular mechanisms underpinning disease. In this scenario, bioinformatics coupled with data-driven frameworks operating at a system or a network level represent an emerging and powerful tool to improve genetic and expression analyses before wet laboratory work is performed, thus constituting a valuable support to guide cellular and biochemical investigations^{97,155}.

1.2.1. Using networks for biological inference

For over a century, reductionism has been the most preeminent approach in biological research, generating an enormous wealth of knowledge on individual cellular components and their functions¹⁵⁴. Despite its vast success, it is becoming increasingly apparent that distinct biological functions only rarely can be ascribed to a single effector

gene product^{4,156}. Instead, most of the biological dynamics arise from complex interactions between the cell's various constituents, including proteins, DNA, RNA and small molecules. As a result, a major challenge in modern biomedicine is now to understand the organisation and the dynamics of the complex inter/intracellular interaction networks that determine the structure and functioning of living cells^{4,156-158}.

Recent advancements in the study of network biology revealed that cellular networks are governed by universal organising principles that are shared to a great extent by other complex systems (e.g. the Internet, computer chips, society), implying that similar principles may govern most complex networks in nature^{156,159}. This research has had huge implications in the cell biology field, indicating that expertise from well-mapped non-biological systems could be implemented to characterise the intricate web of interactions underlying cellular functions.

In summary, by measuring existing or predicted relationships among genes, gene network methods provide an essential organising framework that places each gene in the context of its cellular or tissue system, providing valuable insights into the different levels of molecular organisation within the hierarchy of brain region, cell type, organelle and molecular pathways (**Figure 1.11**)¹⁵⁵.

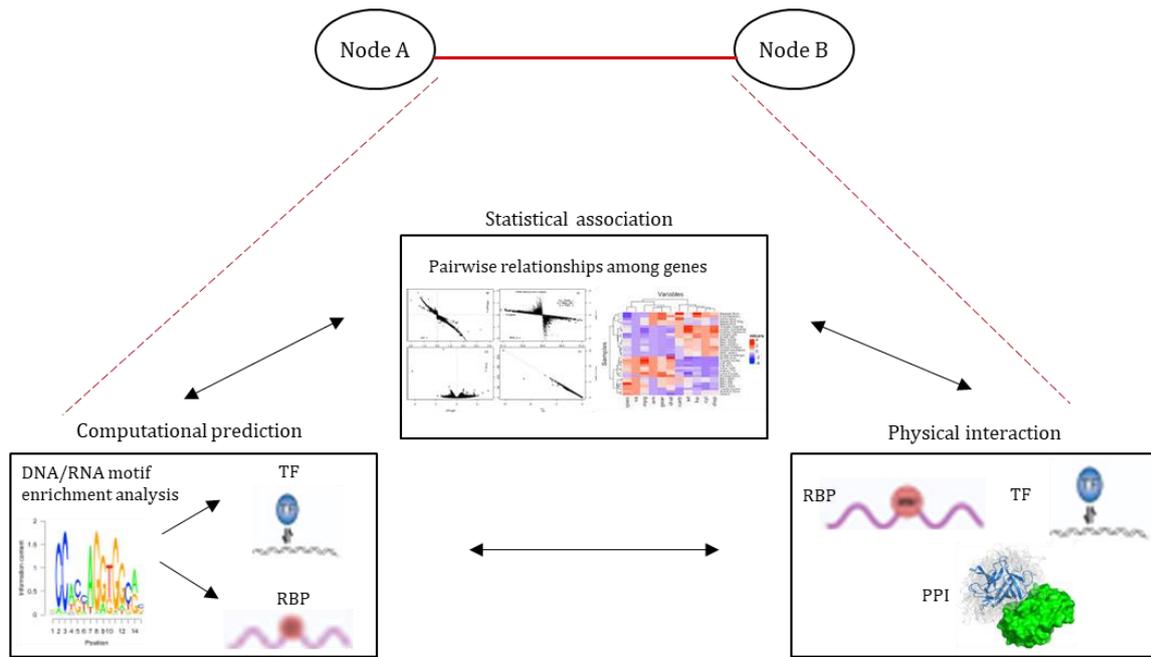


Figure 1.11. Measuring existing or predicted relationships among genes/proteins. Network analysis involves connecting genetic or molecular nodes with information about: i) pairwise relationships (i.e., statistical associations relating molecular patterns measured across experiments such as gene expression level variation or co-expression); ii) physical interaction data from experiments or curated from literature (e.g., transcription factors [TFs], RNA-binding proteins [RBPs] or protein-protein interactions [PPIs]); or iii) computational predictions about TFs or RBPs using motif enrichment analysis.

1.2.2. Network theory and topological properties

Network theory offers several quantifiable tools to explore and understand complex networks and thus the cell's internal organisation and evolution. Within network theory, the network components can be reduced to a series of nodes, or vertices, that are connected to each other by edges, or links, with each edge representing the interaction between two components^{156,158,159}. Mathematically, the network formed by a set of edges and a set of nodes is called 'graph'¹⁶⁰. A node is an entity to which we can ascribe some value and static properties. Edges can be defined as a relation of some sort between two nodes. These relations can be tangible, such as physical protein-protein interactions (PPI), or intangible, as in they may reflect statistical similarity (e.g., correlation) or computational inference¹⁵⁵. The nodes belonging to an edge are called the 'ends', 'endpoints' or 'end-vertices' of the edge. Depending on the nature of the interactions, graphs can be directed or undirected¹⁵⁹. Within an undirected graph, edges have no assigned orientation and represent a mutual binding relationship, meaning that if protein A binds to protein B, then protein B also binds to protein A. Directed graphs are set of

nodes connected by edges that have a well-defined direction associated to them, representing for instance the activation cascade of a signalling pathway; these are typically denoted with arrows indicating the direction¹⁵⁶. In this study, we will be using only undirected networks.

1.2.3. Modularity in biological networks

Networks provide a framework for deriving information from a set of relationships among biological entities¹⁶¹. In this scenario, edges play a key role in the interpretation of the biological implications underlying networks: they define the connectivity of nodes to each other in a network, and such connectivity can be used to organise the nodes into smaller collections of highly interconnected and differentially activated modules^{3,155,161,162}. Inter-modular connectivity reflects the topology and the non-random organisation of biological networks, both of which encode information about how molecular interactions contribute to biological phenotypes, and can identify which genes are biological hubs within modules^{155,161}. In biological networks, hubs are highly interconnected genes that are highlighted as relevant in the molecular process or disease of interest¹⁵⁵. As a matter of fact, molecular interaction networks within the cell have often been shown to be modular, meaning that genes/proteins implicated in the same biological process or disease often segregate to form modules within networks^{4,161,163}. Therefore, in this scenario modularity is very useful because it provides a general organising principle in the study of biological networks by either highlighting key molecular drivers in disease, such as transcriptional regulators that drive genes differential expression, or by functionally annotating a module and suggesting its predominant biological role^{155,161}. As a result, inference based on network architecture can be used to: i) prioritise candidate disease genes and pathways; ii) annotate unknown cell types or biological processes on the basis of their proximity to marker genes of known function (following the ‘guilt by association’ principle); iii) evaluate how a module or the topology of specific genes within a module change in health and disease; iv) elucidate pathway architecture by extracting relevant modules^{161,162,164}.

From a structural perspective, modules in a network can be defined using either a seed-based (prior-based) or a genome-wide approach¹⁵⁵. The seeded approach requires a set

of genes of interest to start with (selected using prior knowledge), expanding edges to bring in additional unannotated genes and clustering highly connected components into modules. The unseeded approach involves starting with unannotated genes, using edges to pinpoint interconnected components (modules) and then assessing where genes of interest fall in the resulting network structure. Importantly, both these approaches allow for integration of external information (e.g., genetic associations, functional enrichments) to further annotate modules and can both be used to prioritise candidate genes and pathways for experimental validation. The prioritisation process generally involves integrating diverse data sources to determine the ranking of the nodes in the network and to identify groups of functionally related genes, down to a smaller set of putative regulatory genes¹⁶⁵.

1.2.4. Gene co-expression networks

Gene co-expression networks (GCNs) are networks based on similarity in gene expression. They are context-specific by definition, as they directly leverage phenotype-specific transcriptomic data to construct the network and then correlations between the expression profiles of each gene pair are calculated³. Briefly, a similarity score (i.e., correlation coefficient) is calculated from the pairwise comparison of the gene expression patterns for each possible pair of genes. Above a certain threshold, gene pairs form a list of nodes and corresponding edges from which the network is constructed. As a rule, the main assumption underlying the interpretation of GCNs is the guilt-by-association principle, which affirms that genes presenting similar expression profiles are generally involved in the same regulatory processes, and hence will form modules within the network. On these grounds, although correlation is not causation, genes with coordinated expression can be assumed to have functional linkages between them, either by being parts of the same molecular complex or mechanism, or by playing a regulatory role modulating their reciprocal expression^{3,166}. This way, within the same module, genes of known function can be used to predict the biological function of unknown co-expressed genes in an intuitive way, help discerning gene transcriptional regulatory mechanisms *in vivo* and prioritising candidate regulatory genes or modules of disease traits^{166,167}. An increasing number of studies have supported the versatility of co-expression analysis for

inferring and annotating gene functions, making them a very promising tool to further enhance the elucidation of gene regulatory relationships¹⁶⁵.

1.2.5. Protein-protein interaction networks

The subcellular environment hosts a dynamic network of molecular events that regulates cell homeostasis and coordinates signal transduction¹⁶⁸. In this scenario, proteins are the main agents of most of the distinct biological phenomena and their functionality is based on their interaction with a variety of other molecules, including other proteins, DNA and RNA¹⁶⁹. The term protein-protein interaction (PPI) refers to a variety of events happening inside the cell involving two or more proteins physically interacting with each other to form stable or transient protein complexes^{168,169}. PPI networks store the information about the protein-protein interactome of a given organism, that is the whole set of its PPIs, by recapitulating evidence of proteins forming complexes through experimentally proven biochemical events and/or electrostatic forces or as a result of a computational prediction¹⁶⁹.

PPI networks are particularly useful to elucidate how PPIs are wired together and mediate cellular response to environmental and genetic cues, and their study is crucial to understand when and where these proteins perform their functions determining healthy and diseased states of organisms¹⁶⁹. Every single interaction within the PPI network can be mathematically evaluated to shed light on the global relationships among candidate contributors to disease mechanisms, where the network is considered an *in silico* model system that can be used to i) better understand the proteome landscape underlying disease, ii) generate novel hypotheses and iii) further support functional research and disease modelling^{153,168}. Similarly to GCNs, the guilt-by-association principle also applies to PPI networks, implying that nodes interacting with each other or topologically segregated within the network generally act together in the same biological processes. Dissecting the biological complexity of these interactomes is crucial to improve prediction of gene function and cellular behaviour in response to diverse signals, but to realize this potential, comprehensive mapping and functional annotation of PPIs is key¹⁷⁰.

1.3. Ongoing and future efforts to tackle FTLD: the rise of systems-level computational approaches

When compared to other neurodegenerative diseases such as AD or PD, FTLD is a rather rare condition with an insidious clinical presentation that embraces a wide spectrum of syndromes. Such complexity translates to several analytical, technical and interpretational challenges which remain a fundamental gap in our understanding of its pathophysiology. Although large-scale genetic studies have begun to decipher the genetic architecture FTLD by identifying hundreds of genetic loci and variants involved in disease risk, their contribution to disease is still unclear. To unravel a hierarchically organised complex system such as the central nervous system (CNS) and understand how genetic variability contributes to disease, experimental scientists currently rely on models that only account for a few features of the CNS at a time¹⁵⁵. While this approach has been successful to characterise some highly penetrant variants that result in clear phenotypes, it has been less fruitful for genetically complex disorders, such as FTLD and other neurodegenerative diseases¹⁵⁵.

Recent studies suggested that the adoption of integrative omics as well as the expansion and wider characterisation of FTLD syndromes cohorts are critical to allow powered genetic and functional studies of sporadic and familial FTLD^{9,77}. In this instance, large international collaborations involving multiple research centres worldwide allow for such expansion. The International Frontotemporal Dementia Genomics Consortium (IFGC; <https://ifgcsite.wordpress.com/>) is among the largest consortia for the study of sporadic FTLD, comprising groups from Europe, North America and Australia who share an interest in the genetics and understanding of sporadic FTLD. Its vision entails the use of: i) genetics, to expand on genes and genetic markers that cause or increase the risk of developing FTLD and ii) bioinformatics, to interpret genetic data and predict risk pathways *in silico*⁹. By allowing well-powered cross-disciplinary studies, this and other consortia clearly represent valuable resources for the wider research community with an interest in FTLD and neurodegeneration. Additionally, these efforts also promote large-scale meta-analyses which help dissecting syndrome-specific genetic fingerprints and molecular phenotypes and allow a better understanding of FTLD and closely related neurodegenerative conditions.

However, we have still not succeeded in translating this wealth of genetic and functional

knowledge into actionable information about impacted biological processes and molecular mechanisms at the basis of disease, and this is currently among the major and most debated topic in biomedical research when investigating complex disorders, including FTLD.

Mendelian FTLD genes are particularly informative to functional biologists to design experiments around their protein products and further characterise them in *in vitro* and *in vivo* models⁹. Although it might be argued that familial genes account for ‘just’ a minority of cases (all together ~30-40% of all FTLD), an intriguing hypothesis proposed by Ferrari et al.⁹ suggests that Mendelian genes could indicate functions and processes whose alteration is necessary and sufficient to trigger FTLD pathogenesis. This would imply that such genes could play a crucial role to define the global pathogenic mechanisms leading to FTLD and, as a result, to be informative also for the vast majority of sporadic cases (~60-70% of all FTLD). In this scenario, in fact, the genetics underlying sporadic cases are still widely understudied especially due to the nature of GWASs, which do not provide clear causative links to specific genes⁹.

The study and functional characterisation of Mendelian and sporadic/GWAS genes’ protein products is also, in the first instance, very informative on the potential impacted biological processes, but similarly to the genetic studies mentioned above, it tends to take into consideration ‘one gene at the time’, resulting reductionist in the long run. Additionally, proteins encoded by candidate disease-risk genes are generally involved in multiple subcellular processes/pathways, making it very hard to determine whether they are truly involved in disease pathogenesis.

While clearly the study of FTLD genes and their protein product has and is still giving an essential contribution to drive research efforts aimed at better characterising FTLD pathology, they tend to lean towards a more reductionist approach, thus substantially failing to globally evaluate the genetic and molecular players contributing to the phenotype. In this scenario, innovative methods relying on data integration and bioinformatics analysis have emerged as an alternative to classical reductionist studies, allowing for a simultaneous and comprehensive evaluation of the genetic and molecular landscape of disease. Application of computational approaches has also allowed for *in silico* simulations of different pathogenic scenarios on the basis of previously generated biomolecular profiles (mRNA expression, miRNA, and noncoding RNA profiling, proteomics and metabolomics measurements) and, as a consequence, for the isolation of

the most promising risk-pathway for validation in a functional setting^{9,163}.

From a genetics perspective, the currently available and emerging high-throughput technologies such as genome-wide, exome-chip arrays and next generation sequencing (NGS) techniques, are becoming more cost-effective, enabling a better characterisation of common and rare genetic variability contributing to complex diseases^{9,163}. These technologies will be essential to fine-map classical GWAS loci for identification of causal variants, to explore in more depth the (likely) oligogenic nature of disease and to aid the identification of novel causative genes, contributing to unfold and define the genetic risk architecture of FTLD^{9,25}. Additionally, implementation of these techniques on larger cohorts representative of the different FTLD subtypes (i.e., clinical variants bvFTLD, SD, PNFA and FTLD-MND as well as pathologically defined cohorts [e.g. FTLD-tau or FTLD-TDP]) will be crucial to aid improvement of genotype-phenotype correlation defining syndrome and/or subtype-specific genetic fingerprinting. Undoubtedly, improved subtype clinical and genetic classification will considerably benefit patient diagnostics as well as our mechanistic understanding of disease, allowing for the identification of suitable patient cohorts qualifying for tailored clinical trials.

Furthermore, better study designs have the potential to reduce the gap between transcriptomics and proteomics when coupled with multi-omics technologies that can be applied to the same disease-relevant tissue(s), minimising inter-sample and tissue specificity issues¹⁵⁰.

From a functional perspective, mechanistic understanding of disease continues to lag behind the pace of gene discovery, and progress in experimental technologies is hampered by specimen availability and cost as well as its time-inefficient nature¹⁵⁵. More time-effective ways to coherently translate genetic into functional knowledge need to be found by using more integrative and systems-level approaches to interpret the present genome data and to guide functional studies for targeted investigation of risk cellular pathways¹⁶³. These include examination of genetic variability's influencing gene expression modulation (e.g. expressed quantitative trait loci [eQTLs], methylation quantitative trait loci [mQTLs]^{150,171}, allele-specific expression, transcriptome-wide association study [TWAS]) and its association with a trait, as well as evaluation of molecular interactions and functional annotation analyses of gene co-expression and protein-protein interaction (PPI) networks that can aid the prioritisation of biological processes that are impacted by genetic variability¹⁵³. New avenues are currently being

explored in the context of single-cell (-nucleus) RNA sequencing (sc or snRNAseq), which has significantly advanced our understanding of several neurodegenerative disorders by identifying major brain cell types affected by degeneration. Importantly, cell-type-specific transcriptomic profiling can help dissecting the complexity and heterogeneity of molecular mechanisms underlying FTLD pathology and suggest basis for selective vulnerability and resilience of brain cell types by delineating cell-type-specific FTLD-linked gene expression and molecular signatures in brain tissues of interest (e.g., FCTX)¹⁷². As a result, undoubtedly this multi-level knowledge will provide solid ground for the development of testable hypotheses and help functional biologists designing more accurate experimental models.

Standardising all such strategies will require some time, nevertheless this paradigm shift is essential to improve basic and translational research and pave the way to for advancements in preventive, monitoring and therapeutic measures for FTLD and other complex diseases.

1.4. Significance of the research and objectives of this thesis

Despite tremendous progress in characterising the genetics of FTLD and related pathology over the past 20 years, the lack of understanding of how genetics, phenotypic, and pathological features are wired by underlying molecular mechanisms represents the major gap to the dissection of FTLD pathogenesis⁹. Importantly, advances in omics technologies — such as genomics, transcriptomics, proteomics and metabolomics — and the development of innovative methods relying on data integration and bioinformatics analyses have begun to aid our understanding of the biological complexity of most human diseases⁷⁷. These are emerging alternatives to the classical studies in that they allow to evaluate altogether the genetic players contributing to the phenotype and to isolate the most likely risk pathways to be validated and tested in the functional setting. For instance, network analyses based on gene co-expression and protein-protein interactions (PPIs) are becoming suitable methods to serve these purposes.

This thesis aims to build on these emerging alternatives to further dissect the genetic and molecular underpinnings of FTLD. Here I present an integrative pipeline that combines

weighted protein-protein interaction network analysis (WPPINA)⁹⁷ and weighted gene co-expression network analysis (WGCNA)¹⁷³ to analyse the specific interactome and co-expression transcriptome of known FTLD-associated genes/proteins. The current work uses current state-of-the-art genetics associated with a trait (i.e., FTLD) and projects it into the transcriptome and protein domains. By integrating genomic, transcriptomic and proteomic data, this cross-disciplinary pipeline supports the global understanding of the FTLD-associated biological processes as well as of the genes/proteins contributing to those processes, and that might therefore play a pivotal role in controlling the shift in balance between the healthy and the pathological scenarios.

Here I hypothesise that there might be a network of genes/proteins functionally linked with the major FTLD Mendelian and GWAS genes which: i) contribute to the alteration of biological processes impacted in FTLD and ii) might hold critical relevance to the functional pathogenesis of FTLD. The major aims of this project are to: i) investigate the functional environment of FTLD *in silico* by identifying disease-specific biological processes and genes/proteins connecting the majority of the FTLD genes; ii) carry the prioritised biological processes and genes/proteins forward for hypothesis-driven functional validation using *in vitro* cellular models. Additionally, a separate study aims at specifically characterising *C9ORF72* repeat expansions as a potential genetic modifier of FTLD clinical variants in relation to genetic ancestry and AAO in an international sporadic FTLD cohort.

Chapter 2

2. Material and methods

2.1. C9ORF72 repeat expansion screening

2.1.1. Cohort, clinical phenotyping and patients consent

FTLD cases were collected between 2016 and 2018 (within the IFGC phase-III project [<https://ifgcsite.wordpress.com/ongoing-projects/>]). The samples were recruited by clinicians and research groups who are part of the IFGC network and based in Italy, Spain, Germany, the Netherlands, Belgium, UK, Sweden, Norway, Slovenia, and USA (**Supplementary Table 1 [Supplementary File 1]**). Patients were diagnosed at each contributing site (**Supplementary Table 2 [Supplementary File 1]**) in a harmonised fashion according to international consensus criteria such as the Neary *et al* (for FTLD), Rascovsky *et al* (for bvFTLD), Gorno-Tempini *et al* (for PPA [SD or PNFA]) and Strong *et al* (for FTLD-MND) criteria^{15,174,175}. Each contributing site obtained written informed consent from all patients to be part of extended genetic studies (and IRB approval #9811/001).

2.1.2. Genotyping, C9ORF72 repeat expansions and analysis cohorts

Thousand four-hundred and fifty-four (1454) cases were successfully genotyped by means of the NeuroArray¹⁷⁶ on the Illumina Infinium platform. Genotypes were used to inform on population substructure *via* standard principal component analysis (PCA) (**Supplementary Figure 1 [Supplementary File 1]**), which led to the exclusion of 44 population outliers, and allowed to address population-substructure within the cohort (we identified 2 distinct ['Nordic' and 'Mediterranean'] clusters; **Supplementary Figure 2 [Supplementary File 1]**). I also assessed cryptic relatedness and excluded 14 first- or second-degree related individuals, leaving a cohort of 1396 cases (*group 0*) – for which *C9ORF72* expansion status (i.e., presence/absence of pathogenic expansions) was known – for analyses. Frequencies of pathogenic expansions were assessed in *group 0* and further analyses were performed in: i) 1295 cases (*group 1*: n = 800 bvFTLDs and n = 495 PPAs) with known *C9ORF72* expansion status; ii) 1179 cases (*group 2*; n = 756 bvFTLDs

and n = 423 PPAs) with known *C9ORF72* expansion status and age at onset (AAO) data available, and; iii) 734 cases (*group 3*; n = 462 bvFTLDs and n = 272 PPAs) with AAO and repeat counts (rc; screened *via* repeat-primed PCR [RP-PCR] [c.f.^{19,33}], see **Supplementary Materials and Methods** and **Supplementary Figure 3 [Supplementary File 1]**) data available (**Figure 2.1**).

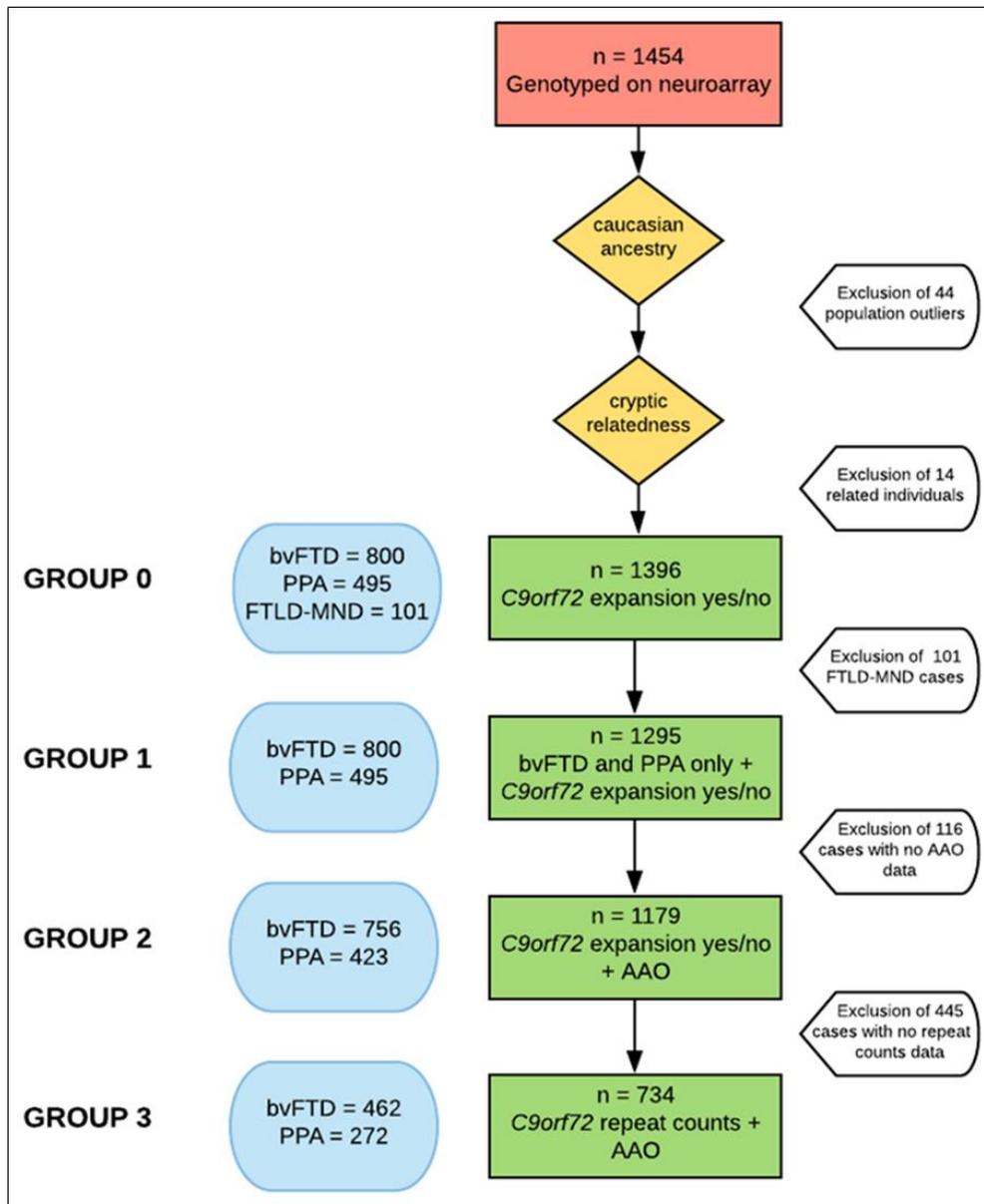


Figure 2.1. Study cohorts. AAO, age at onset; bvFTD, behavioural variant frontotemporal dementia; PPA, primary progressive aphasia; FTD-MND, frontotemporal dementia with motor neuron disease. Adapted from ‘*C9ORF72*, age at onset, and ancestry help discriminate behavioural from language variants in FTL-D cohorts’ by B. Costa, 2020, *Neurology*. Copyright 2022 American Academy of Neurology. Adapted with permission.

2.1.3. Statistical analyses

I first assessed the frequency of pathogenic expansions in the entire cohort (*group 0*). The information on presence/absence of expansions was used as a binary variable (0 = absence of expansion; 1 = presence of expansion). I then investigated differences in the frequencies of pathogenic expansions across bvFTLDs and PPAs, and the 'Nordic' and 'Mediterranean' clusters in *group 1* (Fisher's Exact test) and in *group 3* (logistic regression); in the latter, I used repeat counts (rc) as a categorical variable (using 'no', 'short', 'intermediate' and 'long' as factor levels) considering the following 4 categories: 'no' expansions (rc = 2/3), 'short' expansions ($4 \leq rc \leq 8$), 'intermediate' expansions ($9 \leq rc \leq 24$) and 'long' expansions ($rc \geq 25$), the latter representing expansions in the pathogenic range (c.f.^{35,177}); see also **Supplementary Materials and Methods** and **Supplementary Figure 3 [Supplementary File 1]**.

I then evaluated association between AAO and syndrome, genetic ancestry and expansions (i.e., presence/absence used as a binary variable, see above) alone and with genetic ancestry as a covariate in *group 2* (t-test and logistic regression) and in *group 3* (t-test, ANOVA with Tukey *post-hoc* test, and logistic and linear mixed-effects model). In the latter case, I used rc as a categorical variable (see above).

Finally, I sought to build a model to predict syndrome (bvFTLD vs. PPA) using (i) presence/absence of pathogenic expansions (as binary variable [see above] for *group 2*) or (ii) rc (as categorical variable [see above] for *group 3*), ancestry as binary variable and AAO as continuous variable using logistic regression models (i.e. the leave-one-out cross validation [LOOCV] and the K-fold models). A summary of the analyses workflow can be found in **Figure 2.2**.

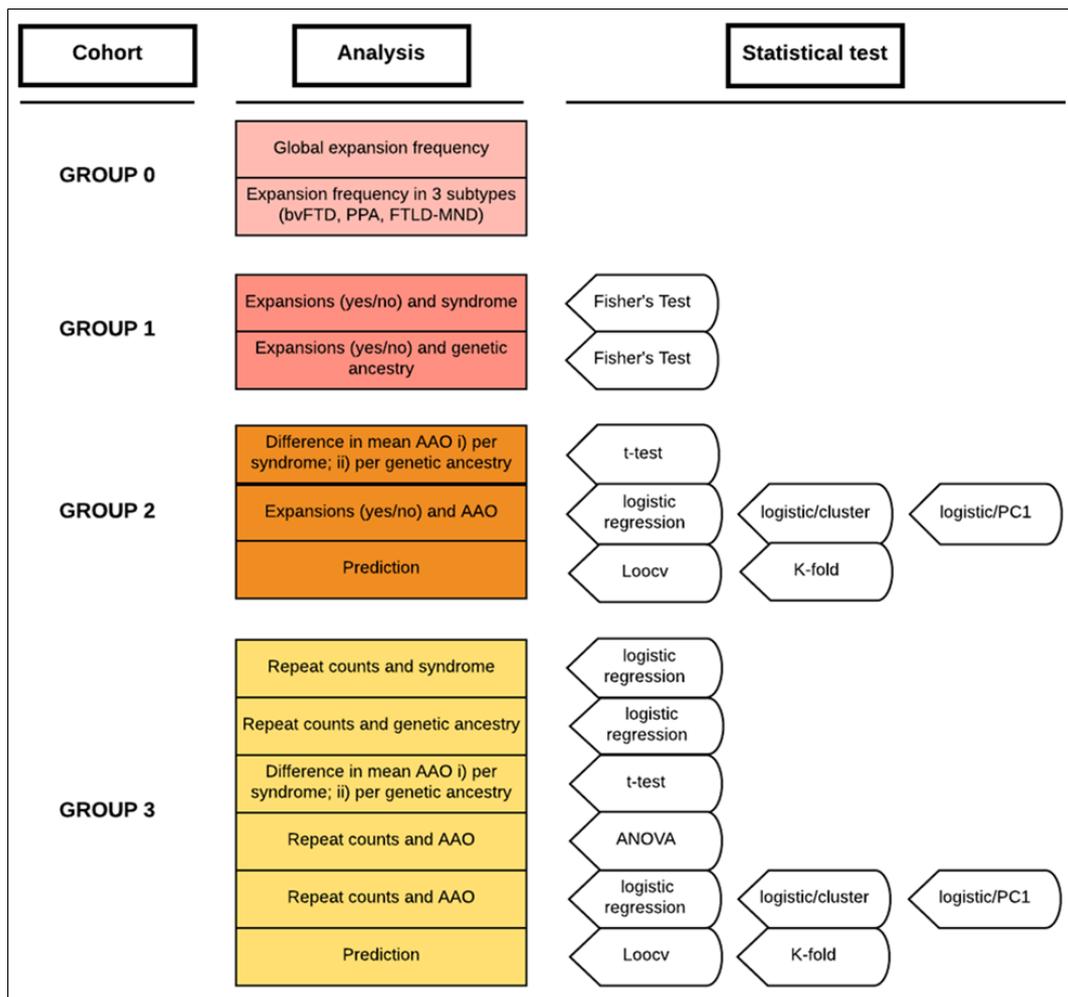


Figure 2.2. Analysis workflow. AAO, age at onset; logistic/cluster = logistic regression using cluster as covariate; logistic/PC1 = logistic regression using PC1 as covariate; LOOCV = leave one out cross validation regression model; k-fold regression model. Adapted from 'C9ORF72, age at onset, and ancestry help discriminate behavioural from language variants in FTLN cohorts' by B. Costa, 2020, *Neurology*. Copyright 2022 American Academy of Neurology. Adapted with permission.

I individually performed the genetic screening of the samples, whilst all the statistical analyses (R studio [version 3.6.0, studio version 1.2.1335]) and data interpretation were carried out in collaboration with Drs. Raffaele Ferrari and Claudia Manzoni. All data generated or analysed during this study are included in **Supplementary File 2**. Supplementary files are available at <https://rdr.ucl.ac.uk/articles/dataset/Supplementary-Files-Neurology-pdf/12418157>.

2.2. Bioinformatics

2.2.1. Weighted Gene Co-expression Network Analysis (WGCNA)

Data for WGCNA were collected from GTEx and UKBEC databases using the CoExp Web application¹⁷⁸. Details of these resources are provided below.

The Genotype-Tissue Expression (GTEx) project is an ongoing effort to build a comprehensive public resource to study tissue-specific gene expression and regulation. Samples were collected from 54 non-diseased tissue sites across nearly 1000 individuals, primarily for molecular assays including WGS, WES, and RNA-Seq. The GTEx Portal provides open access to data including gene expression, expression quantitative trait loci (eQTLs), and histology images.

The UK Brain Expression Consortium (UKBEC) studies the regulation and alternative splicing of gene expression in multiple tissues from human brains.

This dataset currently comprises 134 brains from individuals free of neurodegenerative disorders. Up to twelve brain regions were extracted per brain in parallel for mRNA quantification. DNA was also collected to enable genotyping and eQTL analysis. Data were generated through microarrays. Braineac - the Brain eQTL Almanac – is the web server for data from UKBEC.

Gene co-expression networks were generated by WGCNA and modules of highly correlated genes were determined in an unsupervised manner based on co-expression patterns in frontal cortex (FCTX), the classically affected brain area⁹³. The CoExp Web application (<https://snca.atika.um.es/coexp/Run/Catalogue/>, accessed on 05-12-19) was employed to extract gene co-expression modules, annotate gene sets of interest using GTEx V6 (47 co-expression networks on control tissue including 13 brain areas <http://github.com/juanbot/CoExpGTEx>) and 10UKBEC (10 Illumina microarray based gene expression profiling networks from brain tissue, <http://github.com/juanbot/CoExp10UKBEC>) as catalog categories. To identify highly interconnected genes within each module I used the measure of module membership (MM), a Pearson correlation between gene-expression level and module-eigengene^{93,173}; an elevated MM (≥ 0.5) suggests strong inter-correlations between genes in a module. I

finally evaluated whether modules' consistency and preservation were observed across the two datasets (GTEx vs UKBEC).

Furthermore, I performed functional annotation analysis for the FTLD-genes containing modules to characterise their biological relevance by means of the bioinformatics tool gProfiler (accessed on 10-02-20) (please refer to the functional enrichment section for more details).

2.2.2. Weighted Protein – Protein Interaction Network analysis (WPPINA)

The entire FTLD protein-protein interaction network (entire-FTLD-PPIN) was constructed in a multilayer fashion. FTLD Mendelian, FTLD-GWAS and FTLD-ALS spectrum genes (from now onwards identified as seeds) were used to extract their first layer interactor proteins. The proteins in the first layer were then used as seeds to download a second layer of PPIs. To summarise, the entire-FTLD-PPIN is composed of all FTLD seeds, plus their first layer interactors, plus their second layer interactors. Each seed's complete interactome is constituted by the seed under investigation, plus its first layer of interactor proteins, plus its second layer of interactor proteins.

The web server PINOT (Protein Interaction Networks Online Tool¹⁷⁹; <https://doi.org/10.1186/s12964-020-00554-5>, accessed on 24-12-19) was employed to extract the PPIs associated with each seed protein. PINOT is a web-source that optimises the collection and processing of PPI data from the IMEx consortium associated repositories. PPI data is downloaded live from PSICQUIC (the Proteomics Standard Initiative Common QUery InterfaCe) and then merged, quality checked, and confidence scored based on the number of detection methods and publications in which each interaction has been reported. Of note, all the interactions with a final score < 3 were removed as they are not reproduced in literature (identified with just 1 method or 1 publication or 1 method + 1 publication). Ubiquitin(s) (UBB, UBC) were also removed as they normally conjugate to proteins flagging them for degradation and thus they likely introduced false positive interactions¹⁸⁰.

In line with the pipeline presented by Ferrari et al.^{97,151,153}, I sought to identify inter-interactome hubs (IIHs), which are highly interconnected genes within the entire-FTLD-PPIN that are responsible for keeping network cohesion. The IIHs and their related interactomes were extracted from the entire-FTLD-PPIN to obtain a network of nodes which are highly shared in FTLD (core-FTLD-PPIN) and thus likely to be relevant for FTLD disease phenotype. I performed Gene Ontology (GO) terms enrichment analyses in gProfiler (accessed on 24-12-19) to analyse the functional enrichment of the core-FTLD-PPIN (please refer to the functional enrichment session for more details). The network was visualised using the software Cytoscape 3.7.1.

2.2.3. Functional enrichment analysis

I performed functional enrichment analyses for GO terms in gProfiler (g:GOST, <https://biit.cs.ut.ee/gprofiler/gost>). gProfiler settings were used as follows: enrichment for GO-BP terms only; Fisher's one-tailed tests as statistical method for enrichment, SCS-threshold as multiple testing correction; I considered significant those GO terms with $p < 0.05$. Enriched GO-BP terms were grouped into custom-made classes based on semantic similarity (semantic class). For each semantic class, all of the contributing single GO terms were merged to identify the list of proteins within the network that contribute to the enrichment of that specific semantic class. Semantic classes were categorized into more general "functional blocks", in order to highlight the main biological processes enriched for our genes.

2.2.4. Evaluation of gene expression in brain

Gene expression levels in FCTX were evaluated using both Braineac and GTExV6. We downloaded FCTX expression data from microarray provided by Braineac (available on <http://www.braineac.org/>, accessed on 02-03-20), mapped all the exprIDs to their correspondent HUGO gene names using RStudio and averaged all the expression values derived from each probe and sample. The same procedure was followed for the GTExV6 data, from which I downloaded the RNA-seq data file "Gene TPMs" (Transcripts Per Millions) (available on <https://www.gtexportal.org/home/datasets>, accessed on 02-03-20) and I extracted expression data for my genes of interest using RStudio.

2.2.5. Software

Data were handled, filtered and scored through in-house R scripts (available at https://github.com/bbeacosta/Upgrade_scripts.git) (RStudio version 3.6.0, studio version 1.2.1335). Gene co-expression networks exploration and hierarchical clustering of the modules were performed using the CoExpNets suite of R packages (available at <https://github.com/juanbot/CoExpNets>). The PPI networks were visualised through the freely available Cytoscape 3.7.1 software (<http://www.cytoscape.org/>). Graphs were generated through either RStudio or GraphPad Prism 9, while pipelines were designed using Lucidchart (available at <https://www.lucidchart.com>) or Microsoft PowerPoint.

2.2.6. Re-evaluation of WGCNA networks

After having assessed the functional environment of *CDC37* *in vitro*, I sought to: i) assess whether the relationships found in cellular models were also true in *in silico* models; ii) evaluate *CDC37* co-expression with the major FTLD Mendelian genes to aid FTLD model selection. To this end, I performed follow-up WGCNA analyses to re-evaluate the co-expression networks of *CDC37* specifically in frontal cortex. I selected the gene-co-expression network modules from both gtxv6 and UKBEC where *CDC37* falls into, and I sought to evaluate co-expression and specific modules relatedness of *CDC37* with: i) genes associated to PINK1-dependent mitophagy (i.e., *PINK1*); ii) the major FTLD Mendelian genes (i.e., *MAPT*, *C9ORF72*, *GRN*). Garcia-Ruiz et al.¹⁷⁸ created a suite of R packages to source co-expression networks and modules from gtxv6 and UKBEC databases and assess their relationship. These packages were used to evaluate the relatedness of the different modules between each other within each network: this knowledge was used to infer module proximity, correlation, similarity and thus potential co-expression patterns among their genes.

2.3. Molecular biology

2.3.1. Materials

Reverse transcription: Superscript IV reverse transcriptase, dNTPs, random hexamers,

1M DTT and RNaseOUT RNase inhibitor were all purchased from Invitrogen.

Quantitative PCR: Fast SYBR Green mastermix and probes were purchased from Applied Biosystems.

2.3.2. Quantitative PCR

2.3.2.1. RNA extraction

Cells were immediately lysed for RNA extraction by resuspension in lysis buffer provided with the Monarch® Total RNA Miniprep kit (New England Biolabs, #T2010). Genomic DNA was removed by first spinning through a gDNA binding column and optional on-column DNase treatment. Lysates were then homogenised and RNA was extracted using the using Monarch RNA Purification Columns according to the manufacturer's protocol.

2.3.2.2. Reverse transcription

For production of cDNA, 500 ng extracted RNA was reversed transcribed using SuperScript™ IV First-Strand Synthesis System (ThermoFisher, #18091050), according to the manufacturer's protocol. Briefly, RNA was combined with random hexamer primers (final concentration 2.5 µM) and dNTP mix (final concentration 0.5 mM) in an initial volume of 7 µl and incubated at 65°C for 5 min. The mixture was then incubated on ice for 1 min before addition of 2 µl 5x SSIV buffer, 0.5 µl 100 mM DTT, 0.5 µl RNase OUT and 0.125 µl 200 U SSIV reverse transcriptase enzyme, bringing the total reaction volume to 10 µl. The reaction was incubated at 25°C for 10 min, then at 50°C for 40 min and finally the enzyme was inactivated at 80°C for 10 min.

2.3.2.3. Quantitative real-time PCR

mRNA expression of *STUB1*, *CUL2*, *CDC37*, *PINK1* and *GRN* was measured using SYBR quantitative PCR, performed on a QuantStudio™ 7 Flex system (Applied Biosystems™). 10 ng cDNA were combined with 2x SYBR Green PCR Master Mix (ThermoFisher, #4309155) and 5 µM of forward + reverse primers mix in a total reaction volume of 10 µl in 0.1-ml MicroAmp™ Fast Optical 96-well plates. Thermal cycling conditions are specified in the following **Table 2.1**.

Step	Temperature (°C)	Time (sec)	Cycles (#)
DNA polymerase activation	95	20	HOLD
Denature	95	3	40
Anneal/Extend	60	30	

Table 2.1. Quantitative PCR thermal cycling protocol (SYBR).

Relative quantitation of the target genes was performed against an internal standard to measure gene expression. Data were analysed using the $2^{-\Delta\Delta CT}$ method¹⁸¹, using *RPL18A* as an internal standard.

2.3.3. Cell culture

2.3.3.1. Materials

Cell lines: The cell lines used, their origin and associated mutations are listed in **Table 2.2**.

Line name	Cell type	Disease	Mutation	Origin
SH-SY5Y POE	Human neuroblastoma	N/A	N/A	HPF lab
H4	Human astrocytoma	N/A	N/A	HPF lab
mt-keima SH-SY5Y POE	Human neuroblastoma	N/A	N/A	Luft lab
CRISPR isogenic series	iPSC-derived cortical neurons	FTLD	<i>GRN</i> (R493X)	iNDI

Table 2.2. Characterisation of cell lines used. iNDI, iPSC Neurodegenerative Disease Initiative.

Media for cell lines: For culture of SH-SY5Y Parkin over-expressing (POE), mt-Keima SH-SY5Y POE and H4 cell lines, Dulbecco's modified Eagle medium (DMEM, Gibco, 11995-065) containing high glucose, sodium pyruvate and 2 mM L-glutamine was purchased from Invitrogen and supplemented with 10% (v/v) foetal bovine serum (FBS). 0.25% trypsin-EDTA and sterile phosphate buffered saline (PBS) were purchased from Invitrogen. Right before live-cell imaging, mt-Keima SH-SY5Y POE cells were cultured with phenol red-free DMEM supplemented with 10% FBS. Patient iPSC lines were

cultured in Essential 8 media (Thermofisher). Tris-buffered saline (TBS; Corning) was used to perform washes on iPSC-derived cortical neurons.

Transfection reagents: DharmaFECT1 transfection reagent and siRNAs were purchased as pre-designed siGENOME SMARTpools from Dharmacon (Horizon Discovery): non-targeting (SCR [D-001206-13]), PINK1 (M-004030-02), CDC37 (M-003231-01), GRN (M-009285-02), STUB1(M-007201-02), CUL2 (M-007277-00). siRNAs were diluted in sterile RNase free water to a stock concentration of 20 μ M.

Trypan blue: 0.4% trypan blue solution was purchased from Sigma-Aldrich.

Oligomycin/Antimycin A (O/A): Oligomycin (mitochondrial complex V inhibitor) was purchased from Cayman Chemicals (11341) and from Sigma-Aldrich (O4876), and antimycin A (mitochondrial complex III inhibitor) was purchased from Sigma-Aldrich (A8674).

2.3.3.2. Culture of immortalised cell lines

Neuroblastoma and astrocytoma cell lines were cultured in the appropriate media (see section 2.3.3.1) at 37°C in 5% CO₂/95% air. Cells were passaged every 3-5 days (usually 1:10 for both SH-SY5Ys and H4s). Cells were used at passage numbers no higher than 20. iPSC lines were maintained in Essential 8 media on plates that had been coated overnight with 1:100 geltrex (Thermofisher) in DMEM/F12 at 37°C. The cells were passaged every 3-5 days with 0.5mM EDTA (ThermoFisher) at a split ratio of 1:6 and maintained in small colonies to prevent spontaneous differentiation. Cells were fed daily with Essential 8 media.

2.3.3.3. Differentiation of iPSC into cortical neurons

An isogenic CRISPR iPSC series (control, heterozygous and homozygous *GRN* R493X mutant lines) from the human iPSC Neurodegenerative Disease Initiative (iNDI) was available to the Wray lab. Shi et al.'s¹⁸² dual SMAD (small mothers against decapentaplegic) inhibition protocol was followed to induce neuronal differentiation (**Figure 2.3**).

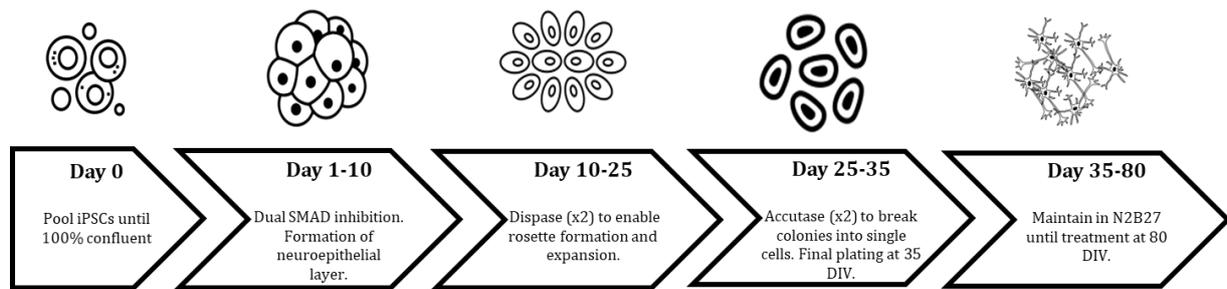


Figure 2.3. Cortical neurons differentiation protocol. DIV, days *in vitro*.

Briefly, iPSCs were pooled into 1:100 geltrex (Thermofisher) coated wells until they reached 100% confluency. They were then fed daily with neural induction media for 10-11 days. The induction media is a 1:1 mix of N2 and B27 (DMEM/F-12 GlutaMAX, 1× N2 supplement, 5 $\mu\text{g ml}^{-1}$ insulin (Sigma-Aldrich), 1mM L-glutamine, 100 μM nonessential amino acids, 100 μM 2-mercaptoethanol, 50 U ml^{-1} penicillin and 50 mg ml^{-1} streptomycin. B27: Neurobasal, 1× B27, 200 mM l-glutamine, 50 U ml^{-1} penicillin and 50 mg ml^{-1} streptomycin [all from Thermofisher unless stated otherwise]), supplemented with 1 μM Dorsomorphin (Tocris) and 10 μM SB431543 (Tocris), dual SMAD inhibitors.

200 μl of dispase (10mg/ml) (Invitrogen) was added to the media to lift the neuroepithelial layer, which was then broken into clumps with approximately 300-500 cells to allow rosettes to form and enable neurogenesis of cortical progenitor cells. From this point on, the cells were maintained in N2B27 on laminin (Sigma-Aldrich) coated 12 well plates (1:50 laminin for the first dispase, then 1:100 laminin for other splits). The cells were passaged with dispase 2-4 times at approximately 10, 18 and 24 days *in vitro* (DIV), to expand the cultures. They were then passaged twice with accutase (Invitrogen) to dissociate the cultures into single cells, including the final passage at day 35 when they were plated on poly-L-ornithine/ 1:100 laminin coated plates. Finally, cortical neurons were maintained in N2B27 until treatment at 80 DIV.

2.3.3.4. Characterisation of cortical neurons derived from PSCs

At day 30 after neural induction, PSC-derived cortical neurons were assessed by immunofluorescence staining to characterise their neuronal fate, as per Shi et al.'s protocol¹⁸². Briefly, cells were fixed in culture in 4% (wt/vol) PFA for 10 min. Following removal of PFA, cells were washed three times for 5 min each with TBS, permeabilised with three 5-min washes with 0.3% (vol/vol) Triton X-100 in TBS and then blocked for 1h with 5% (vol/vol) donkey serum in 0.3% (vol/vol) Triton X-100-TBS. Cortical neurons were then stained with primary antibodies to MAP2, Tuj1, FOXg1 (Abcam, ab18259), PAX6 (Covance, PRB-278P) and OTX2 (Millipore, AB9566) diluted in blocking solution at 4 °C overnight. Visualisation of the neuron-specific markers by confocal microscopy was used to determine their neuronal fate.

2.3.3.5. Freezing and thawing

For long-term storage of immortalised cell lines, the cells were detached by incubation in 0.025% trypsin, resuspended in warm medium and pelleted by centrifugation at 240 g for 5 min. The cell pellet was resuspended in 90% FBS with 10% sterile DMSO and transferred to cryovials for freezing at a rate of -1°C/min using a Nalgene Mr Frosty at -80°C. After 24 hr, the cells were transferred to liquid nitrogen for storage.

To resuscitate frozen cells, cells were thawed rapidly and resuspended in 5 ml warm growth medium. Cells were again pelleted by centrifugation at 400 g for 5 min, resuspended in fresh growth medium and transferred to a cell culture flask.

2.3.3.6. Cell counting

Where necessary, cells were counted using a Neubauer haemocytometer. Dead cells were detected using 0.4% trypan blue solution at a 1:1 ratio with the cell suspension. Trypan blue is excluded from intact cells, therefore any stained cell was disregarded.

2.3.3.7. siRNA-mediated gene silencing

Cells were reverse-transfected (meaning transfected when still in suspension as opposed to forward-transfected when already plated) with siRNA at 80-90% confluency using

DharmaFECT according to the manufacturer's protocol. Briefly, for one well of a 6 well plate 10 µl of 20 µM siRNA and 6 µl DharmaFECT were each diluted in a total of 200 µl serum free media. The two solutions were incubated separately for 5 min at room temperature, then combined and incubated for a further 20 min at room temperature during which time the cells were trypsinised and counted. After 20 min, 1.6 ml of cell suspension was added dropwise to the entire 400 µl siRNA/DharmaFECT solution to get a final concentration of 0.55×10^6 /ml. For transfection in a 24 well plate, the protocol was as described above except all volumes were divided by four.

2.3.3.8. Mitophagy induction

For all western blotting and immunocytochemistry experiments, after 72h siRNA transfection cells were treated with 1 µM oligomycin/antimycin (O/A) to depolarise mitochondria.

2.3.4. Protein biochemistry

2.3.4.1. Materials

Buffers:

- Whole cell lysis buffer: 50 mM Tris (pH 7.4), 0.1 mM EGTA, 0.27 M sucrose, 1% Triton-X100, 1mM EDTA in ddH₂O. Protease and phosphatase inhibitor cocktail (Roche) were added immediately before cell lysis to get 1X concentration.
- Mitochondrial fractionation buffer: 10 mM Tris base, 0.25 M sucrose, 1 mM EDTA (pH 7.4) in ddH₂O. Protease and phosphatase inhibitor cocktail (1/2 protease tablet and 1 phosphatase tablet for 50 ml of buffer) (Roche) were added immediately before cell lysis.
- Sample loading buffer: NuPAGE 4x LDS Sample buffer (Invitrogen) was combined with 40 mM Dithiothreitol (DTT) to get a final concentration of 1 mM. which was then diluted 1:4 to get 1X LDS in 10mM DTT.
- Gel running buffer: NuPAGE MOPS SDS (Invitrogen).
- Transfer buffer: Tris-glycine (National Diagnostics, Georgia, USA) plus 20% (v/v) methanol.

- Hydrophilic polyvinylidene fluoride (PVDF) membrane (Sigma-Aldrich).
- PBST: 1x phosphate buffered solution (PBS) was made from tablets (Invitrogen) dissolved in deionised ultrapure water with 0.1% (v/v) Tween-20 detergent (Sigma-Aldrich).
- Milk: Marvel skimmed milk powder was dissolved in PBST to the appropriate percentage.

Antibodies: A list of primary and secondary antibodies used for immunoblotting is shown in **Table 2.3** overleaf. Primary antibodies were prepared at the specified dilution in 1-3% (w/v) milk/PBST whilst secondary antibodies were diluted in PBST.

Name	Species	Dilution	Cat. No.	Supplier
GAPDH	Mouse	1:10,000	ab8245	Abcam
Mfn2	Rabbit	1:1000	9482	Cell Signaling
Pink1	Rabbit	1:1000	-	Takeda
TOM20	Rabbit	1:5000	sc-17764	Santa Cruz
TIM23	Mouse	1:1000	611223	BD Biosciences
Phospho-ubiquitin (Ser65)	Rabbit	1:1000	37642	Cell Signaling
Stub1/Chip	Rabbit	1:1000	ab2917	Abcam
Cul2	Rabbit	1:1000	A302-476A	ThermoFisher
Cdc37	Mouse	1:1000	sc-17758	Santa Cruz
TBK1	Rabbit	1:1000	3504	Cell Signaling
Phospho-TBK1 (Ser172)	Rabbit	1:1000	5483	Cell Signaling
IRDye 800CW anti rabbit	Donkey	1:20,000	925-32213	LI-COR Biosciences
IRDye 680LT anti mouse	Donkey	1:20,000	925-68022	LI-COR Biosciences

Table 2.3. List of antibodies for immunoblotting.

Consumables: Bio-Rad DC (detergent compatible) protein assay kit was purchased from Bio-Rad (CA, USA). NuPAGE 4-12% Bis-Tris protein gels, NuPAGE LDS sample buffer was purchased from Invitrogen and the Chameleon Duo Pre-stained protein ladder was purchased from LICOR. The iBlot 2 Dry Blotting System and the pre-packaged ready-to-use transfer stacks were purchased from Thermo Fisher Scientific.

2.3.4.2. Harvesting and lysing of cells

Cells were first washed with PBS, then harvested by adding whole cell lysis buffer with protease inhibitors on ice and scraping into an eppendorf. Lysates were vortexed and incubated on ice for 20 min 3 consecutive times to ensure complete lysis. Insoluble cellular components were removed by centrifugation at 16,000 g for 20 min at 4°C, then the supernatant was transferred to a fresh eppendorf and stored at -20°C. Protein concentration was measured using the Bio-Rad DC protein assay kit, using protein standards prepared by diluting BSA to concentrations of 1.0, 2.0, 3.0, 4.0 and 5.0 mg/ml in lysis buffer.

2.3.4.3. Mitochondrial fractionation for biochemistry

To isolate mitochondria from cultured cells, the media was aspirated from the cells and the monolayer rinsed once with PBS. 100 µl of mitochondrial fractionation buffer was added directly to the cells, they were scraped into an eppendorf and the well was washed with another 100 µl of mitochondrial fractionation buffer to collect cell residues. The cells pelleted at 1500 g for 20 min to remove intact cells and debris. The supernatant was then centrifuged at 12,500 g for 20 min to separate the mitochondrial pellet from the cytosolic fraction (the supernatant). The cytosolic fraction was removed, quantified and diluted directly in sample buffer for western blot analysis, while the mitochondrial pellet was rinsed twice with ice-cold mitochondrial fractionation buffer, resuspended in 1X LDS 10mM DTT and sonicated before loading.

2.3.4.4. Protein electrophoresis and western blotting

Samples were diluted in 4x sample buffer containing DTT, boiled at 100°C for 5 min and loaded on a precast 4-12% NuPAGE polyacrylamide gel. Bands were separated by electrophoresis at 180 V, then transferred to a PVDF microporous membrane for 7 min at 80 V. The membrane was blocked by incubation in PBST/5% milk for 30 min, then incubated in primary antibody either overnight at 4°C or for 2 hr at room temperature. The membrane was washed three times in PBST, incubated with the appropriate secondary antibody for 1 hr at room temperature, and washed again two times in PBST and one last time in PBS before imaging. Imaging of the membrane and quantification analysis of the bands were carried out by using the LICOR western blot imager and LICOR

Image Studio Lite 5.2 software respectively.

2.3.5. Immunocytochemistry

2.3.5.1. Materials

Antibodies: Stains, primary and secondary antibodies used for immunocytochemistry are listed in **Table 2.4** below.

Name	Species	Dilution	Cat. No.	Supplier
Phospho-ubiquitin (Ser65)	Rabbit	1:1000	37642	Cell Signaling
TOM20	Mouse	1:1000	ab56783	Abcam
TOM20	Rabbit	1:1000	sc- 17764	Santa Cruz
CDC37	Mouse	1:1000	sc-17758	Santa Cruz
p62	Mouse	1:1000	ab155686	Abcam
Hoechst 33342	-	1:2000	875756-97-1	Sigma Aldrich
AlexaFluor 488 anti rabbit	Goat	1:2000	A11008	ThermoFisher
AlexaFluor 568 anti mouse	Goat	1:2000	A11004	ThermoFisher

Table 2.4. Primary antibodies used for immunofluorescence.

Reagents: 4% paraformaldehyde (PFA) solution was made with 37% formaldehyde in 1X PBS buffer.

Consumables: FBS and PBS were made as described previously (section 2.3.3.1). Triton X-100 and Hoechst 33342 were purchased from Sigma-Aldrich.

2.3.5.2. Cell staining

POE SHSY5Y cells for immunofluorescent analysis were cultured in a 96-well plate, transfected as previously described (section 2.3.3) and fixed for 15 min in 4% (w/v) PFA/PBS. Unless otherwise stated, for all of the incubation steps the 96-well plate was placed in an orbital shaker (40 rpm) and covered in tinfoil to protect it from light and prevent fading of the signal. Cells were permeabilised and blocked with 0.5% (v/v) triton X-100/PBS solution in 10% (v/v) FBS/PBS for 30 min to prevent non-specific antibody binding. Primary antibodies were diluted in 10% (v/v) FBS/PBS and pipetted into each well to get a 40 µl final volume and incubated for 2h at room temperature. Cells were

rinsed three times in PBS, then incubated for 1h at room temperature with fluorescently labelled secondary antibodies and Hoechst in 10% (v/v) FBS/PBS. Excess antibody was removed by washing the wells three times with PBS; Finally, 70 μ l of PBS was added to all of the wells and the plate was kept at 4°C until imaging.

2.3.5.3. Visualisation of protein relocation and image analysis

Stained 96-well plates were imaged using the Opera Phenix (Perkin Elmer): 5x fields of view and 4x 1 μ m Z-planes were acquired per well, using the 40X water immersion objective. Excitation/emission wavelengths are shown in **Table 2.5** below.

Dye	Excitation (nm)	Emission (nm)
Hoechst 33342	350	510-540
Alexa Fluor 488 (anti-rabbit)	488	496-573
Alexa Fluor 568 (anti-mouse)	568	573-630

Table 2.5. Excitation and emission wavelength for immunocytochemistry.

Image analysis was performed in an automated way using Columbus 2.8 analysis system (Perkin Elmer) to measure the integrated intensity of pUb(Ser65) within the whole cell. Briefly, images were loaded in Columbus as maximum projections and the Hoechst 33342 channel was used to find the nuclei and exclude those at the image border. The cytoplasm was found using Hoechst 33342 + Alexa Fluor 568 tagging TOM20, whereas pUb (Ser65) spots were identified within the whole cell using Alexa Fluor 488 and their intensity measured. The mean integrated pUb (Ser65) intensity was calculated as follows:

$$\text{Total cell area covered by PuB spots} \quad \times \quad \text{Corrected PuB spots intensity}$$

The outputs of the analysis were i) the number of selected nuclei and ii) the mean integrated pUb (Ser65) spot intensity.

2.3.6. Live cell imaging: mt-Keima

2.3.6.1. Materials

Reagents:

- Hoechst 33342 was used to stain cells nuclei at a concentration of 1:2000.
- O/A was diluted in phenol-free DMEM to obtain a concentration of 1 μ M to depolarise mitochondria.

2.3.6.2. Cells treatment and imaging

Keima-expressing SHSY5Y cells were plated in a 96-well plate. After 72h siRNA transfection, cells were treated with 1 μ M O/A and were imaged using Opera Phenix with a 63x water objective. The Opera Phenix was set so that CO₂ concentration was 5% at and the temperature was 37°C at the start of imaging. Pictures were taken at 5 different timepoints (0, 2, 6, 9, 12h).

2.3.7. Statistical analysis

Unless otherwise stated, experiments were performed at least three times ($n \geq 3$) and data is presented as mean \pm standard deviation (StD). A two-way ANOVA was used to measure differences between groups for all the wet-lab experiments except for qPCR experiments, where an ordinary one-way ANOVA was used. For all statistical tests, multiplicity adjusted p-values were computed for each comparison using Dunnett's multiple comparisons test. Scientific significance was assumed at p-value < 0.05 .

Chapter 3

3. Results

3.1. *C9ORF72*, AAO, and genetic ancestry discriminate behavioural from language variant FTLD syndromes

3.1.1. Introduction

3.1.2. *C9ORF72* repeat expansions

It has become clear over the past several decades that expansions of simple intronic sequence repeats, such as microsatellites, correlate with the onset of many human diseases in individuals of European descent¹⁸³. Trinucleotide repeat expansions were the first to be discovered and are the most frequent cause of many common genetic disorders: specifically, to date, tri-, tetra-, penta-, hexa-, and even dodeca-nucleotide repeat expansions have been identified as the cause of human diseases such as FTLD/ALS, Huntington disease, Friedreich ataxia and fragile X syndrome.

Repeat expansion diseases arise from normally existing polymorphic repeats and, despite their sequence and size heterogeneity, they show multiple phenotypic similarities.

The mutation related to FTLD/ALS onset is a hexanucleotide (GGGGCC)ⁿ expansion whose pathogenicity can vary depending on a broad spectrum of determining factors, such as repeat length, its position within the gene, the age of the patient and epigenetic modulation⁹⁷. It has been established that a convenient cut-off to discriminate between 'normal' repeat alleles and pathogenic expanded repeats could be 30 repeats (using the repeat-prime PCR method), being typically unaffected the individuals carrying fewer than 30 repeats^{30,177}. In fact, patients with ALS and/or FTLD pathologies have shown to harbour several hundreds to thousands of repeats.

Repeat expansions in *C9ORF72*¹⁸⁴ have been formerly reported to occur in ~25%^{9,14,35,62,185} of familial and ~6%⁶³ of sporadic FTLD cases (i.e., individuals with no clear familial history and/or genetic aetiology²⁹).

FTLD patients with abnormal *C9ORF72* repeat expansions exhibit considerable phenotypical and pathological heterogeneity, therefore suggesting presence of additional (genetic and environmental) modifiers⁵⁵. In spite of contradictory studies reporting either positive or inverse correlation between repeat length and age at onset (AAO),

C9ORF72 expansions have been suggested to act as a genetic modifier of AAO¹⁸⁶⁻¹⁹⁰.

I here analysed 1396 FTLD cases gathered through the IFGC (International FTLD-Genetics Consortium; <https://ifgcsite.wordpress.com/>) phase-III initiative, aiming at (i) characterising *C9ORF72* expansions in relation to genetic ancestry and AAO, and (ii) assessing the usefulness of these parameters in discriminating the behavioural from the language variant syndrome.

3.2. Results

3.2.1. *C9ORF72* expansions frequency and syndromes

I assessed the frequency of pathogenic expansions in the entire cohort and across the different syndromes in the *group 0* cases (**Figure 2.1**). Four percent of all cases (56/1396 [4%]) carried pathogenic expansions. These were most frequent in FTLD-MNDs (12/101 [11.9%]) followed by bvFTLDs (40/800 [5%]) and PPAs (4/495 [0.8%]). The higher prevalence of pathogenic expansions in bvFTLDs vs. PPAs was statistically significant (Fisher's Exact test: $p = 2.17 \times 10^{-5}$; OR = 6.4; 95% CI: 2.31 – 24.99 **Table 3.1**).

Cohort	n of cases	Expansion carriers	Frequency
bvFTLD	800	40	5%*
PPA	495	4	0.8%*
FTLD-MND	101	12	11.9%
Total	1396	56	4%

Table 3.1. Frequency of expansion carriers in the entire cohort and by syndrome. Summary of expansions carriers frequency in the entire cohort (n = 1396) and across syndromes. The higher prevalence of expansion carriers in bvFTLD vs. the PPA is statistically significant: *Fisher's exact test performed to statistically evaluate the difference between the occurrence of pathogenic-expansions in the bvFTLD vs. the PPA syndromes: $p = 2.17 \times 10^{-5}$; odds ratio (OR) = 6.4; 95% confidence interval (CI): 2.31-24.99. Adapted from '*C9ORF72*, age at onset, and ancestry help discriminate behavioural from language variants in FTLD cohorts' by B. Costa, 2020, *Neurology*. Copyright 2022 American Academy of Neurology. Adapted with permission.

I further explored this finding in the *group 3* cases using logistic regression to assess association between expansion length (represented by 4 repeat counts [rc] factor levels – ‘short’, ‘intermediate’ and ‘long’ expansions, tested against ‘no’ expansions) and syndromes (bvFTLD vs. PPA). Expansion length discriminated bvFTLD from PPA with a trend that was significant in the ‘intermediate’ ($p = 4.7 \times 10^{-2}$; OR = 1.6; CI: 0.0061 [2.5%] – 0.94 [97.5%]) and ‘long’ ($p = 1.9 \times 10^{-3}$; OR = 7.2; CI: 0.86 [2.5%] – 3.45 [97.5%]) rc ranges (with a ~90% probability of a bvFTLD diagnosis supported by the latter; **Supplementary Table 3 [Supplementary File 1]**; Supplementary files are available at <https://rdr.ucl.ac.uk/articles/dataset/Supplementary-Files-Neurology-pdf/12418157>).

3.2.2. C9ORF72 expansions (and repeat counts [rc]) and genetic ancestry

I performed PCA (PC1 vs. PC2, **Supplementary Figure 2A [Supplementary File 1]**; PC1 vs. PC3, **Supplementary Figures 2B [Supplementary File 1]**) to cluster the *group 1* cases based on their genetic make-up. There were 2 major clusters: *cluster-1* (‘Mediterranean’) included most of the cases (439/500 [87.8%]) recruited from Southern European sites (Italy and Spain); *cluster-2* (‘Nordic’) included most of the cases (627/795 [78.8%]) recruited from Central and Northern European sites (Belgium, The Netherlands, Germany, UK, Norway and Sweden). Samples recruited from Eastern European (Slovenia) and North American sites distributed across both clusters – although with a higher prevalence within *cluster-2* (167/795 [21%]) vs. *cluster-1* (42/500 [8.4%]).

I observed a significantly higher prevalence of pathogenic expansions in the ‘Nordic’ (35/795 [4.4%]) vs. the ‘Mediterranean’ (9/500 [1.8%]) cluster (Fisher’s Exact test: $p = 1.1 \times 10^{-2}$; OR = 2.5; CI: 1.17 – 5.99 **Table 3.2**).

Genetic ancestry	n of cases	Expansion carriers	Frequency
Nordic	795	35	4.4%*
Mediterranean	500	9	1.8%*

Table 3.2. Frequency of expansion carriers in the ‘Nordic’ and ‘Mediterranean’ clusters. The higher prevalence of expansion carriers in the ‘Nordic’ vs. the ‘Mediterranean’ cluster is statistically significant: *Fisher’s exact test: $p = 1.1 \times 10^{-2}$; OR = 2.5; 95% CI: 1.17-5.99. Adapted from ‘*C9ORF72*, age at onset, and ancestry help discriminate behavioural from language variants in FTLD cohorts’ by B. Costa, 2020, *Neurology*. Copyright 2022 American Academy of Neurology. Adapted with permission.

I further evaluated this finding in the *group 3* cases using logistic regression to assess association between expansion length (see above) and genetic ancestry. Expansion length discriminated the ‘Nordic’ from ‘Mediterranean’ cluster with a trend that was significant in the ‘intermediate’ ($p = 9.7 \times 10^{-4}$, OR = 2.2; CI: 0.32 [2.5%] – 1.25 [97.5%]) and ‘long’ ($p = 4.7 \times 10^{-4}$, OR = 9.3; CI: 1.12 [2.5%] – 3.7 [97.5%]) rc ranges (with a ~90% probability of ‘Nordic’ ancestry supported by the latter; **Supplementary Table 4 [Supplementary File 1]**).

Provided differences in syndromes prevalence and distribution across the ‘Nordic’ and ‘Mediterranean’ clusters – bvFTLD (469/795 [59%] vs. 331/500 [66.2%]) and PPA (326/795 [41%] vs. 169/500 [33.8%]), respectively (**Supplementary Table 5 [Supplementary File 1]**) – I analysed the distribution of pathogenic expansions across syndromes and clusters. Stratified Fisher’s Exact test showed significant differences in the distribution of the pathogenic expansions between bvFTLD and PPA in the ‘Nordic’ (but not the ‘Mediterranean’) cluster ($p = 1 \times 10^{-4}$; OR = 7.87; 95% CI: 2.43 – 40.52), and between the ‘Nordic’ and the ‘Mediterranean’ clusters for the bvFTLD (but not PPA) syndrome ($p = 1.9 \times 10^{-2}$; OR = 2.95; 95% CI: 1.31 – 7.52), suggesting that ancestry (‘Nordic’) and syndrome (bvFTLD) are independently associated with pathogenic expansions (**Table 3.3**).

Subtype/Ancestry	Expansion range		Fisher's Exact Test
	pathogenic	non-pathogenic	
bvFTLD			
Mediterranean	8	323	p = 1.9x10 ^{-2*}
Nordic	32	437	
Ancestry/Subtype	Expansion range		Fisher's Exact Test
Mediterranean	pathogenic	non-pathogenic	
bvFTLD	8	323	p = 1
PPA	1	168	
Nordic			
bvFTLD	32	437	p = 1x10 ^{-4#}
PPA	3	323	

Table 3.3. Stratified Fisher's exact tests comparing prevalence of pathogenic expansions across bvFTLD and PPA and the 'Nordic' and 'Mediterranean' clusters. P-values presented in the table are corrected for multiple testing statistics. Prior correction p-values were as follows: * (uncorrected) Fisher's exact test: p = 4.7x10⁻³; OR = 2.95; 95% CI: 1.31-7.52 → significant difference in the prevalence of bvFTLD expansion carriers in the 'Nordic' vs. the 'Mediterranean' cluster; # (uncorrected) Fisher's exact test p = 2.7x10⁻⁵; OR = 7.87; 95% CI: 2.43-40.52 → significant difference in the prevalence of expansion carriers in bvFTLDs vs. PPAs within the Nordic cluster. Adapted from '*C9ORF72*, age at onset, and ancestry help discriminate behavioural from language variants in FTL D cohorts' by B. Costa, 2020, *Neurology*. Copyright 2022 American Academy of Neurology. Adapted with permission.

3.2.3. C9ORF72 repeat expansions (and counts [rc]) and age at onset (AAO)

I assessed AAO in the *group 2* cases (**Figure 2.1**). Mean AAO was significantly different between the bvFTLD (61.7) and PPA (64) syndromes (t-test: $p = 1.86 \times 10^{-5}$; CI: -3.34 - -1.25), and the 'Nordic' (61.3) and 'Mediterranean' (64.3) clusters (t-test: $p = 1.16 \times 10^{-7}$; CI: 1.86 - 4.03) (**Supplementary Table 6A and B [Supplementary File 1]; Figure 3.1[A]**); Supplementary files are available at https://rdr.ucl.ac.uk/articles/dataset/Supplementary-Files-Neurology_pdf/12418157).

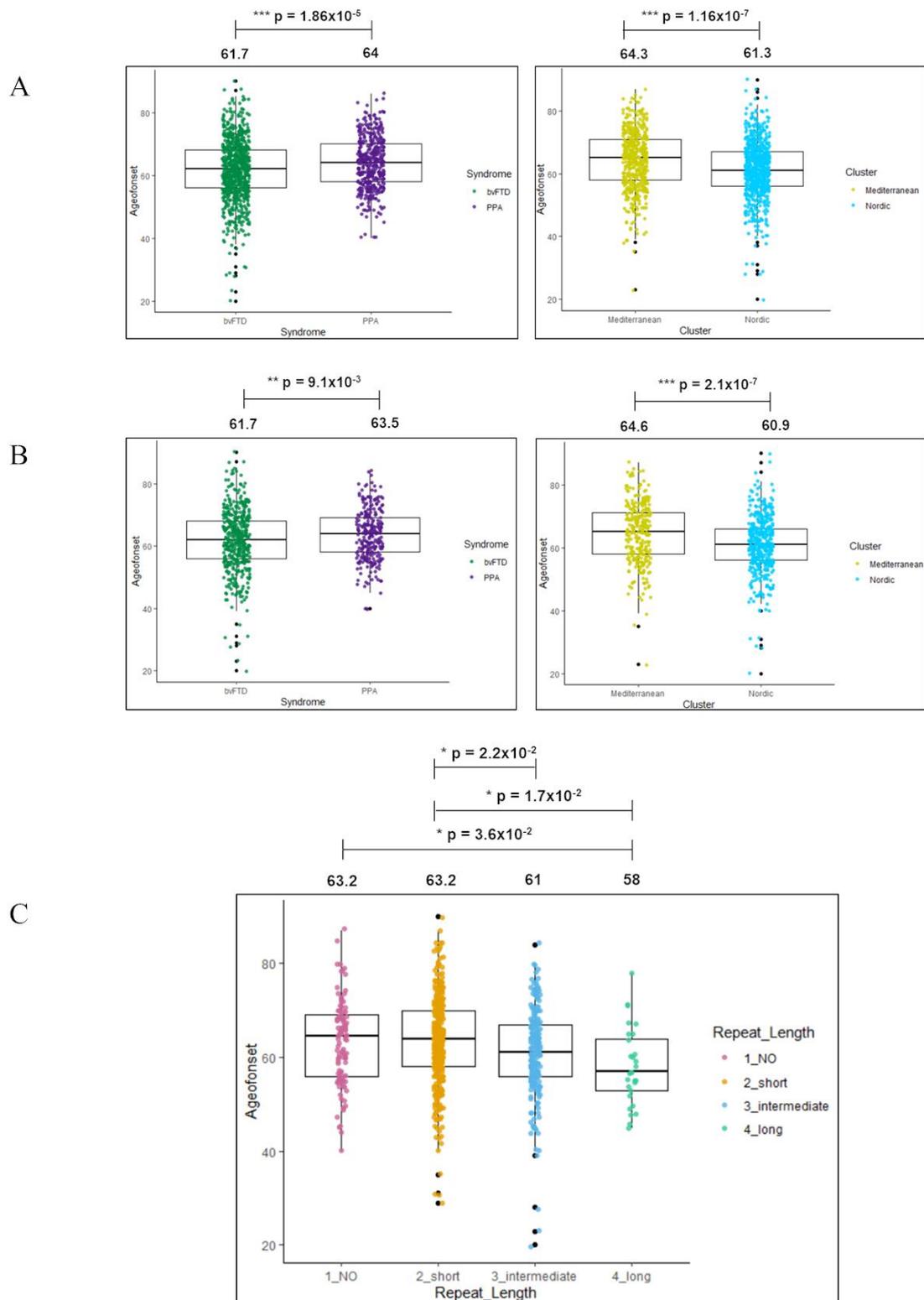


Figure 3.1. Association between AAO and: ancestry; syndrome; expansion length. A. AAO in the group 2 cases. Mean AAO bvFTLD (61.7) and PPA (64) (t-test: $p = 1.86 \times 10^{-5}$; CI: -3.34--1.25); mean AAO 'Nordic' (61.3) and 'Mediterranean' (64.3) clusters (t-test: $p = 1.16 \times 10^{-7}$; CI: 1.86-4.03); B. AAO in the group 3 cases. Mean AAO bvFTLD (61.7) and PPA (63.5) (t-test: $p = 9.1 \times 10^{-3}$; CI: -3.11--0.44), mean AAO 'Nordic' (60.9) and 'Mediterranean' (64.6) (t-test: $p = 2.1 \times 10^{-7}$; CI: 2.32-5.09); C. AAO in the group 3 cases. Mean AAO for both 'no' and 'short' expansions (63.2), for 'intermediate' expansions (61) and for 'long' expansions (58) evaluated via ANOVA test. Adapted from '*C9ORF72*, age at onset, and ancestry help discriminate behavioural from language variants in FTLD cohorts' by B. Costa, 2020, *Neurology*. Copyright 2022 American Academy of Neurology. Adapted with permission.

I then assessed the relationship between pathogenic expansions and AAO *via* logistic regression. First, I identified a significant correlation between a decrease in AAO and presence of pathogenic expansions ($p = 7.7 \times 10^{-4}$; $R^2 = 0.008$; CI: -8.05 [2.5%] – -2.13 [97.5%]). When I included genetic ancestry in the model we observed a significant correlation with a decrease in AAO, no difference in using either cluster ($p = 2.3 \times 10^{-3}$; CI: -7.5 [2.5%] – -1.63 [97.5%] for pathogenic expansions; $p = 2.3 \times 10^{-7}$; CI: -3.9 [2.5%] – -1.77 [97.5%] for cluster; $R^2=0.03$) or PC1 ($p = 2.1 \times 10^{-3}$; CI: -7.5 [2.5%] – -1.66 [97.5%] for pathogenic expansions; $p = 6.4 \times 10^{-7}$; CI: 30.1 [2.5%] – 68.9 [97.5%] for PC1; $R^2=0.028$) as covariate and an almost 4-fold goodness of fit increase (**Supplementary Table 7A, B and C [Supplementary File 1]**). Of note, when comparing the two regression models (with/without genetic ancestry as covariate) through the log-likelihood R^2 ratio test, the difference (between the 2 models) appeared not to be due to chance ($p < 10^{-12}$) (**Supplementary Table 7B and C [Supplementary File 1]**).

I further evaluated the relationship between expansion length (represented by 4 repeat counts [rc] factor levels – ‘short’, ‘intermediate’ and ‘long’ expansions, tested against ‘no’ expansions) and AAO in the *group 3* cases (**Figure 3.1**). First, I independently analysed association between AAO and: i) genetic ancestry – mean AAO 60.9 and 64.6 in the ‘Nordic’ and ‘Mediterranean’ cluster, respectively (t-test: $p = 2.1 \times 10^{-7}$; CI: 2.32 – 5.09; **Supplementary Table 8A [Supplementary File 1]**); ii) syndrome – mean AAO 61.7 and 63.5 in the bvFTLD and PPA syndromes, respectively (t-test: $p = 9.1 \times 10^{-3}$; CI: -3.11 – -0.44; **Supplementary Table 8B [Supplementary File 1]**), and; iii) expansion length – mean AAO 63.2 for both ‘no’ and ‘short’ expansions, 61 for ‘intermediate’ expansions and 58 for ‘long’ expansions (ANOVA: $p = 3.6 \times 10^{-2}$; CI: -10.2 – -0.23 for ‘long’ vs. ‘no’ expansions) (**Supplementary Table 8D [Supplementary File 1]**; **Figure 3.1 B and C**).

I then assessed the relationship between expansion length (see above) and AAO *via* logistic regression. First, I identified a significant correlation between a decrease in AAO and both ‘intermediate’ and ‘long’ expansions ($p = 4 \times 10^{-2}$; CI: -4.36 [2.5%] – -0.96 [97.5%] for ‘intermediate’ and $p = 7 \times 10^{-3}$; CI: -9.05 [2.5%] – -1.43 [97.5%] for ‘long’ expansions; $R^2 = 0.017$) (**Supplementary Table 9A [Supplementary File 1]**). When I included genetic ancestry in the model I observed a significant correlation with a decrease in AAO, no difference in using either cluster ($p = 4.7 \times 10^{-2}$; CI: -7.65 [2.5%] – -0.05 [97.5%] for ‘long’ vs. ‘no’ expansion; $p = 2.38 \times 10^{-6}$; CI: -4.73 [2.5%] – -1.97 [97.5%] for cluster; $R^2 =$

0.045) or PC1 ($p = 5.98 \times 10^{-2}$; CI: -7.5 [2.5%] – 0.14 [97.5%] for ‘long’ vs. ‘no’ expansion; $p = 1.2 \times 10^{-6}$; CI: 39.8 [2.5%] – 92.9 [97.5%] for PC1; $R^2=0.047$) as covariate and an almost 3-fold goodness of fit increase (**Supplementary Table 9A, B and C [Supplementary File 1]**). Of note, when comparing the two regression models (with/without genetic ancestry as covariate) through the log-likelihood R^2 ratio test, the difference (between the 2 models) appeared not to be due to chance ($p < 10^{-12}$) (**Supplementary Table 9B and C [Supplementary File 1]**). These findings were further supported by non-linear mixed-effects model regression using genetic ancestry as random effect covariate (for ‘long’ vs. ‘no’ expansion; see **Supplementary Table 10 [Supplementary File 1]**).

3.2.4. Syndrome prediction

I then sought to build a model to predict syndrome (bvFTLD vs. PPA) and assess its accuracy. I analysed both *groups 2* and *3* cases using expansion status (presence/absence of expansion for *group 2*, and the 4 rc factor levels for *group 3* [see materials and methods]), genetic ancestry (using either ‘cluster’ or ‘PC1’) as binary variables, and AAO as a continuous variable in logistic regression models. I observed an accuracy of ~ 0.64 (*group 2*; **Supplementary Table 11 [Supplementary File 1]**) and ~ 0.62 (*group 3*; **Supplementary Table 12 [Supplementary File 1]**) in predicting bvFTLD, whilst there were no differences in the outcome when using either ‘cluster’ or ‘PC1’ as covariates in both (LOOCV and K-fold) models.

3.3. Discussion and conclusions

This study aimed to characterise *C9ORF72* expansions in relation to genetic ancestry and age at onset (AAO), and to assess the usefulness of these parameters in discriminating the behavioural from the language variant syndrome, in a large pan-European cohort of 1396 FTLD cases.

To the best of my knowledge, the present study is unique in that, prior characterising the expansions, I excluded population-substructure bias using genome-wide genotyping data to cluster the cases on the basis of their genetic ancestry. I performed principal component analysis (PCA) and I identified two distinct clusters including samples with geographical ancestry corresponding to Southern Europe (‘Mediterranean’ cluster) and

Central/Northern Europe ('Nordic' cluster). These analyses showed that patients from the 'Nordic' cluster presented significantly higher frequency of pathogenic *C9ORF72* expansions compared to the 'Mediterranean' cluster.

Globally, I found a pathogenic expansions frequency of about ~4% and that the proportion of expansion carriers was significantly higher in bvFTLDs compared to PPAs. These findings report significant association between pathogenic expansions, a diagnosis of bvFTLD and Central/Northern European ancestry – findings for the most in line with previous studies^{34,35,42,55,62,191-194} – suggesting that *C9ORF72* expansions might serve as valuable genetic marker to define subpopulations of FTLD patients (**Figure 3.2**).

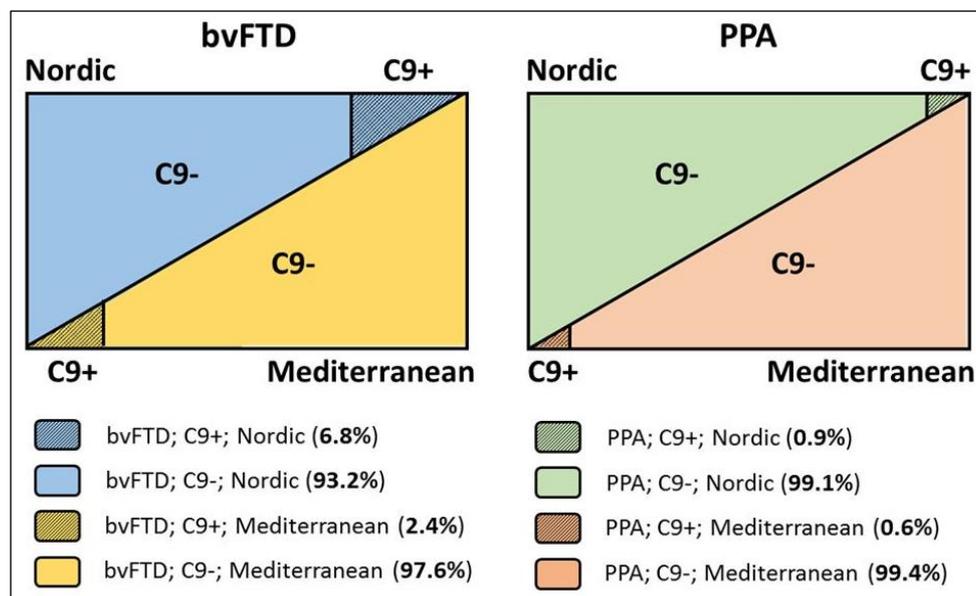


Figure 3.2. Patient subpopulations (bvFTLD and PPA syndromes) based on *C9ORF72* expansions genetic signatures and ancestry. Adapted from '*C9ORF72*, age at onset, and ancestry help discriminate behavioural from language variants in FTLD cohorts' by B. Costa, 2020, *Neurology*. Copyright 2022 American Academy of Neurology. Adapted with permission.

Of note, I observed a trend of association with syndrome (bvFTLD) and genetic ancestry (Central/Northern European) already supported by the 'intermediate' repeat counts ($9 \leq rc \leq 24$) category. This finding seems in line with previous studies suggesting that individuals with 7 to 24 alleles might have an increased risk to become carriers of pathological repeat expansions^{35,188} and may, altogether, be useful information in the context of diagnostics.

Despite some previous conflicting reports of direct (or inverse) correlation between *C9ORF72* expansions and AAO^{186,187,189}, I (as others^{188,190}) found a significant inverse correlation between *C9ORF72* expansion length and AAO.

Additionally, and interestingly, these data also indicates that Central/Northern European genetic ancestry contributes to a decreased AAO (independently from the expansions) possibly implying to a more complex genetic signature (or architecture), and subsequently molecular mechanisms, underlying this very feature. Undoubtedly, disease mechanisms that involve *C9ORF72* expansion length and AAO are complex, thus it is likely that additional factors might further modulate their relationship and effect on the phenotype¹⁸⁴.

I used expansion length, genetic ancestry and AAO in a regression model to discriminate behavioural from language variant subtypes, and I found that such parameters did support a prediction of bvFTLD with 64% accuracy.

These results have several implications. First, provided that significant variation exists in the genetic architecture of the Caucasian population¹⁹⁵, genetic variability characterising and differentiating 'Nordic' vs. 'Mediterranean' subjects (such as in the case of our cohort) might influence predisposition to harbouring longer repeat expansions. In other repeat expansion diseases – e.g. Huntington's disease (HD) or other microsatellite diseases, including myotonic dystrophy and spinocerebellar ataxias¹⁹⁵ – the presence of specific haplogroups in Western European populations occurs with a manifold increase in prevalence of repeats compared to other ethnic groups and populations¹⁹⁶. Second, different genetic risk-architectures underlying different (and possibly genetically more homogeneous) subpopulations of patients may exist within the FTL D population.

In summary, these findings indicate that a significantly higher proportion of FTL D cases, with 'Nordic' rather than 'Mediterranean' genetic ancestry, is likely to develop bvFTLD in presence of 'intermediate' and 'long' (pathogenic) expansions, whilst 'long' (pathogenic) expansions are (almost) negligible in PPAs, regardless of ancestry. Clearly, multiple factors including genetic heterogeneity, epigenetic changes, ethnicity, as well as environmental factors and habits that may exist within and across multicultural cohorts, all together, contribute to disease predisposition, onset and progression^{76,188,197}. These concepts, reinforced by this study, warrant further characterisation of genetic, environmental, and additional clinical measures to fine-tune models able to predict

disease outcome to complement diagnostic criteria, and possibly assist, in the near future, in the identification of informative cohorts for tailored clinical trials and the development of effective personalised therapies.

Chapter 4

4. Results

4.1. Stratification of FTLD-risk candidate genes using WPPINA and WGCNA

4.1.1. Introduction

As described in section 1.4, this thesis proposes a multi-omics network approach integrating genetics, transcriptomics and proteomics to investigate the common pathogenic mechanisms and their associated genes/proteins underlying Mendelian and sporadic forms of FTLD. To this end, protein-protein interaction (PPINs) and gene co-expression (GCNs) networks were used to identify the susceptibility processes conserved across familial and sporadic FTLD cases through the recently developed WPPINA and WGCNA pipelines.

4.1.2. WPPINA and WGCNA

In 2017 Ferrari et al.⁹⁷ presented a novel approach, Weighted Protein-Protein Interaction Network Analysis (WPPINA), to highlight key functional players within relevant biological processes associated with a given trait. This approach is based on the current state of the art of genetics and proteomics, using genes known to be associated with a given trait (i.e. FTLD) to build multiple layers interactomes and subsequently to identify proteins that represent the backbone of biological processes (BP) likely impacted in disease pathogenesis⁹⁷. Most importantly, while such approach generates additional knowledge, it also fosters cross-disciplinary work (e.g., genomics, transcriptomics, proteomics)¹⁹⁸.

Similarly, other system biology methods have been developed to improve our understanding of disease pathogenesis; the possibility of using gene expression profiling (Weighted Gene Co-expression Network Analysis [WGCNA]) for diagnostic and prognostic purposes has generated much excitement in past few years and has proven to be critical to model and characterise biological systems in novel ways¹⁵². WGCNA describes the correlation patterns among genes across microarray or sequencing data and it can be used to find clusters (modules) of highly correlated genes, to relate modules to one another and to a certain trait, to calculate module membership measures¹⁷³.

Correlation networks are valuable tools to facilitate network-based gene screening methods that have been already successfully used to identify candidate biomarkers and therapeutic targets in several biological contexts^{10,173}. In this study I used microarray expression data generated from 134 control individuals (UKBEC) and sequencing (whole exome sequencing [WES], whole genome sequencing [WGS], RNA-sequencing [RNA-seq]) expression data generated from nearly 1000 non-diseased brain tissue (GTEx) to: i) evaluate co-expression patterns of FTLD Mendelian genes in frontal cortex (FCTX); ii) annotate and highlight biological processes potentially implicated in FTLD pathogenesis; iii) identify novel potential risk factors for FTLD. While both these methods have been known for some time, the novelty of this approach lies in the unprecedented integration of WPPINA and WGCNA, providing a framework for interpreting and leveraging big genomic, transcriptomic and proteomic datasets in neurobiology (**Figure 4.1**).

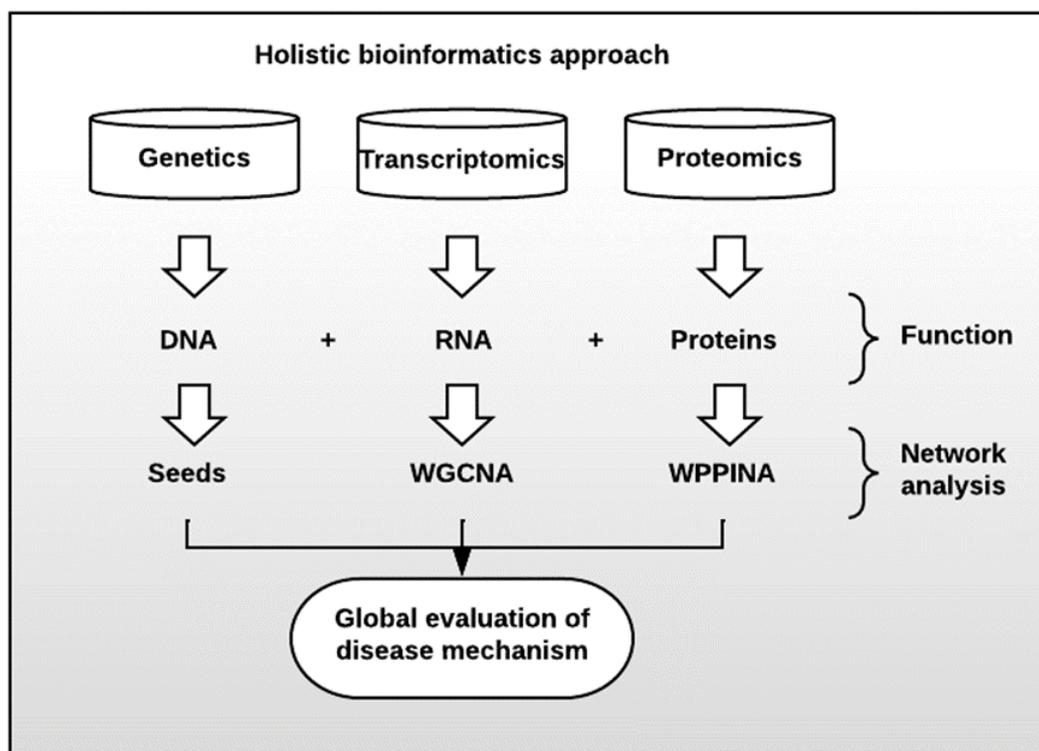


Figure 4.1. Holistic bioinformatics approach employed in this chapter. The scheme exemplifies the holistic approach used to define communal functional features across multiple gene(s)/risk marker(s), as opposed to studying one gene/risk marker at the time, to interpret genetics and subsequently assist and drive functional studies.

The aims of this chapter is to gain a global perspective on FTLD pathogenesis by identifying: i) disease-specific biological processes on the basis of known FTLD

Mendelian and GWAS genes and ii) their associated molecular players, to be carried forward for hypothesis-driven functional validation using *in vitro* cellular models.

4.2. Results

4.2.1. Seeds selection

For the purpose of this study, 18 Mendelian (familial), and GWAS genes associated with the pure FTLN and FTLN-ALS spectrum were selected as ‘seeds’ (**Table 4.1**) to generate two seeded networks (GCN and PPIN). The seeds were selected on the basis of their relevance and degree of characterization as reported in the most recent FTLN literature reviews⁹ and FTLN GWAS¹⁹⁹ at the beginning of the project (04-02-2019).

gene name	frequency (Mendelian%)	pathology	phenotype
<i>C9ORF72</i>	common (7-20%)	FTLN-TDP	pure FTLN
<i>GRN</i>	common (5-11%)	FTLN-TDP	pure FTLN
<i>MAPT</i>	common (2-11%)	FTLN-Tau	pure FTLN
<i>VCP</i>	rare (<1%)	FTLN-TDP	FTLN-ALS
<i>CHMP2B</i>	rare (<1%)	FTLN-UPS	FTLN-ALS
<i>SQSTM1</i>	rare (<1%)	FTLN-TDP	FTLN-ALS
<i>UBQLN2</i>	rare (<1%)	FTLN-TDP	FTLN-ALS
<i>IFT74</i>	rare (<1%)	not known; possibly FTLN-TDP	FTLN-ALS
<i>OPTN</i>	rare (<1%)	FTLN-TDP	FTLN-ALS
<i>CHCHD10</i>	rare (<1%)	FTLN-TDP	FTLN-ALS
<i>DCTN1</i>	rare (<1%)	not understood; possibly FTLN-TDP	FTLN-ALS
<i>FUS</i>	rare (<1%)	FTLN-FUS	FTLN-ALS
<i>TARDBP</i>	rare (<1%)	FTLN-TDP	FTLN-ALS
<i>TBK1</i>	rare (<1%)	not known	FTLN-ALS
<i>TIA1</i>	rare (<1%)	not known	FTLN-ALS
<i>RAB38</i>	GWAS	not known	pure FTLN
<i>HLA-DRA</i>	GWAS	not known	pure FTLN
<i>TMEM106B</i>	GWAS	FTLN-TDP	pure FTLN

Table 4.1. Mendelian (familial) and GWAS FTLN-ALS genes used as seeds for downstream bioinformatics analyses.

Of note, although mutations in the *FUS* and *TARDBP* genes are more commonly found in ALS and ALS-FTLD cases, their protein products FUS and TDP-43 are both known pathological hallmarks of the FTL spectrum, thus these genes are likely to hold functional relevance in the pathogenesis of FTL. For this reason, they were included in the downstream bioinformatics analyses (**Figure 4.2**).

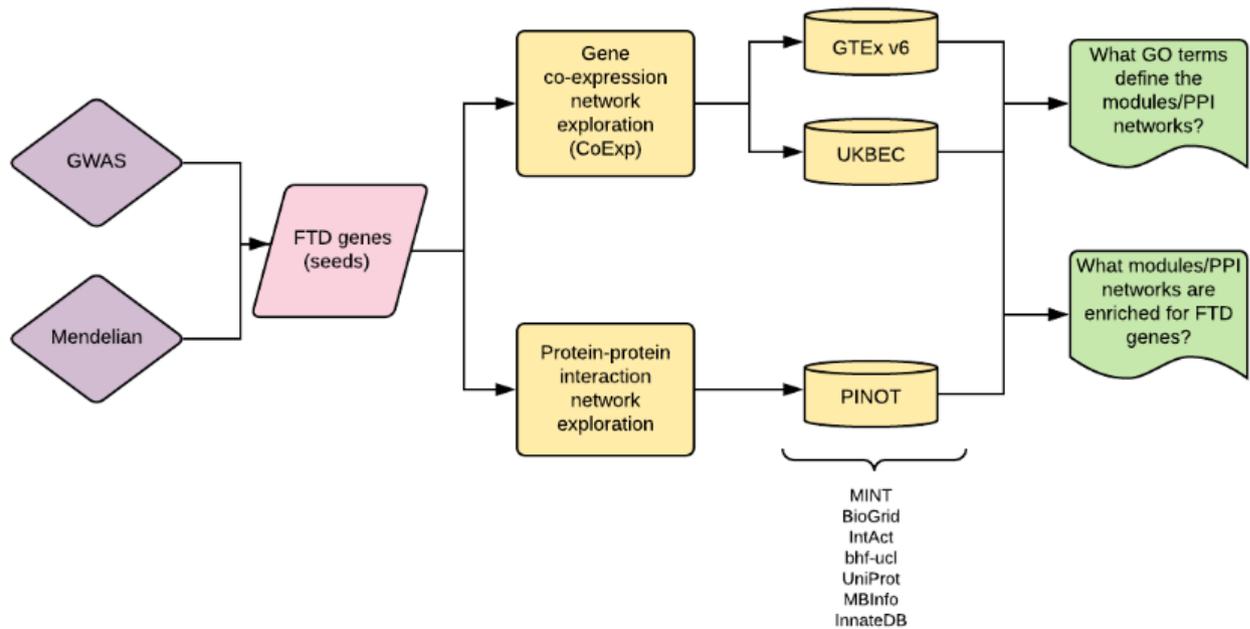


Figure 4.2. Workflow exemplifying the bioinformatics pipeline implemented in the present study.

4.2.2. Construction of the gene co-expression networks (WGCNA)

I performed the WGCNA with a primary focus on frontal cortex (FCTX) to investigate any relationship between FCTX expression networks and FTL Mendelian genes and to identify gene co-expression groups (modules). I used the CoExp Web application¹⁷⁸ to extract the gene co-expression modules (in FCTX) containing the FTL seeds and retrieved 19 distinct modules of interest (**Figure 4.3**).

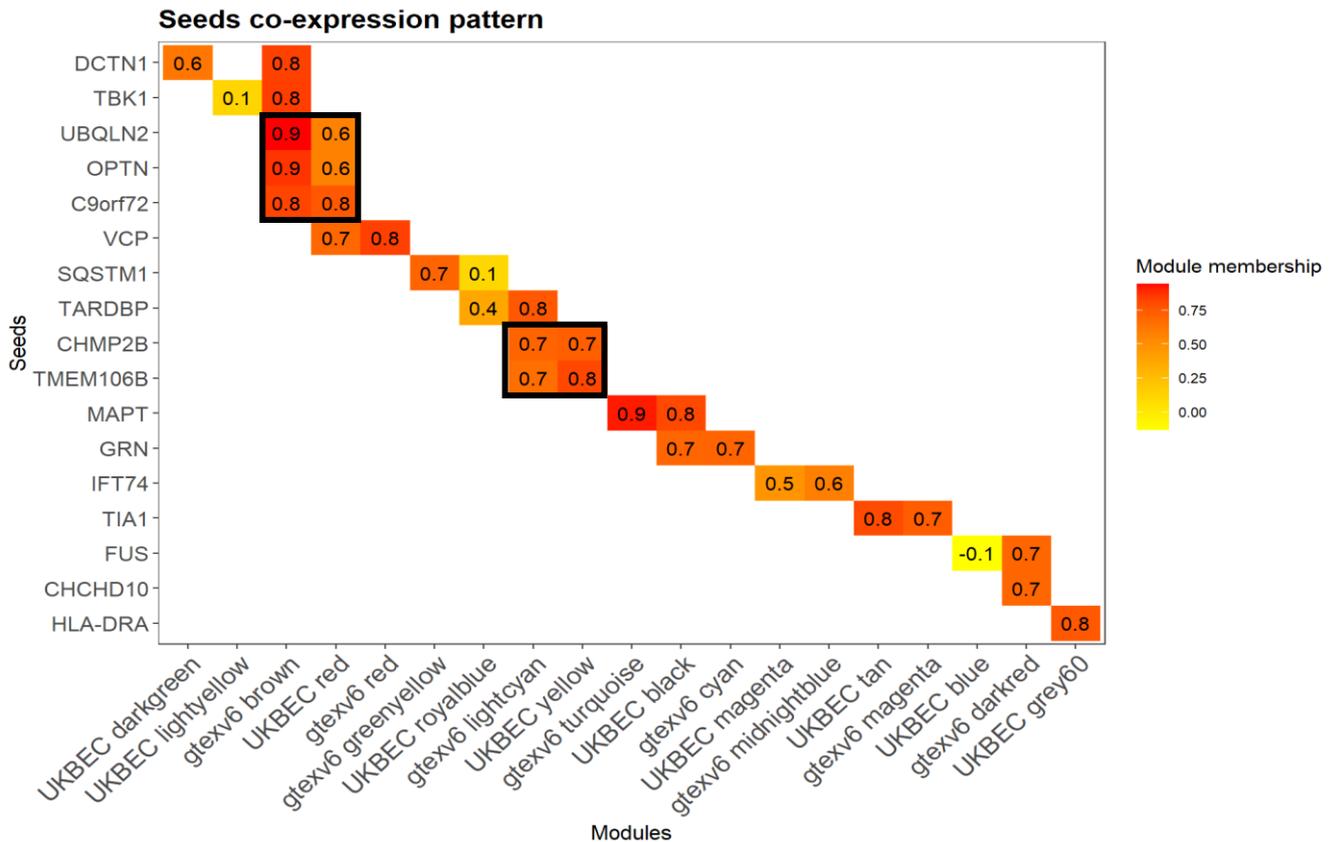


Figure 4.3. Gene co-expression patterns for the FTLD genes (seeds) and their relevance within modules. Genes showing a substantial co-expression overlap across the gtexv6 and UKBEC databases are framed. Module membership (MM) values are displayed.

Of note, co-expression data were sourced from two different transcriptomics databases: GTEx V6 from which 9/19 modules were obtained and UKBEC providing 10/19 modules. The genes that were not found in both datasets were excluded from further analyses. These were *CHCHD10*, which was found in GTEx V6 only ('gtexv6 darkred'), *HLA-DRA*, which was found in UKBEC only ('UKBEC grey60'), and *RAB38*, which was not present in any of the two. As the 'UKBEC grey60' module did not present with the expression of any other FTLD Mendelian gene, it was excluded from further analyses. As a result, I obtained in a list of 18 modules of interest.

I assessed co-expression profiles for the most relevant modules ('relevant' are defined those modules containing one or more FTLD-genes with $MM \geq 0.5$; see Methods section for details)⁹³. Interestingly, I identified the 'conserved' modules as those pairs of modules that, independently from their derivation (from GTEx or UKBEC) showed similar results. In particular 'gtexv6 brown' and 'UKBEC red' were identified as a conserved pair both

containing *UBQLN2* ($MM_{\text{gtexv6}}=0.9441$; $MM_{\text{UKBEC}}=0.5721$) *C9ORF72* ($MM_{\text{gtexv6}}=0.8238$; $MM_{\text{UKBEC}}=0.7562$) and *OPTN* ($MM_{\text{gtexv6}}=0.8715$; $MM_{\text{UKBEC}}=0.5724$). I observed similar co-expression patterns also in the modules 'gtexv6 lightcyan' and 'UKBEC yellow', both of which included *CHMP2B* ($MM_{\text{gtexv6}}=0.7075$; $MM_{\text{UKBEC}}=0.7284$) and *TMEM106B* ($MM_{\text{gtexv6}}=0.6635$; $MM_{\text{UKBEC}}=0.8173$).

4.2.3. Construction of the PPI FTLD network (WPPINA)

The FTLD seeds (**Table 4.1**) were used to download the direct PPI from peer review literature (first layer FTLD interactome). The first layer interactome for each seed was built based on experimentally proven interactions (i.e., biochemical, physical, imaging), whilst interactions inferred from text mining algorithms or *de novo* prediction were excluded. PPIs were downloaded from the PINOT platform and filtered as described in the **Material and methods** section to remove non-validated interactions and ubiquitins.

First layer interactors of each seed were subsequently employed for downloading the second layer of interactions via the PINOT server. The second layer was then filtered and processed as described for the first layer. The complete FTLD-PPIN was therefore composed of the FTLD seeds, their direct interactors (first layer FTLD interactome) and the proteins interacting directly with the first layer (second layer interactome). The complete FTLD-PPIN was composed of 13274 single nodes and 87543 edges (**Figure 4.4**) and represents the state-of-the-art protein interactors gravitating around the FTLD-spectrum genes/proteins.

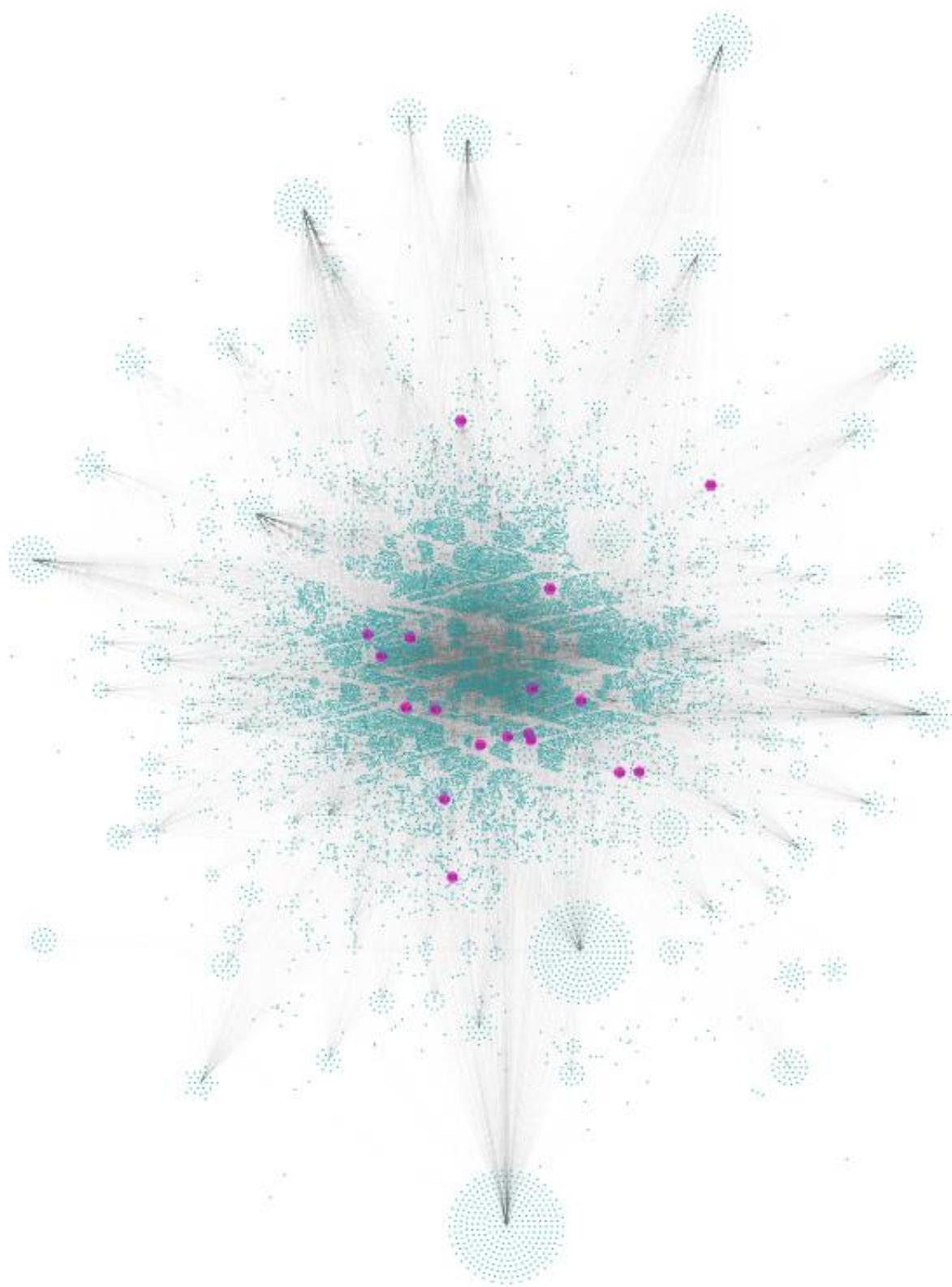


Figure 4.4. Entire-FTLD-PPIN. The network shows the seeds and their first and second layer interactors. pink, seeds; turquoise, first + second layer interactors.

There were no disconnected components in the second layer, showing great cohesion of the complete FTLD-PPIN. Notably, the second layer construction is not directly influenced by the seeds, because the nodes within such layer are extracted on the basis of the first layer interactors rather than the seeds, that indeed no longer act as hubs (this role being undertaken by the first layer genes/proteins). Inter-interactome hubs (IIHs, i.e., ‘hub’ proteins having special topological and functional significance) were extracted from the complete FTLD-PPIN by creating a pivot table to calculate the recurrence of every node in each seed’s interactome (inter-interactome degree) and plotting a degree distribution graph (**Figure 4.5**).

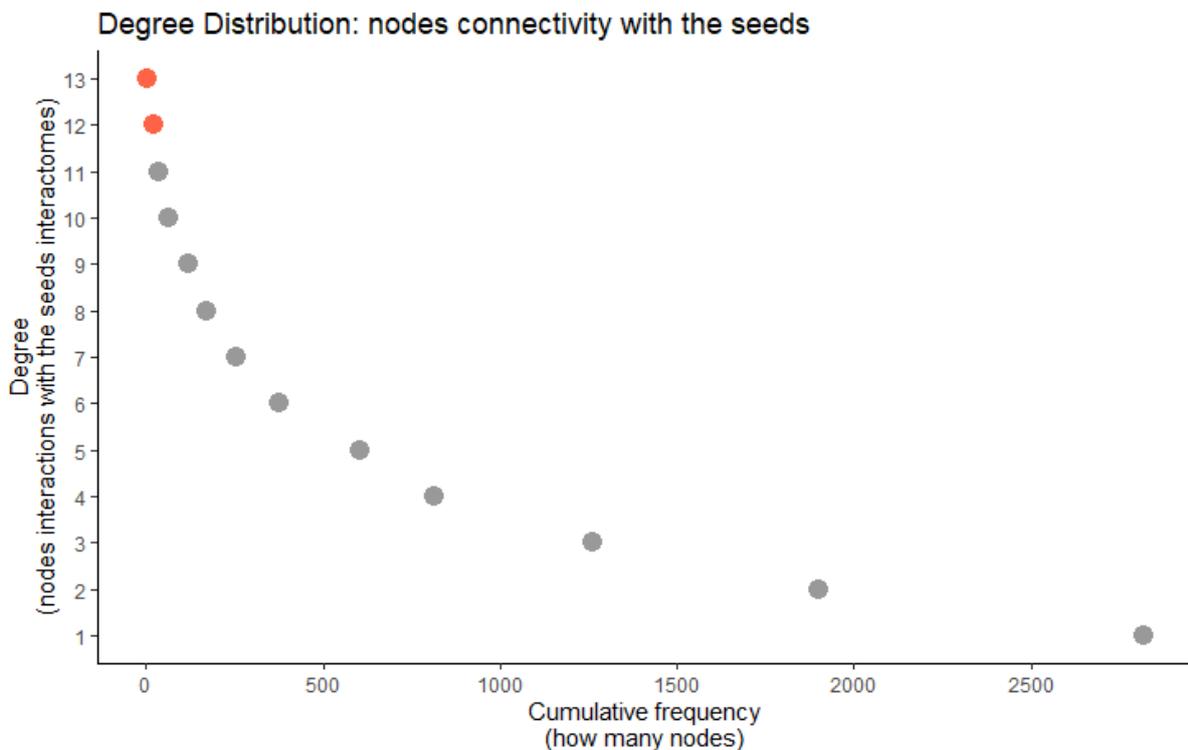


Figure 4.5. Inter-interactomes degree distribution. The number of nodes (x axis) is plotted as a function of the number of seed interactomes they belong to (y axis). The interactors shared across a minimum of 12 different seeds interactomes (highlighted in red) were selected as IIHs.

Specifically, IIHs are defined as those proteins within the FTLD-PPIN that are interactors shared across a minimum of 12 different seeds (12/18); IIHs are therefore able to bridge and keep the cohesion for at least 66% of the entire network. The analysis led to the selection of 26 nodes to be defined IIHs. I combined the list of different seeds bridged by

the IIHs reaching a total number of 17/18 seeds connected via the 26 IIHs (each of the IIHs connecting at least 12 different seeds) (**Table 4.2**).

IIHs	inter interactome degree	<i>C9ORF72</i>	<i>CHCHD10</i>	<i>CHMP2B</i>	<i>DCTN1</i>	<i>FUS</i>	<i>GRN</i>	<i>HLA-DRA</i>	<i>IFT74</i>	<i>MAPT</i>	<i>OPTN</i>	<i>SQSTM1</i>	<i>TARDBP</i>	<i>TBK1</i>	<i>TIA1</i>	<i>TMEM106B</i>	<i>UBQLN2</i>	<i>VCP</i>
EGFR	14	1	-	-	1	1	1	1	1	1	1	1	1	1	1	1	-	1
PDHA1	14	1	1	1	1	1	1	-	1	-	1	1	1	1	1	-	1	1
VCAM1	14	-	1	1	1	1	1	-	1	1	1	1	1	1	1	-	1	1
VCP	14	-	1	1	1	1	1	-	1	1	1	1	1	1	1	-	1	1
HSPA5	14	-	-	-	1	1	1	1	1	1	1	1	1	1	1	-	1	1
ARRB2	13	-	1	1	1	1	1	-	-	1	1	1	1	1	1	-	1	1
CTNNB1	13	1	1	-	1	1	-	-	-	1	1	1	1	1	1	1	1	1
ESR1	13	-	1	1	-	1	1	-	1	1	1	1	1	1	1	-	1	1
GRB2	13	-	1	-	1	1	1	-	1	1	1	1	1	1	1	-	1	1
HNRNPD	13	-	1	1	1	1	1	-	1	1	1	1	1	1	1	-	1	-
HSP90AA1	13	-	1	1	1	1	1	-	1	1	1	1	1	1	1	-	-	1
HSPA4	13	-	1	-	1	1	1	-	1	1	1	1	1	1	1	-	1	1
HSPA8	13	-	1	1	1	1	1	-	1	1	1	1	1	1	1	-	-	1
SQSTM1	13	1	1	-	1	1	1	-	1	1	1	1	1	1	-	-	1	1
YWHAZ	13	-	1	-	1	1	1	-	1	1	1	1	1	1	1	-	-	1
ACTB	12	-	-	1	1	1	1	-	1	1	1	1	1	1	-	-	1	1
BAG6	12	-	-	-	1	1	1	-	1	1	1	1	1	1	1	-	1	1
CDK2	12	-	-	1	1	1	1	-	-	1	1	1	1	1	1	-	1	1
EP300	12	-	-	1	1	1	-	-	1	1	1	1	1	1	1	-	1	1
FLNA	12	-	-	1	1	1	1	-	1	1	1	1	1	1	1	-	-	1
HSP90AB1	12	-	-	-	1	1	1	-	1	1	1	1	1	1	1	-	1	1
HSPB1	12	-	-	-	1	1	1	-	1	1	1	1	1	1	1	-	1	1
TARDBP	12	-	-	-	1	1	1	-	1	1	1	1	1	1	1	-	1	1
TCP1	12	-	-	1	1	1	1	-	1	1	1	1	1	1	1	-	-	1
TP53	12	1	-	1	-	1	1	-	1	1	1	1	1	1	-	-	1	1
XRCC6	12	-	-	1	1	1	1	-	1	1	1	1	1	1	1	-	-	1

Table 4.2. Inter-interactome hubs (IIHs) (n=26) interacting with ≥ 12 seeds interactomes.

Notably, the interactomes of RAB38 showed no interaction with any of the IIHs. This could be due to the fact that: i) these proteins do not genuinely share interactors with this FTLD seed or ii) literature does not report much information about it because it has not been extensively studied yet (ascertainment bias). Intriguingly, as shown in the previous section, gene co-expression analyses (WGCNA) showed poor co-expression module assignment for the same seed.

The IIHs were used to extract the most interconnected core of the network (core-FTLD-PPIN) containing the 26 IIHs, the seeds they connect to and any additional protein that act as a bridge between the seeds and the IIHs. The core-FTLD-PPIN was composed of 410 single interactors, connected by 1570 edges and 17 seeds (out of the total number of 18 FTLD seeds) (**Figure 4.6**).

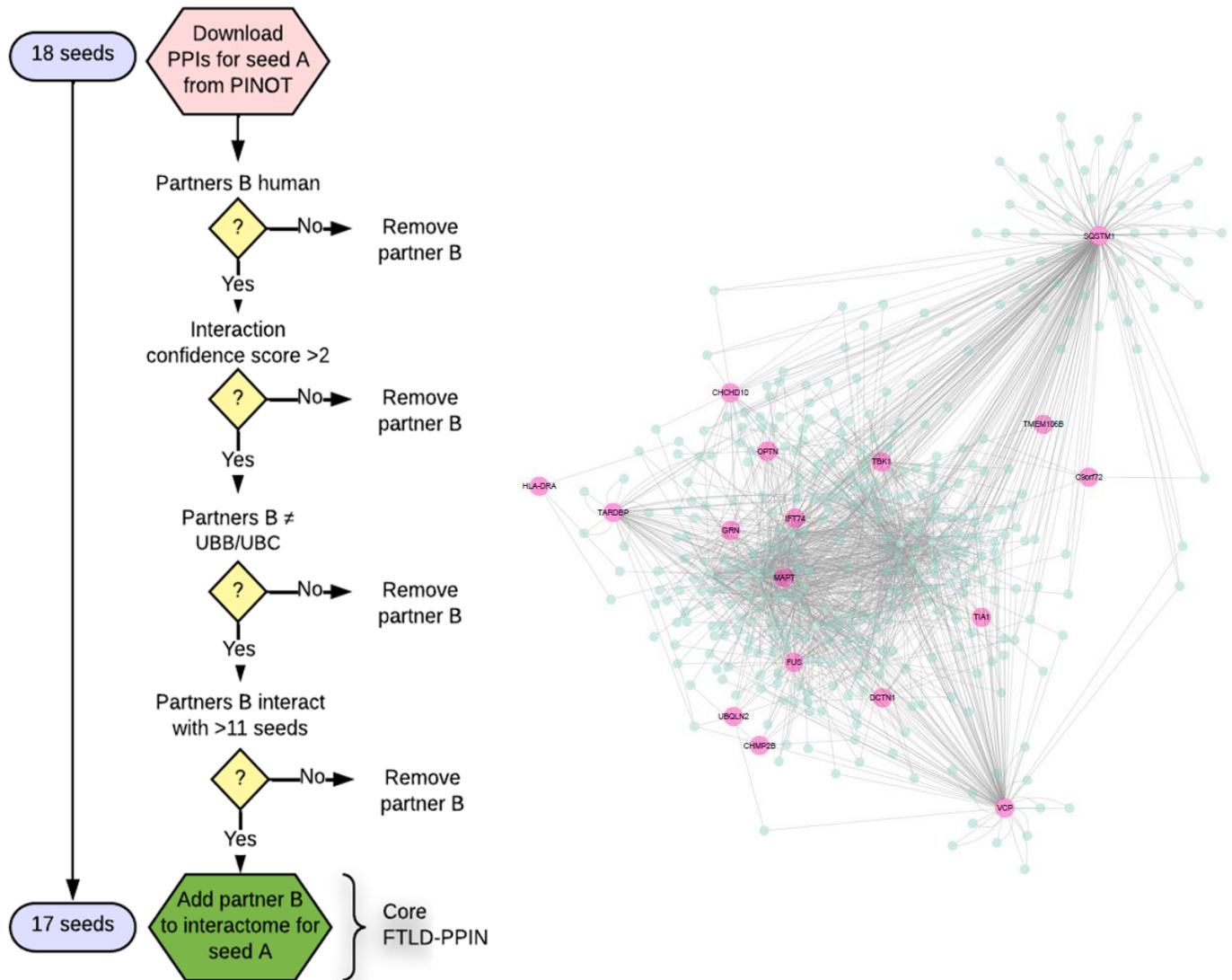


Figure 4.6. Workflow to generate the core-FTLD-PPIN. IIHs plus their first- and second-layer interactors combined in the core-FTLD-PPIN (n=410). Briefly, seeds' human protein-protein interactions (PPI) were downloaded from the PINOT platform. Only PPIs presenting with a confidence score > 2 (e.g., one publication + experimental validation) were retained. PPIs with ubiquitins (UBB, UBC) as well as nodes interacting with less than 11 seeds were removed from the analyses. The resulting PPIs formed the core FTLD-PPIN. Pink = seeds, green = core FTLD interactome.

While often biological networks can be partitioned into topological and /or functional clusters (modules) on the basis of the connectivity and functional similarity of their nodes, the core-FTLD-PPIN constitutes already a functional module *per se*, as it recapitulates all of the (known) genetic and pathological features of FTLN. Indeed, it presents with a high density of connections among the seeds and relatively few private partners, showing great cohesiveness and connectivity among the core FTLD proteins

and thus supporting the hypothesis that they could share a number of common mechanisms underlying FTLD pathology.

4.2.4. Functional enrichment and prioritisation of FTLD-risk genes/pathways

Functional enrichment was performed for the proteins in the core-FTLD-PPIN using gProfiler. All of the significantly enriched Gene-Ontology Biological Processes (GO-BP) terms ($p < 0.05$) were grouped firstly into semantic classes and then into functional blocks to ease the evaluation of the results. The analysis of the structure and content of the functional blocks revealed notable BPs, collectively implicating: (i) *adhesion*, (ii) *cell cycle*, (iii) *cell death*, (iv) *chromatin*, (v) *development*, (vi) *DNA metabolism*, (vii) *enzyme*, (viii) *exocytosis*, (ix) *histone*, (x) *immune system*, (xi) *intracellular organisation*, (xii) *membrane*, (xiii) *motility*, (xiv) *protein metabolism*, (xv) *response to stimulus*, (xvi) *RNA metabolism*, (xvii) *synapse*, (xviii) *telomers*, (xix) *transport*, (xx) *waste disposal* (**Figure 4.7**).

Functional enrichment of WPPINA proteins

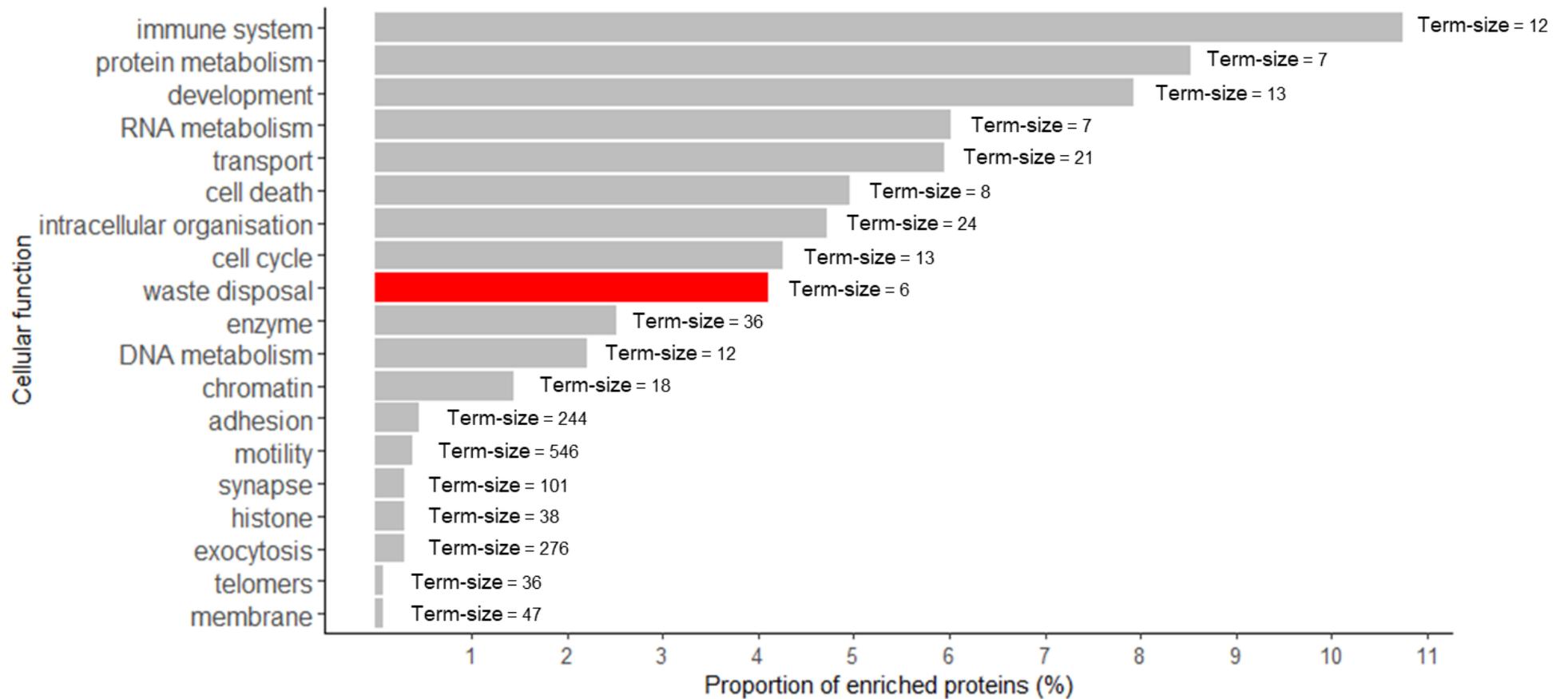


Figure 4.7. Functional enrichment analysis of core-FTLD-PPIN. Functional enrichment analysis extrapolates the functions that are driven by proteins shared at least by 66% of the seeds interactomes. The bar charts show the proportion of proteins enriched for specific cellular processes and their associated GO term size, which corresponds to the number of genes/proteins annotated for a specific GO term. Waste disposal showed the smallest term size (term size = 6), followed by protein and RNA metabolism (term size = 7) and cell death (term size = 8), implying that these processes are likely highly specific to FTL D pathogenesis.

On the basis of this analysis, these functional blocks indicated susceptibility processes shared by at least 66% of the core-FTLD-PPIN seeds suggesting processes of critical relevance to deepen our understanding of the molecular mechanisms underlying FTLD pathogenesis. The semantic classes contributing to (iii), (xiv), (xv) and (xx) revealed convergent information, particularly *cell death*-related activities indicated *apoptotic processes* in response to both *ER* and *oxidative stress* and included *regulation of mitochondrial activity* in the *apoptotic signalling pathway*. Interestingly, these combinations suggest a univocal process, such as that of endogenous/exogenous *cellular stress response* resulting in either the *positive regulation of mitochondria degradation pathway* or *cell death (apoptosis)*, as relevant BPs underlined by elements of the core-FTLD-PPIN. As a result, these semantic classes suggest the critical importance of *waste disposal* and *apoptosis*-related pathways in cell survival and homeostasis in an FTLD-risk scenario. To further dissect the relevance of the semantic classes for FTLD, I not only evaluated the proportion of enriched proteins (%) enriched for a specific cellular function but also ranked the ‘term size’ of each enrichment. The term size is a GO attribute which illustrates the number of genes/proteins annotated for a specific GO term (or pathway, i.e., to FTLD-specific processes): the greater the size of the GO term, the broader and non-specific it will be (many genes have been annotated for that term), while the least genes/proteins annotated for a specific term, the more specific the enrichment will be (few genes have been annotated for that term). Therefore, it could be argued that ‘term size’ is a proxy for the specificity of the GO term and that it is inversely proportional to the specificity. Out of a total of 19 cellular processes, *waste disposal* showed the smallest term size (term size = 6), followed by *protein* and *RNA metabolism* (term size = 7) and *cell death* (term size = 8), implying these four pathways as highly specific key biological processes implicated in FTLD pathogenesis and potential candidates for functional validation.

I then proceeded with functional enrichment for every WGCNA module to gain insight into their biological significance in the context of FTLD using gProfiler and then grouped all of the significantly enriched GO-BP terms ($p \leq 0.05$) into semantic classes, similarly to what was done to WPPINA enrichments (**Figure 4.8**).

Functional enrichment of WGCNA modules

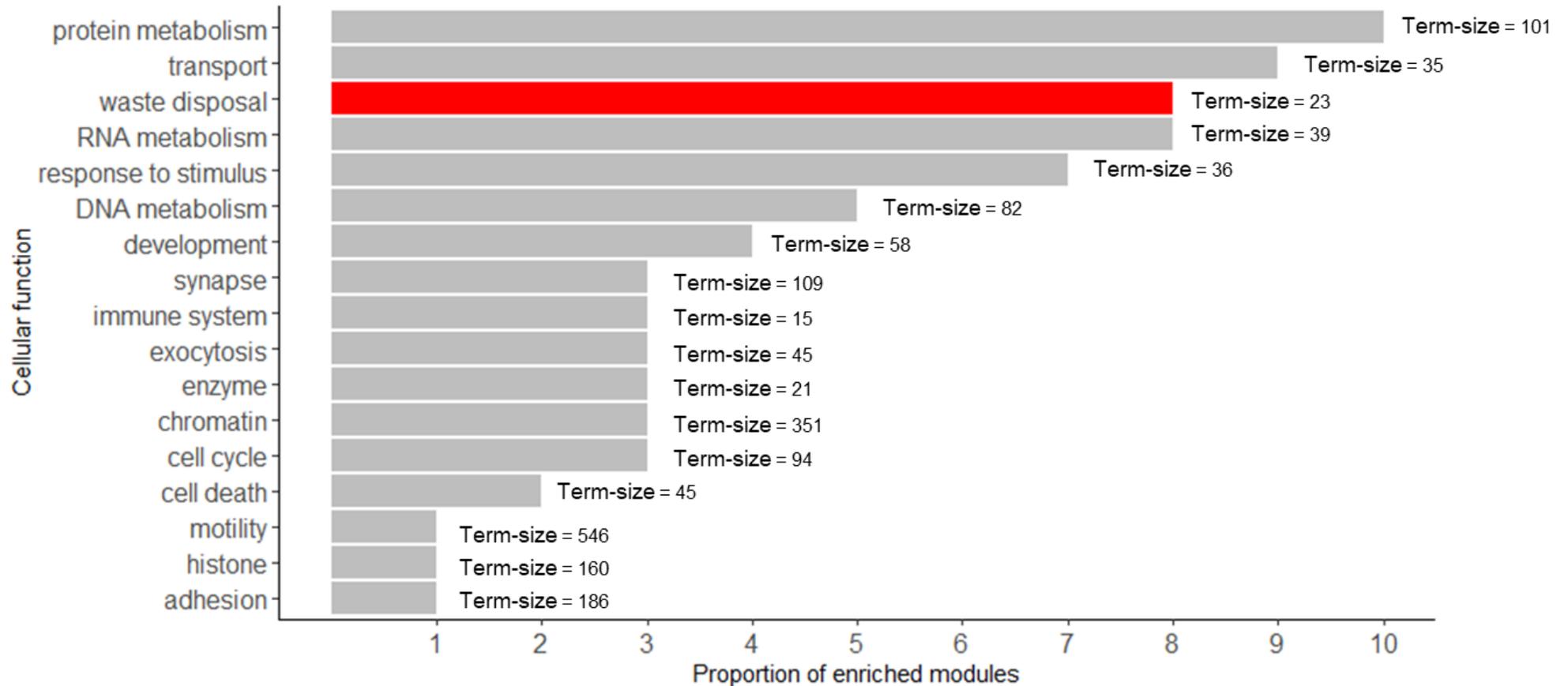


Figure 4.8. Functional enrichment analysis of GCNs. The bar chart shows the number of modules enriched for a specific cellular process their associated GO term size, which corresponds to the number of genes/proteins annotated for a specific GO term. Immune system showed the smallest term size (term size = 15), followed by enzyme (term size = 21) and waste disposal (term size = 23), implying that these processes are likely highly specific to FTLD pathogenesis.

Between 10 and 30% of the genes in each module was lost because it was not recognised by gProfiler, due to ambiguous GO ID or misleading annotation within GTEx and UKBEC datasets. Two of the modules ('UKBEC royalblue'; 'UKBEC blue') were removed from the analysis because gProfiler did not provide any significant enrichment. The module 'gtexv6 magenta' was also removed because the GO-BPs did not survive quality control. As a result, I obtained a list of GO-BPs enriched for 15 modules of interest (**Table 4.3**).

Module	Genes lost in gProfiler	Module size	% Loss
gtexv6 brown	306	2305	13.27%
UKBEC black	106	808	13.11%
UKBEC blue	440	1816	24.22%
gtexv6 cyan	110	952	11.55%
UKBEC darkgreen	99	634	15.61%
gtexv6 greenyellow	134	1218	11%
gtexv6 lightcyan	160	835	19.16%
UKBEC lightyellow	89	669	13.30%
gtexv6 magenta	363	1143	31.75%
UKBEC magenta	101	785	12.86%
gtexv6 midnightblue	215	919	23.39%
gtexv6 red	257	1464	17.55%
UKBEC red	107	877	12.20%
UKBEC royalblue	172	561	30.65%
UKBEC tan	142	580	24.48%
gtexv6 turquoise	459	2045	22.44%
gtexv6 darkred	291	1474	19.70%
UKBEC yellow	92	719	12.79%

Table 4.3. Absolute values and associated percentages of gene losses in gProfiler for each module. Framed are the modules that were not enriched for any particular BP and that were dropped as a result.

After having performed functional enrichment in gProfiler, generic terms (classified in the functional blocks of: *General – Metabolism – Physiology*) were discarded from further analyses. When considering the totality of the modules, functional enrichment showed the higher proportion of modules being related to *protein metabolism* (n = 10 modules) and *transport* (n = 9 modules), followed by *waste disposal* and *RNA metabolism*, which were both enriched for by n = 8 modules. Functional annotation analysis for the modules

(‘gtexv6 brown’, ‘UKBEC red’) containing *UBQLN2*, *C9ORF72* and *OPTN* mainly pointed to (i) *protein metabolism* entailing proteolysis and protein localisation, (ii) *transport*, particularly hinting to exocytosis and ER to Golgi vesicle-mediated transport, (iii) *intracellular organisation* associated with organelle biogenesis and localisation, as well as *waste disposal*-related processes (i.e. macroautophagy, cellular response to unfolded proteins and ubiquitin-dependent protein catabolic processes) when analysing the ‘UKBEC red’ module. Interestingly, the *CHMP2B-TMEM106B*-containing modules showed a similar functional enrichment, indicating *protein metabolism*, *transport* and *waste disposal* as the most significantly enriched BPs; major overlap between GO terms was observed across the four modules, particularly intriguing is the frequent recurrence of terms related to Golgi vesicle-trafficking activities (in *transport*), which could hint to a subtle involvement/function of the Golgi apparatus in FTLN-related pathways.

Also in this instance, I sought to further dissect the relevance of the semantic classes for FTLN by evaluating the ‘term size’ of each enrichment, which revealed high enrichment specificity for modules associated to *immune system* (n = 15), *enzyme* (n = 21) and *waste disposal* (n = 23). It is interesting to notice how *waste disposal* clearly seems a recurring biological process in two independent analyses, which cross-support each other and provide substantial evidence that this process might underpin FTLN pathogenesis.

Therefore, I pursued the *waste disposal*-related pathways for functional validation, which are among the processes that have been shown by several studies to hold high relevance in neurodegeneration, including FTLN^{78,200}, and that I here particularly prioritized in the context of FTLN. All of the GO terms assigned to the *waste disposal* functional block, as well as the complete list of proteins directly contributing to such enrichment (n=165) were extracted from the enrichment results obtained from the core-FTLN-PPIN functional annotation (**Figure 4.7**). Similarly, all the GO terms classified as *waste disposal* from the functional annotations of the WGCNA modules containing the core-FTLN-PPIN seeds were extracted (**Figure 4.8**).

Of note, the modules (i) 'UKBEC darkgreen', (ii) 'gtex lightcyan', (iii) 'UKBEC lightyellow', (iv) 'gtex midnightblue", (v) 'UKBEC tan', (vi) 'gtex turquoise', were not enriched for *waste disposal*, so they were not included in further analyses. Finally, the *waste disposal* GO-BP terms obtained from the analysis of the core-FTLD-PPIN and the co-expression modules were compared. The greatest level of overlap was identified for two critical semantic classes: *autophagy – mitophagy (Atg – Mito)* and *ubiquitin proteasome system (UPS)* (**Figure 4.9**).

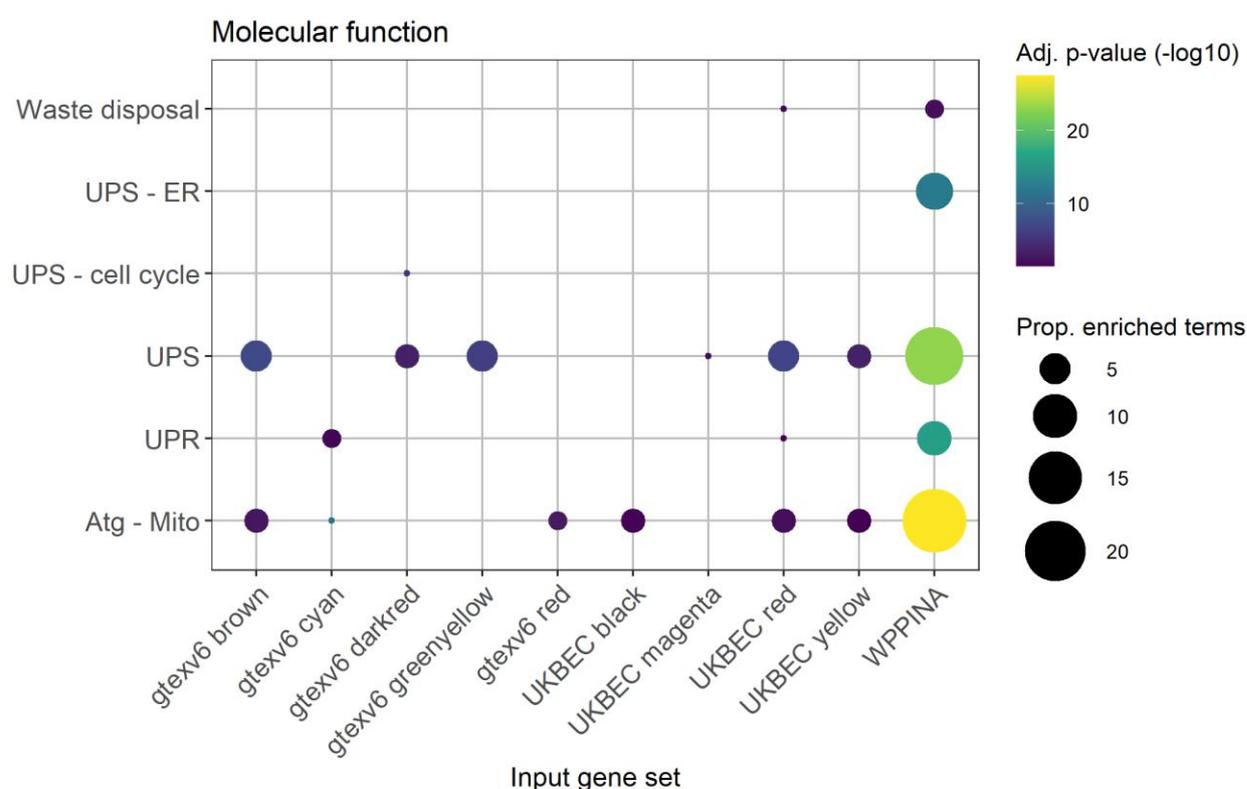
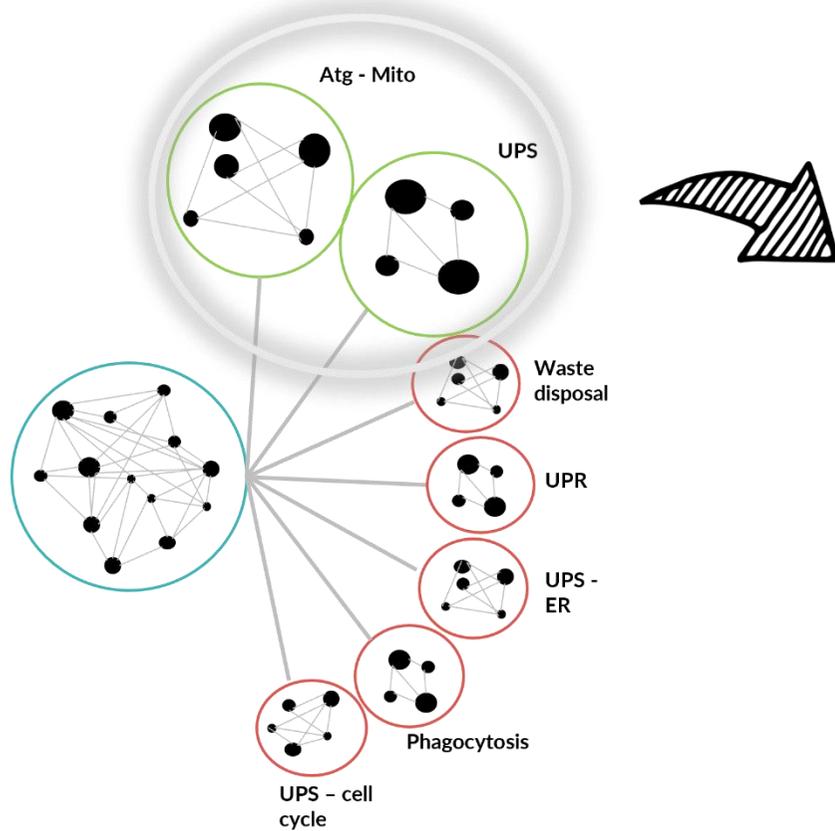


Figure 4.9. Functional enrichment for waste disposal of WPPINA and WGCNA. WGCNA modules and WPPINA proteins (y axis) enrichments are shown for the most relevant waste disposal semantic classes (x axis). The strength of the contribution to the enrichment is shown by the colour gradient, with yellow indicating a stronger enrichment (more GO terms associated to a specific BP) and dark purple indicating a lower enrichment (fewer GO terms associated to a specific BP). Blank fields indicate no enrichment at all for those semantic classes/modules.

Figure 4.9 indicates there are overlaps between functional blocks in WGCNA and WPPINA functional enrichments. Their associated GO terms were manually curated in order to double-check that the semantic classes had been accurately associated to the GO

terms. Finally, the 157 genes/proteins contributing to the functional enrichment in *Atg-Mito/UPS* pathways were used to generate a highly specific network (Atg-Mito/UPS FTLD PPIN) shown in Cytoscape (**Figure 4.10**).

Waste disposal macro-network - n=312



Atg-Mito/UPS FTLD PPIN network - n=157

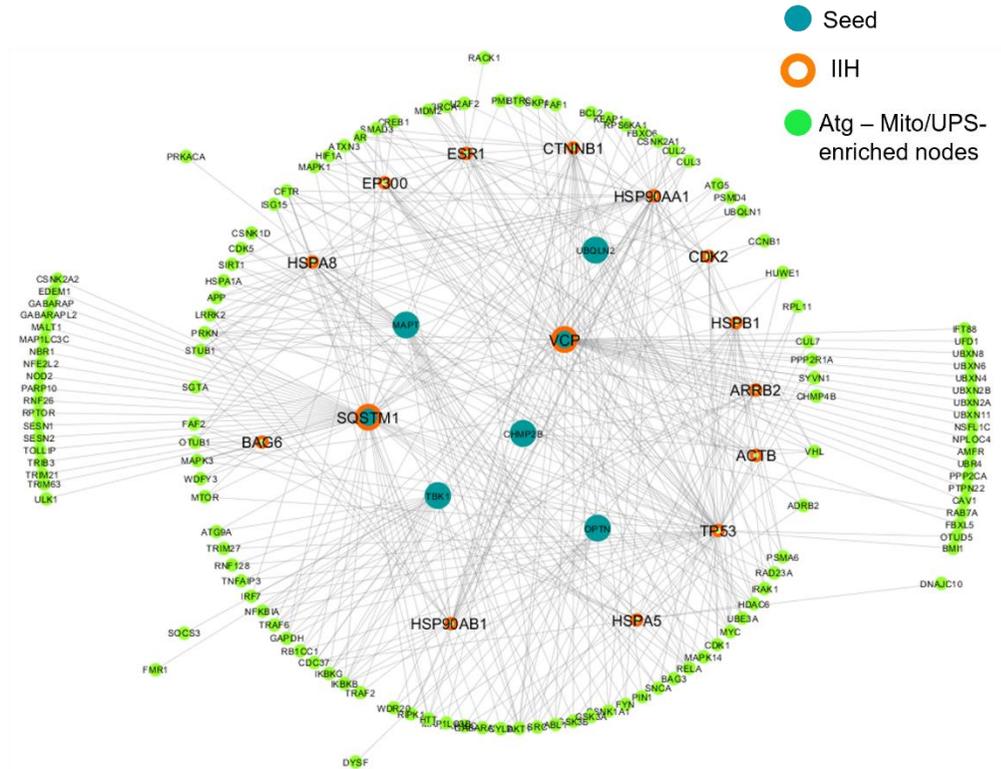


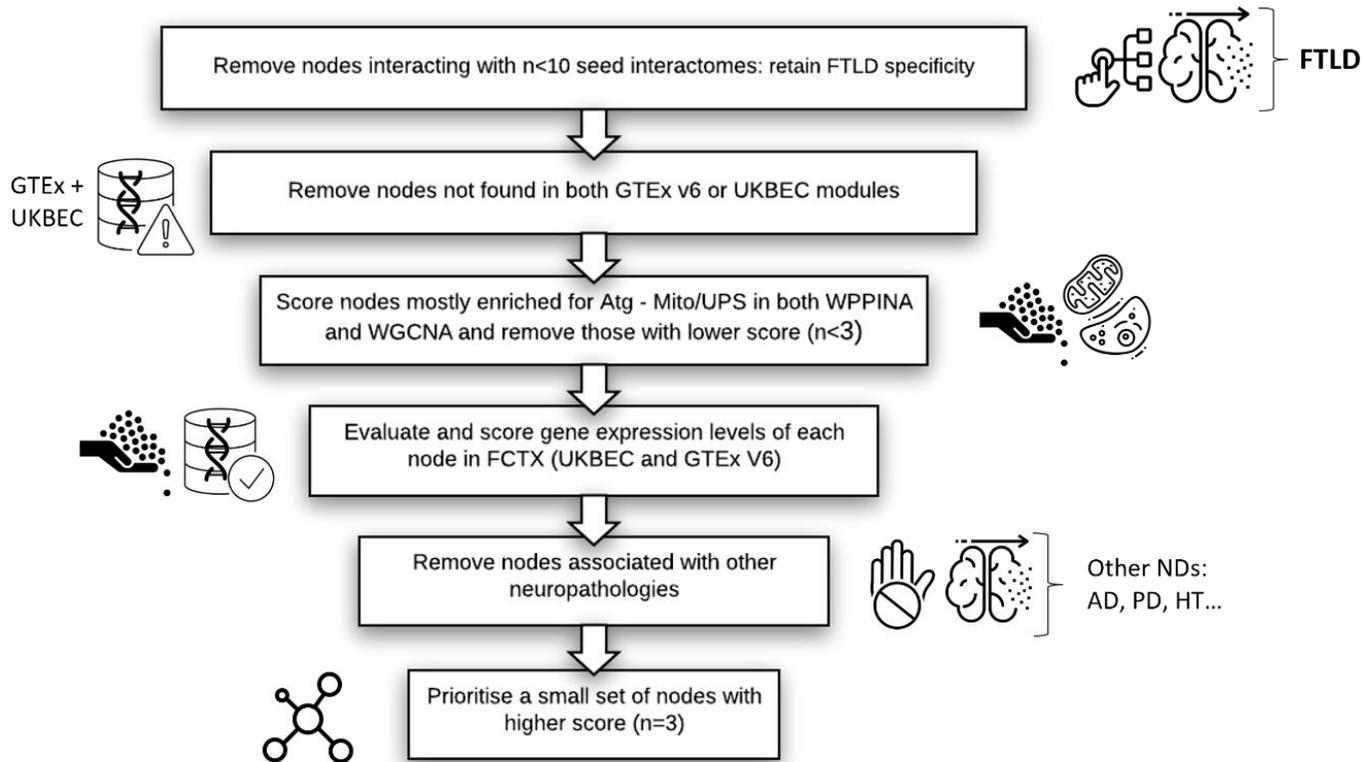
Figure 4.10. Topological features and functional patterns of the Atg - Mito/UPS-FTLD-PPIN. The Atg - Mito/UPS FTLD PPIN network is a subset of the *waste disposal* macro-network including only nodes that show functional enrichment for the Autophagy - Mitophagy (Atg - Mito) pathways and the Ubiquitin Proteasome System (UPS) (all shown in green). Some of the nodes composing this network had previously identified as seeds (blue, n=7) and IIHs (orange circle, n=15).

In order to functionally validate the *waste disposal*-related pathways prioritized in the context of FTLN, I sought to prioritize a small set of candidate FTLN-risk genes/proteins (from the Atg-Mito /UPS-FTLN-PPIN) to be modelled in a wet-lab experimental setting.

To aid the prioritisation of the candidate genes/proteins, I created a number of ranking variables to filter the Atg-Mito /UPS-FTLN-PPIN, which we applied to each node as follows: **(i)** how many seeds interactomes is each node connecting ($n=x$), **(ii)** is the node enriched for *Atg - Mito* in WPPINA analyses (yes=1, no=0), **(iii)** is the node enriched for *UPS* in WPPINA analyses (yes=1, no=0), **(iv)** is the node found in any UKBEC module (yes=1, no=0), **(v)** is the node found in any GTE_x module (yes=1, no=0), **(vi)** how many seeds are co-expressed in its UKBEC module ($n=x$), **(vii)** how many seeds are co-expressed in its GTE_x module ($n=x$), **(viii)** is the node enriched for *Atg - Mito* in its UKBEC module (yes=1, no=0), **(ix)** is the node enriched for *UPS* in its UKBEC module (yes=1, no=0), **(x)** is the node enriched for *Atg - Mito* in its GTE_x module (yes=1, no=0), **(xi)** is the node enriched for *UPS* in its UKBEC module (yes=1, no=0), **(xii)** is the node expressed in FCTX in UKBEC and to what extent ($n=x$), **(xiii)** is the node expressed in FCTX in GTE_x and to what extent ($n=x$).

The genes/proteins that were previously shown to be pathogenic for neurodegenerative disorders (i.e., *APP*, *LRRK2*, *APOE*, *SNCA*, *HTT*) were removed, as there is already strong evidence supporting their involvement in the pathogenesis of one or more neurodegenerative diseases. A final score was assigned to each node based on the sum of the binary scores of the functional enrichments obtained from WPPINA+WGCNA. The top nine genes/proteins with the highest final score were selected as prioritised genes. Finally, their expression levels in brain were checked (in GTE_x and UKBEC) to make sure that the values were high enough to allow successful *in vitro* modelling of the candidate genes/proteins. I then highlighted the three candidate genes/proteins with the highest score ($n \geq 3$): *CUL2*, *STUB1* and *CDC37* (**Figure 4.11**).

Prioritisation pipeline



Prioritised nodes

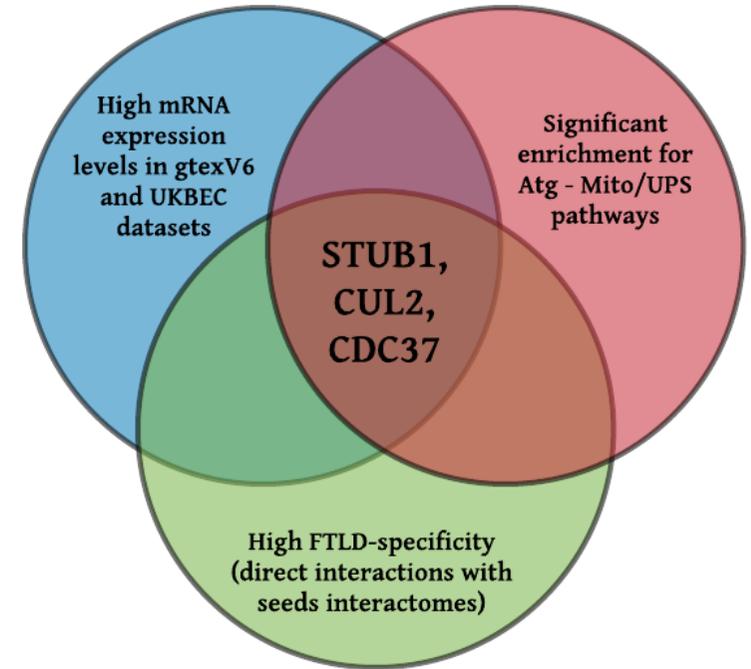


Figure 4.11. Workflow to prioritise FTLN-risk candidate genes and Venn diagram showing the three prioritised nodes (STUB1, CUL2 and CDC37).

These genes represent the molecular bridges connecting the majority of the FTLD Mendelian genes and have been prioritised for their involvement in the UPS and Atg – Mito pathways, which are the most relevant GO terms within the FTD functional enrichment. As a result, all three genes are predicted to be relevant in both biological processes, which means that their functional validation can be conducted in both pathways. In this study, the prioritised genes/proteins were first characterised in the Atg – Mito pathway; future work will be necessary to evaluate their role in the UPS pathway.

4.2.5. Re-evaluation of WGCNA to improve FTLD *in vitro* modelling

As FTLD presents with a particularly heterogeneous clinical and pathological phenotype, I sought to further define the functional environment of the prioritised candidate genes/proteins by investigating: i) their relationship with the Atg – Mito pathways; ii) their potential association with the main FTLD Mendelian genes (e.g., *C9ORF72*, *MAPT*, *GRN*). This re-evaluation will guide and support a more hypothesis-driven investigation prior to functional validation by suggesting a set of genetic and molecular markers likely to be involved in the functional environment of the candidate genes/proteins in the context of FTLD/Atg – Mito.

First, I sought to investigate the relationship between the prioritised candidate genes/proteins and mitophagy, using the *PINK1* gene as a representative marker for PINK1-dependent mitophagy. For this analysis, I used the previously generated GCNs in both GTEX V6 and UKBEC for FCTX (**Table 4.4**).

Gene	Category	MM
PINK1	gtexv6 darkred	0.7
	UKBEC darkturquoise	0.2
STUB1	gtexv6 darkred	0.8
	UKBEC magenta	0.5
CDC37	gtexv6 red	0.8
	UKBEC black	0.5
CUL2	gtexv6 brown	0.9
	UKBEC red	0.8

Table 4.4. Comparing co-expression patterns between PINK1 (Mito marker) and the prioritised genes in FCTX. For each gene, GTEX V6 and UKBEC module names and associated module membership (MM) values are reported.

By further dissecting the associated GCNs, it can be observed that *PINK1* and *STUB1* fall in the same ‘gtexv6 darkred’ module ($MM_{PINK1}=0.7$; $MM_{STUB1}=0.8$), suggesting a potential functional link between these two genes. Interestingly, in this instance UKBEC did not replicate these results.

By contrast, *CDC37* and *CUL2* can be found into the 'gtexv6 red' and 'gtexv6 brown' modules respectively ($MM_{CDC37}=0.8$; $MM_{CUL2}=0.9$), implying that neither of them might be associated to the mitophagy pathway.

Another useful feature of the CoExp web application is that GCN inter-modular similarity in terms of gene expression can be visually inspected using the homonym package in R. To investigate the higher order organisation among gene co-expression modules in GTEx V6 I hierarchically clustered module eigengenes (MEs), which are defined as the first principal component of the expression matrix of the corresponding module. To guarantee scale free topology for the network, WGCNA generates an adjacency matrix and a Topology Overlap Matrix (TOM), with $1-TOM$ being used as the 'distance' or 'height' between modules for hierarchical clustering. The smaller the distance ($1-TOM$, y-axis) between modules for hierarchical clustering. The smaller the distance ($1-TOM$, y-axis) between different modules, the greater their similarity. In summary, the modules hanging from the same root are very similar in terms of expression while they tend to be increasingly different the further apart they lie (**Figure 4.12**).

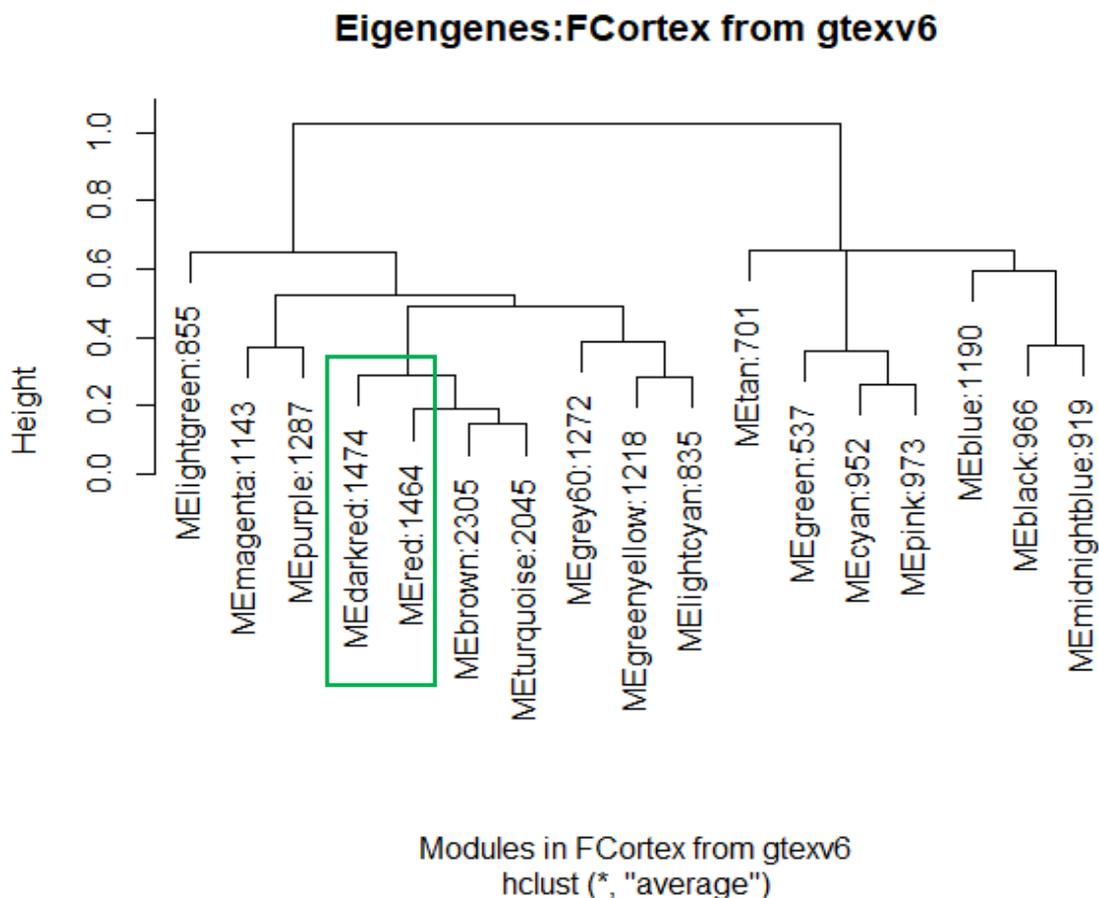


Figure 4.12. Dendrogram comparing inter-modular relationships among GTE_x V6 modules. The dendrogram shows the results of hierarchically clustering gte_xv6 module eigengenes (MEs) to examine higher-order relationships between the modules. “Leaves” along “branches” represent MEs. The y-axis represents network distance as determined by 1 – TOM, where values closer to 0 indicate greater similarity between main sources of expression perturbation in the modules.

In this instance I observed that the ‘gte_xv6 darkred’ (which includes both *PINK1* and *STUB1*) and the ‘gte_xv6 red’ (which includes *CDC37*) modules are closely related in the Eigengene correlation dendrogram (height [network distance, 1-TOM] = 0.1), indicating greater similarity between main sources of expression perturbation in such modules, and therefore in the genes composing them. By contrast, the ‘brown’ modules in gte_xv6 (where *CUL2* falls into) seems to not to be related in any way to the ‘gte_xv6 darkred) module as it presents with a considerable network distance. Notably, further investigation of the inter-modular connectivity between UKBEC modules did not lead to validation of what found in GTE_x V6, as the ‘UKBEC black’ and the ‘UKBEC darkturquoise’ modules, where *CDC37* and *PINK1* are found respectively, are very distant from each other (**Figure 4.13**).

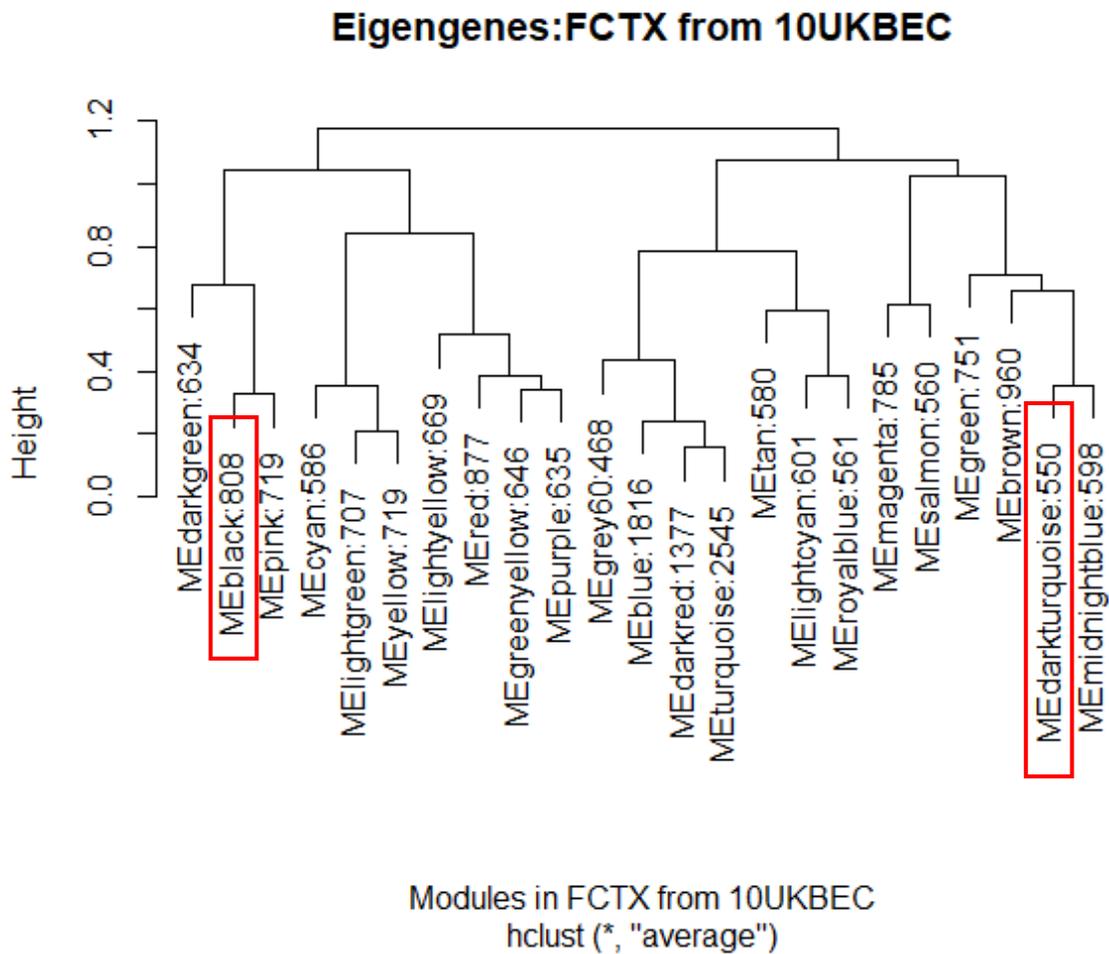


Figure 4.13. Dendrogram comparing inter-modular relationships among UKBEC modules. The dendrogram shows the results of hierarchically clustering UKBEC module eigengenes (MEs) to examine higher-order relationships between the modules. “Leaves” along “branches” represent MEs. The y-axis represents network distance as determined by $1 - \text{TOM}$, where values closer to 0 indicate greater similarity between main sources of expression perturbation in the modules.

Furthermore, similarly to previous analyses, I sought to investigate the relationship between the prioritised candidate genes/proteins and the main Mendelian FTLD genes, namely *C9ORF72*, *MAPT* and *GRN*, in order to guide a more targeted experimental validation in an FTLD model (**Table 4.5**).

Gene	Category	MM
<i>C9ORF72</i>	gtexv6 brown	0.8
	UKBEC red	0.7
<i>MAPT</i> (LOC100130148)	NA	NA
	UKBEC black	0.8
<i>GRN</i>	gtexv6 cyan	0.7
	UKBEC black	0.7
<i>STUB1</i>	gtexv6 darkred	0.8
	UKBEC magenta	0.5
<i>CDC37</i>	gtexv6 red	0.8
	UKBEC black	0.5
<i>CUL2</i>	gtexv6 brown	0.9
	UKBEC red	0.8

Table 4.5. Comparing co-expression patterns between the FTLD Mendelian and the prioritised genes in FCTX. For each gene, GTEx V6 and UKBEC module names and associated module membership (MM) values are reported.

The two more notable features that can be observed from **Table 4.5** are that: i) *C9ORF72* and *CUL2* co-expression pattern in the GTEx V6 dataset is fully mirrored by UKBEC, respectively in the ‘gtexv6 brown’ ($MM_{C9ORF72}=0.8$; $MM_{CUL2}=0.9$) and ‘UKBEC red’ ($MM_{C9ORF72}=0.7$; $MM_{CUL2}=0.8$) modules, strongly implying that they could be co-expressed and thus involved in the same biological processes; ii) *CDC37* is co-expressed with both *MAPT* and *GRN* in the ‘UKBEC black’ ($MM_{CDC37}=0.5$; $MM_{MAPT}=0.8$; $MM_{GRN}=0.7$) module, suggesting that its involvement in FTLD might be associated to mutations associated to either gene. Notably, LOC100130148 currently acts as a surrogate gene ID for *MAPT* in the CoExp application and has been found to be expressed in the UKBEC dataset only.

4.3. Discussion

The objective of my PhD project was to test whether a holistic approach integrating genetics, transcriptomics and proteomics is an effective strategy to predict disease-specific processes on the basis of known FTLD-risk genes and, as a result, to guide and support hypothesis-driven functional validation in a cell biology setting. Future implementation of this very pipeline to identify and further investigate other key biological process implicated in the FTLD-trait (i.e., *protein metabolism*, *DNA damage response*, *immune system*) might be suggested if this approach is successful.

In this chapter, I used a WGCNA-WPPINA-combined approach to integrate knowledge generated within the related yet distinct field of genomics and proteomics, aiming at identifying and highlighting key protein players and biological processes in FTLD pathogenesis.

As reported above, I constructed the entire-FTLD-PPIN first and second layer interactors for 18 seeds (*C9ORF72*, *DCTN1*, *FUS*, *GRN*, *IFT74*, *MAPT*, *OPTN*, *SQSTM1*, *TARDBP*, *TIA1*, *TBK1*, *VCP*, *TMEM106B*, *UBQLN2*, *CHMP2B*, *HLA-DRA*, *CHCHD10* and *RAB38*). I first analysed the entire-FTLD-PPIN and identified 26 IIHs - the nodes that bridge 66% of the complete network - that hold structural, and possibly functional, significance in the disease context. Interestingly, several genetic and functional studies reported that many of such IIHs (*VCAM1*, *VCP*, *HSPA8*, *CDK2*, *EP300*, *TCP1*, *ESR1*, *TP53*, *TARDBP*, *SQSTM1*, *CDK2*, *PDHA1*, *GRB2*, *YWHAZ*, *ARRB2*, *FLNA*, *HSP90AA1*, *HSP90AB1*, *HSPA4*, *HSPB1*; 20/26 [77%])^{93,201-209} have been widely associated with dementia (across the spectrum of AD, dementia with Lewy bodies [DLB], vascular, and HIV-1-induced dementia). It is remarkable to note that 9 (*VCAM1*, *VCP*, *HSPA8*, *CDK2*, *EP300*, *TCP1*, *ESR1*, *TP53*, *CDK2*) IIHs replicated the findings from a previous WPPINA-WGCNA study⁹³, further validating the inference power of our approach. I expect that the remainder of the IIHs (*HSPA5*, *HSPA8*, *EGFR*, *ACTB*, *BAG6*, *XRCC6*) for which no link to dementia is established yet, might hold relevance to FTLD or dementias and should be prioritised in both genetic assessments and for further investigation and characterisation at the molecular level.

From a functional perspective, I assessed BPs for the nodes composing the relevant WGCNA modules, the core-FTLD-PPIN and subsequently the Atg - Mito/UPS-FTLD-PPIN. The first analysis revealed up to 21 BPs, strongly implying the following susceptibility processes: (i) *protein metabolism*, (ii) *transport*, (iii) *intracellular organisation*, (iv) *waste disposal*. Even more strikingly, I observed that these four functional blocks highlighted by my gene co-expression data through WGCNA replicated the results I obtained through WPPINA, in particular from the functional enrichment of the core-FTLD-PPIN, implying the potential extent of the translation of co-expression patterns into the protein domain. I therefore gather that these four functional blocks are, collectively, the most common BPs associated with the entire-FTLD-PPIN and are therefore likely to represent the core functional architecture impacted in FTLD for a proportion of cases. Among these functional blocks, I identified *waste disposal* as the most biologically intriguing and data-driven BP. *Waste disposal* encompasses GO terms, such as *Atg - Mito* and *UPS* (particularly

referring to ubiquitin-dependent catabolic processes, macroautophagy and cellular [ER] response to misfolded proteins [UPR]), that seem to be acting in a synergistic fashion: activation of UPR causes transient attenuation of protein synthesis, increased capacity for protein trafficking and processing through ER-associated degradation (ERAD) and macroautophagy^{78,200}. These are the BPs I suggest being worthy of attention for further investigation and characterisation at the molecular level.

I further assessed WGCNA and WPPINA together and not only I identified, again, a high level of overlap between *Atg - Mito* and *UPS* -associated GO terms within the *waste disposal* functional block, but also a considerable number of nodes shared across the two BPs, further confirming that *Atg - Mito* and *UPS* might represent sensitive FTLD-risk pathways to be targeted for wet-lab validation.

These results imply that the identical biological processes and/or nodes defining the overlapping genes/proteins between WGCNA and WPPINA represent the elements where the transcriptome (WGCNA) and the proteome (WPPINA) fully match in a cross-supportive fashion⁹³, lending support to the inference that these very elements are key functional factors in FTLD to be carried forward in the cell biology setting for hypothesis-driven functional validation. This all together reinforces the view that this pipeline is able to extract and integrate orthogonal information contributing to the generation of novel and more comprehensive awareness about the implication of genes/proteins in the pathobiology of a trait.

Furthermore, I developed a pipeline which integrates convergent multi-omics data to prioritise key functional players implicated in complex disease traits. This pipeline is helpful to emphasise pivotal components/hallmarks (genes, proteins, BPs) of disease pathogenesis (i.e., FTLD) and yields a comprehensive view of the biological context of disease. In the current study, I identified three candidate genes/proteins (STUB1, CDC37, CUL2) that are highly shared, through WPPINA and WGCNA analyses, by the FTLD seeds interactomes and transcriptomes, and that are enriched for FTLD-risk processes, corroborating their structural significance within such processes and, most importantly, suggesting that they may hold functional relevance in FTLD. These three genes will be carried forward for functional validation in a wet-lab experimental setting.

At present, *in vitro* modelling of complex diseases, such as neurodegenerative disorders, presents with many challenges. The fact that multiple genetic and environmental factors variably contribute to disease pathogenesis makes it very difficult to design a functional model that truly recapitulates most of the pathogenic signatures of disease. As FTLD presents with a particularly heterogeneous clinical and pathological phenotype, I sought to further define the functional environment of the prioritised candidate genes/proteins by re-evaluating their co-expression profiles and their associated modules' relatedness. Several studies have shown that having more information about the connection (adjacency) between two nodes and/or two modules and about their topological overlap provides a more robust and sensitive measure of interconnectedness, informing on their degree of belonging to the same functional class²¹⁰. With this in mind, I was particularly interested in investigating the candidate genes/proteins' relationship with the Atg – Mito pathways and their potential association with the main FTLD Mendelian genes (e.g., *C9ORF72*, *MAPT*, *GRN*), in order to better guide and support a more hypothesis-driven investigation prior to functional validation. This characterisation will be helpful to predict which genetic and molecular markers are likely to be involved in the functional environment of the candidate genes/proteins and leverage them in the context of FTLD/Atg – Mito during experimental validation.

A further dissection of the FTLD-GCNs in FCTX revealed that *PINK1* and *STUB1* fall into the same 'gtexv6 darkred' module, suggesting that they likely share similar transcriptional expression patterns and therefore they are likely to regulate similar biological processes (i.e., PINK1-dependent mitophagy). Few studies have already reported about the role of *STUB1* in mitochondrial regulation and clearance in response to stress^{106,211}, in mitochondrial homeostasis, morphological reorganisation and relocalisation within the cell^{106,211}, proteasomal degradation (UPS)²¹² and autophagy^{213,214}, cross-supporting the co-expression data I obtained and thus reinforcing its strong functional implication in mitophagy. It is worth noting that, out of the three candidate genes/proteins, *STUB1* is the one that has been most extensively studied and well-characterised, which implies that I might be incurring in an ascertainment bias by prioritising this gene over the others.

By contrast, *CDC37* and *CUL2* fall into the 'red' and 'brown' modules respectively in GTEX V6, suggesting different co-expression patterns from *PINK1*. Interestingly, by hierarchically clustering the GTEX V6 modules in FCTX, I observed that the 'gtexv6 red'

and the 'gtexv6 darkred' modules, where *CDC37* and *PINK1* (and *STUB1*) fall into respectively, are very closely related in the Eigengene correlation dendrogram, indicating a great degree of topological overlap and thus functional similarity across the genes composing them. These results are supported, although to a lesser extent than for *STUB1*, by literature, which shows that *PINK1* is a client kinase of *CDC37* and their interaction is crucial for *PINK1* stability and as a result for mitochondrial homeostasis²¹⁵⁻²¹⁸. Although not fully characterised yet, the relationship between *CDC37* and mitophagy is intriguing and it would be interesting to follow up in a wet-lab experimental setting.

By contrast, *CUL2* exhibits the weakest link to *PINK1* and generally mitophagy: as an E3 ubiquitin ligase, it has been shown to act as a key player in responding to hypoxic stress^{219,220} and clearance of misfolded protein aggregates²²¹ but according to current literature it does not seem to be involved in mitochondrial clearance in any way.

These results further reinforce the hypothesis of a potential involvement of these three genes/proteins, with a particular focus on *STUB1* and *CDC37*, in *PINK1*-dependent mitophagy and supports their further investigation in this molecular pathway.

Further re-evaluation of the GCNs showed that *C9ORF72* and *CUL2* are co-expressed in both the GTEx V6 and the UKBEC datasets. Replication of these data in two different datasets (GTEx V6 and UKBEC) increasingly strengthens the idea of a potential link between *C9ORF72* and *CUL2*. Interestingly, Uchida et al.²²¹ had identified *CUL2* as a novel ubiquitin ligase for fragmented forms of TDP-43, which is a typical protein inclusion characterising FTLD-ALS patients carrying *C9ORF72* repeat expansions^{222,223}. Another important finding is that *CDC37* is found in the same module as both *MAPT* and *GRN*. Although no significant link has been made with these specific genes yet, its involvement in the molecular mechanisms of neurodegenerative diseases has been reported several times. Specifically, *CDC37* has been suggested to play a critical role in regulating the phosphorylation of disease-associated proteins by a number of kinases, thus provoking their pathological aggregation and potentially contributing to neurodegeneration by: i) altering tau and TDP-43 clearance, hence supporting a link with FTLD-ALS and AD^{218,224}; ii) contributing to the regulation of α -synuclein phosphorylation, which is linked to its pathological aggregation in PD²²⁴.

Given the complexity of genetic and molecular landscape of FTLD and the lack of disease models that accurately recapitulate the complex pathologies of disease¹⁴¹, these results have a number of implications when projecting them on a potential *in vitro* model for FTLD; when generating or choosing *in vitro* models of genetic FTLD and other complex disorders, this type of analysis can guide the most adequate choice by aiding the selection of the FTLD Mendelian genes mutation on the basis of its association with the prioritised candidate genes/proteins. In this instance, for example, I would expect to find changes in CDC37 transcriptional regulation or proteostasis in a *GRN* or *MAPT* model rather than a *C9ORF72* model. Similarly, if I would decide to further investigate CUL2 in a disease-relevant model I would probably select a *C9ORF72* mutant rather than any other. Additionally, increasing progress in annotating cell-type-specific transcriptome data will further improve the accuracy of models for complex diseases by recapitulating the cell type(s) that degenerate in disease and in specific genetic scenarios.

All this taken together opens the way for at least two opportunities: (i) to further explore and corroborate susceptible BPs and their associated key protein players using *in silico* approaches (i.e. Panther, David, Reactome) and validate such pathways *in vitro* through hypothesis-driven cell biology investigations; (ii) to integrate FTLD-risk BPs and associated proteins with pharmacogenomics data by using compounds/drugs that already modulate relevant protein targets (or BPs) and could potentially positively correct disease trait.

In summary, this study presents a novel integrative approach to improve our understanding of neurological disorders by guiding wet-laboratory experiments with multidisciplinary knowledge and by supporting drug discovery and prospective implementation for patient benefit.

It is important to note that a number of limitations apply to this approach and to WPPINA particularly. First, this is a seeded approach, which means that it carries a fundamental bias associated to the choice of the genes used to create the networks. Secondly, in the context of neurodegenerative diseases it is very difficult to define which genes are 'Mendelian': most of these diseases are relatively rare, with few families carrying a specific genetic mutation, therefore there are widespread disagreements across the scientific community regarding which genes are to be considered Mendelian. Also, FTLD specifically is a very heterogeneous disease from a clinical, genetic and pathological

perspective⁹, and strikingly it presents with a clinical spectrum overlapping with ALS, which makes it even harder for researchers to select genes truly associated to FTLN pathogenesis. To minimise this bias as much as possible, I selected the genes using the most recent review of FTLN Mendelian genes⁹ at the time. Furthermore, since I acknowledge that in the past three years more genes could have been identified as potentially associated to FTLN pathology, I performed a literature review to pinpoint any additional FTLN gene. A recent genetic study by Wagner and colleagues²²⁵ detected through exome sequencing analysis pathogenic variants in the *CTSF* and *ABCA7* genes in a German FTLN cohort, suggesting that they might be linked to risk of developing FTLN. Notably, several lines of evidence (e.g., GWASs, sequencing) had previously identified *ABCA7* as a risk locus for AD, reinforcing its possible implications in NDS²²⁶⁻²³⁰. In the same study, mutations in the *RBM33* and *TET2* genes were found in a number of FTLN and/or ALS cases, indicating a potential causal role for these genes in FTLN. Altogether, these genes might confer risk of developing FTLN and should be considered as FTLN candidate seeds for future re-implementation of this approach.

With regards limitations applying to WPPINA, it is important to note that PPI network data are essentially a static representation of these interactions and do not include information about the organisation of protein dynamics as well as the effect of post-translational modifications (PTMs)¹⁵⁷. Second, many of the PPI data are collected from wet-lab approaches that are low-throughput and hypothesis-driven, therefore they are biased by a limited number of proteins selected on the basis of research-driven priorities. As a result, PPI networks generally contain false positive and false negative interactions, distorting our understanding of the cell's functional organisation^{93,157}. To mitigate the impact of this issue, I complied with the approach pursued by Ferrari et al.⁹³, which constructed the first layer of the network as a basis of the second layer to (i) dilute the above-mentioned bias and (ii) minimise the effect of seed-centrality, an issue that is intrinsic in networks of this kind (i.e. generated on the basis of trait-specific known associated genes). Additionally, the same bias applies to the functional enrichment analysis because GO is based on current literature, therefore some of the known PPIs might not have been comprehensively studied and/or annotated, and thus the enrichment analysis queries partial annotation data (ascertainment bias).

Chapter 5

5. Results

5.1. Validation of FTLD-risk candidate genes *in vitro*

5.1.1. Introduction

Several diseases have been associated with alterations of the biological processes related to *waste disposal*, such as defective autophagosome formation and/or fusion with the lysosome or deficiency in autophagosomal/lysosomal enzymes, chaperones or receptors^{9,78,199}. As a result, a deeper understanding of the mechanisms and factors involved in the maintenance of a balanced proteome and their dysregulation can provide invaluable insights into the possible application of therapeutic strategies in disease. FTLD pathology is characterised by the presence of intracellular, insoluble inclusions composed of misfolded proteins which is associated to a number of cellular defects including ER and/or oxidative stress, excitotoxicity, disruption of the ER-Golgi trafficking, RNA processing defects and mitochondrial dysfunction (**Figure 5.1**).

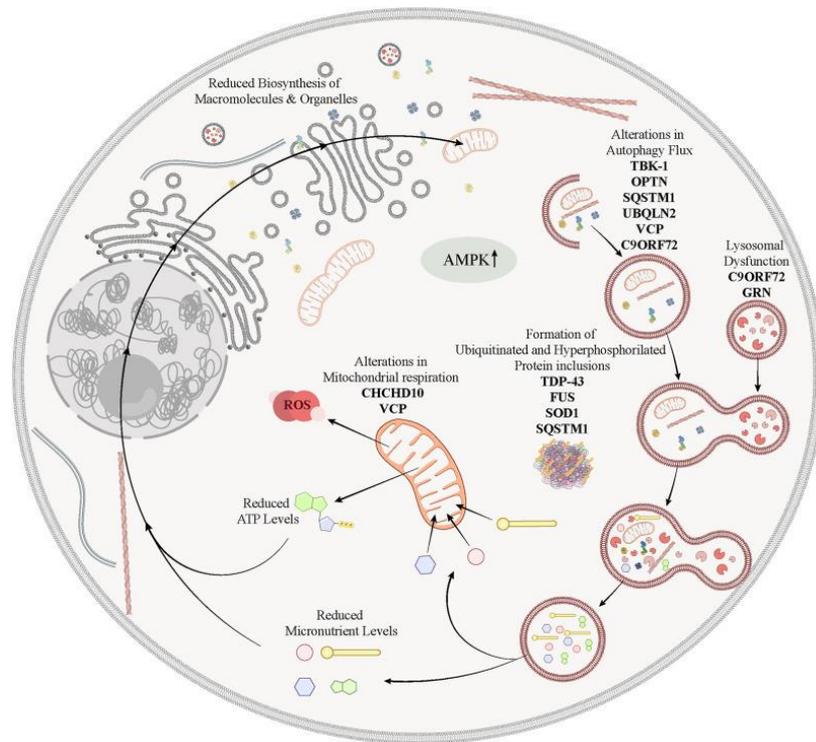


Figure 5.1. Schematic representation highlighting the involvement of various genetic risk factors associated to ALS/FTD in defective mechanisms of nutrient sensing and autophagy. In ALS/FTD, alterations in the autophagic flux caused by mutations in *TBK1*, *OPTN*, *SQSTM1*, *UBQLN2*, *VCP*, *C9ORF72* or *GRN* may lead to the accumulation of misfolded proteins into pathological inclusions, and the improper recycling of damaged organelles, as for example mitochondria, to produce micronutrients. Damaged mitochondria or mutations in *CHCHD10* or *VCP* that are associated with defective mitochondrial phenotypes, would further aggravate the energy crisis that is precipitated by the failure of energy-sensing pathways, such as defective lysosomal signaling or dysregulated AMPK activation. Adapted from 'Defects of Nutrient Signaling and Autophagy in Neurodegeneration' by J. Ondaro, 2022, *Frontiers in Cell and Developmental Biology*. Adapted with permission.

In this thesis, I implemented a bioinformatics pipeline to prioritise FTLD-risk candidate genes/proteins to guide their experimental validation *in vitro*. Protein-protein interaction (PPIN) and gene co-expression (GCN) networks resulted to be enriched for a number of biological processes (BPs) associated with FTLD: BPs associated to *waste disposal* and in particular to the autophagy-mitophagy (Atg – Mito) and ubiquitin-proteasome system (UPS) pathways were highlighted as disease (FTLD)-specific processes. Additionally, three candidate genes/proteins were identified as likely relevant molecular players within these FTLD-specific processes. As a result, I sought to further characterise the three candidate genes/proteins (*STUB1*, *CUL2*, *CDC37*) in the context of PINK1-dependent mitophagy, which is a type of selective autophagic degradation of damaged mitochondria (see Section 1.1.4.4). Association between the three candidate genes/proteins and mitophagy has been partially documented and further gene co-

expression network analyses (refer to Section 5.1.1.1) suggested that they are likely to regulate similar biological processes and that they might be relevant in the mitophagy pathway.

5.1.1.1. The prioritised genes/proteins: STUB1, CUL2 CDC37

The integrated bioinformatics approach deployed in this study and outlined in **Chapter 4** allowed for systematic prioritisation of three candidate genes/proteins that might be involved in the pathogenesis of FTL. Their main structural and functional features as well as their potential implications in neurodegenerative disorders are reported below.

STIP1 Homology And U-Box Containing Protein 1 (STUB1), also known as CHIP, (C-terminus of HSC70-interacting protein), is a ubiquitously expressed cytosolic 35 kDa protein with dual function of co-chaperone and E3 ubiquitin ligase activity, thus acting as a connecting link between molecular chaperones and proteasomes^{231,232}. From a structural perspective, STUB1 contains three tetratricopeptide repeats (TPR domains) at its N-terminal and a U-box domain at its C-terminal. The TPR domain mediates the interaction with major cytoplasmic chaperones such as HSC/HSP70 and HSP90, while the U-box modulates the proteasomal degradation of numerous chaperone-bound misfolded protein substrates via ubiquitination. Through the action of both domains, STUB1 plays a pivotal role in mediating triage decisions between protein refolding and degradation, as well as a regulatory role in immunity and necroptosis²³². Consistently, deficiency or mutations in the *STUB1* gene can lead to disorders which involve misfolding and aggregation of proteins, such as spinocerebellar autosomal-recessive ataxia type 16 (SCAR16), Gordon-Holmes syndrome, and spinocerebellar ataxia type 48 (SCA48)^{231,233}. Interestingly, several studies have shown that STUB1 overexpression can have neuroprotective functions in neurological diseases by enhancing clearance of pathogenic A β , α -synuclein and mutant huntingtin deposits respectively in AD, PD, and HD, making it a promising therapeutic strategy to rescue disease toxicity^{103,232,234,235}.

Cullin 2 (CUL2) is a core component of multiple cullin-RING-based ECS (ElonginB/C-CUL2/5-SOCS-box protein) E3 ubiquitin-protein ligase complexes. Cullin-RING ligases (CRLs, the largest family of E3 ubiquitin ligases), are multiprotein complexes involved in

protein degradation through 26S proteasome²³⁶. They are assembled on a cullin scaffold (CUL1, CUL2, CUL3, CUL4A, CUL4B, and CUL5) and contain a RING finger protein (RBX1 or RBX2), a substrate receptor and adaptor proteins^{236,237}. CUL2 serves as a rigid scaffold in the CRL2 complex and contributes to the catalysis through positioning of the substrate and the ubiquitin-conjugating enzyme. The E3 ubiquitin-protein ligase activity of the complex is dependent on the neddylation (the process of conjugating the ubiquitin-like protein NEDD8 onto a substrate) by the cullin subunit leading to substrate proteasomal degradation. The most well-known CRL2 substrate recognition receptor is the tumour suppressor protein VHL, that is mutated in von Hippel–Lindau (VHL) syndrome, and that interestingly has been reported to ubiquitinate misfolded forms of TDP-43 and promote its clearance²²¹. Additionally, a recent study by Yasukawa et al.²³⁷ showed that CRLs are involved in the regulation of A β production and fibrillation in neurons by targeting BRI2 and BRI3 (i.e., physiological inhibitors of amyloid precursor protein [APP] processing and A β oligomerisation) for degradation. Finally, hypoxia inducing factors (HIFs) have been shown to undergo CRL2-VHL complex-dependent degradation, and it is the most studied role of CRL2 ubiquitin ligase in tumorigenesis²²¹.

Cell division cycle 37 (CDC37) is a ubiquitous protein that often forms a complex with the 90 kDa heat shock protein (Hsp90), a molecular chaperone that requires a partner (or co-chaperone) to assist with client triage¹⁰². The Hsp90/Cdc37 complex interacts with ~60% of the kinome and by doing so promotes the stabilisation and activation of a variety of client kinases including CDK4, CDK6, SRC, RAF-1, MOK, as well as eIF2 alpha kinases, that are collectively critical regulators of cell cycle progression, signal transduction and transcriptional regulation²²⁴. To date, several studies have shown that CDC37 is implicated in cancer and neurodegenerative disease-related pathways, as its depletion can stabilise tau and TDP-43 by promoting their clearance via autophagy^{102,218,238}. Additionally, the Hsp90/CDC37 complex has been shown to regulate ULK1 and other kinases involved in macroautophagy, suggesting a potential role of CDC37 in autophagy/mitophagy and, in the context of NDs, as a candidate drug target for clearance of toxic accumulation of misfolded proteins^{102,239}. Interestingly, while CDC37 mainly functions in conjunction with Hsp90 as an active scaffold for protein kinase recruitment, it seems to also have independent chaperone activity, thanks to which it maintains the homeostasis of a number of proteins, such as the beta-galactosidase enzyme and the mitochondrial serine/threonine-protein kinase PINK1, which has been shown to depend

on CDC37 for stabilisation on the OMM upon mitochondrial stress and correct subcellular distribution^{215,216,240,241}.

5.1.1.2. Modelling PINK1-dependent mitophagy

As described in section 1.1.4.4, PINK1-dependent mitophagy is the selective autophagy of mitochondria which plays a key role in maintaining the quality of the mitochondrial pool as well as regulating the abundance of intracellular mitochondria upon environmental cues¹³³. Nowhere is the requirement for effective mitochondrial quality control systems more important than in neurons, where proper mitochondrial function is paramount to satisfy the high energetic demands and need for high calcium-buffering capacity due to action potential-driven calcium influxes. Such strong dependence on mitochondrial function exponentially increases neuronal vulnerability to mitochondrial insults, and, in turn, makes efficient and properly functioning mitochondrial quality control pathways essential for their survival¹²⁸. In this scenario, it is evident that when investigating the implications of PINK1-dependent mitophagy in neurodegenerative disease, characterisation of the pathway in the relevant cell-type is key.

Although studies examining PINK1-dependent mitophagy in heterologous cell cultures have provided seminal insights on the molecular mechanisms of PINK1 and Parkin activation and action in the past, the existence and relevance of such a pathway in neurons has remained elusive, making its characterisation in neuronal disease models very challenging^{87,128}. Generally, studies investigating PINK1-dependent mitophagy predominantly rely on Parkin over-expression (POE) models in combination with chemical uncouplers to induce mitochondrial depolarisation^{242,243}; however studies conducted in neuronal models were not able to show recruitment at mitochondria of overexpressed Parkin nor endogenous Parkin-mediated mitophagy, raising an ongoing debate on a possibly different role for Parkin recruitment in neuronal PINK1-dependent mitophagy^{87,128,244,245}. Clearly, these data indicate that the physiological and pathophysiological role of Parkin recruitment and, as a result, of PINK1-dependent mitophagy in neurons differs from that of cultured cell lines; a possibility is that differences in bioenergetics between neurons and cultured cell lines might contribute to these different responses, as suggested by Van Laar et al.²⁴⁵ An additional factor worth

considering is that neuronal culture protocols carry many more variables than those for immortalized cells (e.g., neuronal media components), and many of them could influence Parkin translocation¹²⁸. Cai and colleagues²⁴³ for instance used a cocktail of apoptosis inhibitors in their neuronal culture (e.g., the caspase inhibitor Z-VAD-FMK) to counter the effects of high doses of chemical uncouplers triggering apoptosis in an environment devoid of protective glia. While these conditions do lead to parkin translocation in neurons, they may also mask the normal physiological reaction of neurons to gross depolarization of the mitochondrial network. It is unlikely that neurons have evolved to adapt to this type of insult, and apoptosis may be the resulting physiological response. By contrast, the absence of antioxidants (in the form of the B-27 supplement) showed a significantly increased ratio of Parkin recruitment^{128,246}. Taken together, these factors might explain why some studies show such variability in detecting Parkin translocation upon mitochondrial depolarization in neurons.

This study aims to investigate the molecular implications of the three prioritised candidate FTLN-risk genes/proteins in PINK1-dependent mitophagy using biochemical techniques and fixed/live cell imaging approaches to dissect mitophagy and mitochondrial clearance upon genetic silencing of the prioritised genes. PINK1-dependent mitophagy will be investigated by pharmacologically disrupting $\Delta\psi_m$ using the mitochondrial uncoupler O/A (oligomycin/antimycin), which has been shown to provide a robust and effective experimental paradigm, although PINK1-dependent mitophagy has also been observed under less severe conditions¹²⁸. Immortalised as well as iPSC (induced pluripotent stem cell) lines will be used to model neurons as closely as possible.

5.1.2. Results

5.1.2.1. *CDC37* KD increases pUb (Ser65) accumulation in SH-SY5Y cell line and increases *PINK1* mRNA expression

Based on my bioinformatics analyses (see **Results** section in **Chapter 4**), I sought to validate the three candidate genes/proteins (STUB1, CUL2, CDC37) for their involvement in autophagy/mitophagy processes by evaluating their possible implication in PINK1-dependent mitophagy.

In order to test whether the three prioritised genes may play a role in PINK1-dependent mitophagy, I individually knocked them down (KD) using siRNA in Parkin over-expressing (POE)-SH-SY5Y human neuroblastoma cells and immunoblotted for phosphorylation of ubiquitin on serine 65 (pUb [Ser65]), a PINK1-dependent mitophagy marker¹⁴⁰, following mitochondrial depolarization. Increased mitochondrial clearance following mitochondrial depolarization induced by treatment with 1 μ M of oligomycin/antimycin (O/A; mitochondrial stressors inhibiting respectively complex V and III of the electron transport chain) for 3h was validated as an endpoint for mitophagy. To confirm siRNA efficiency, individually transfected POE SH-SY5Y cells were immunoblotted with specific STUB1, CUL2, and CDC37 antibodies. Extensive testing by the HPF lab and others has shown endogenous PINK1 is difficult to detect in untreated whole cell lysate (WCL) as it is rapidly degraded under basal conditions. *PINK1* KD in WCL was therefore assessed either by qPCR to measure *PINK1* mRNA levels or by assessing PINK1-dependent phosphorylation of ubiquitin at Ser65 (pUb [Ser65]). Hits were identified as those candidate genes for which associated O/A-induced pUb (Ser65) was significantly altered following specific siRNA KD when compared scrambled (SCR) negative control siRNA. Throughout this chapter, additional mitochondrial markers were employed to evaluate relevant phenotypes: i) the OMM protein TOM20 was used to assess mitochondrial load in the sample; ii) TIM23, a major protein translocase embedded in the IMM²⁴⁷, was used to assess mitochondrial degradation; iii) MFN2, a mitochondrial protein that participates in mitochondrial, which is phosphoubiquitinated by PINK1/Parkin and rapidly degrades during mitophagy²⁴⁸ and whose degradation was used to assess mitophagy initiation upon O/A treatment (refer to section **1.1.4.4**).

Out of the three candidate genes tested, only *CDC37* KD showed significant changes in accumulation of pUb (Ser65) at mitochondria. IB in WCL showed that upon *CDC37* KD, pUb (Ser65) accumulation at mitochondria was significantly increased following O/A-induced mitochondrial depolarisation (**Figure 5.2**). These results suggest that *CDC37* might play a regulatory role in PINK1-dependent mitophagy. As a result, *CDC37* was taken forward for further functional analyses.

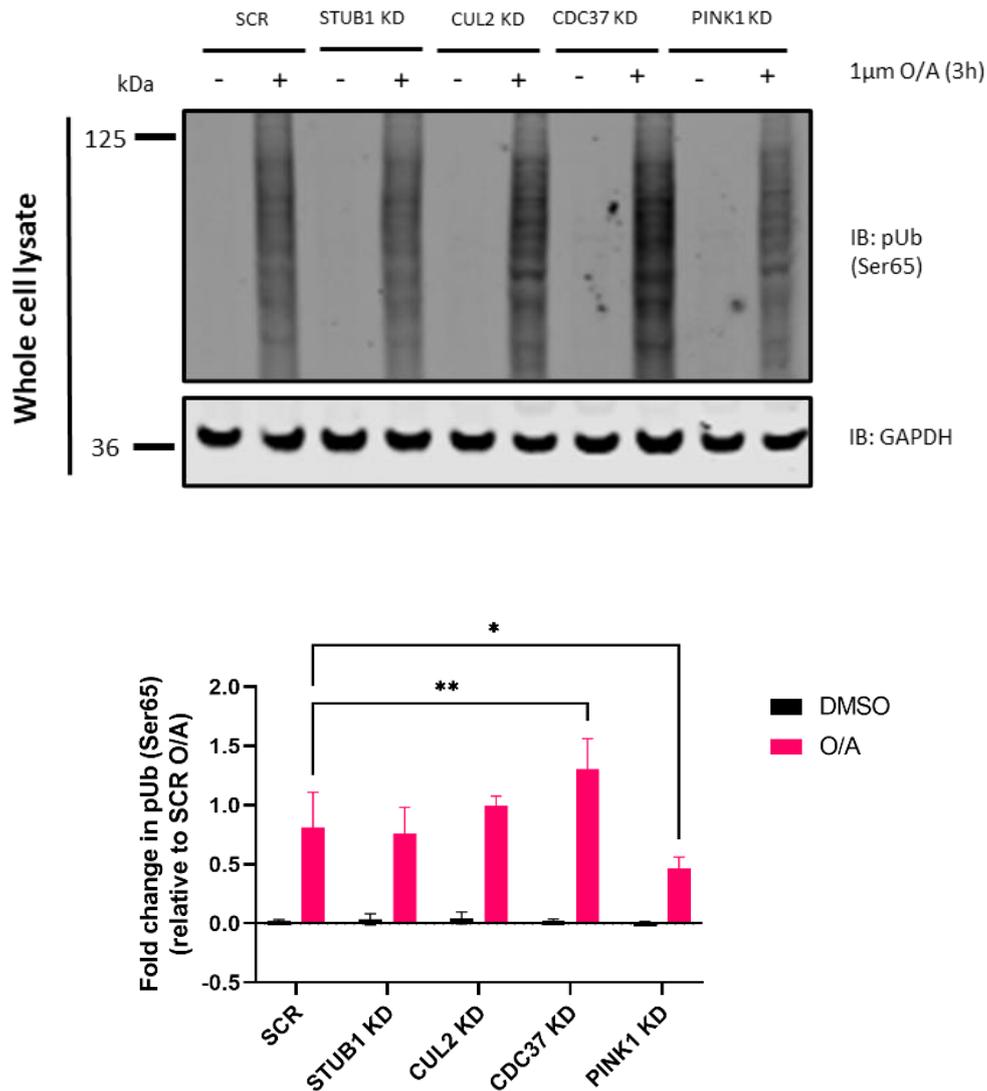


Figure 5.2. *CDC37* KD increases pUb (Ser65) accumulation in POE SH-SY5Y cells (WCL). Representative IB of WCL from SCR, *STUB1*, *CUL2*, *CDC37*, *PINK1* siRNA KDs in POE SH-SY5Y treated with 1 μM O/A for 3 h. IB for phospho-ubiquitin (pUb) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) control was performed. pUb protein expression upon siRNA KDs was quantified by densitometry and expressed as fold change relative to the SCR negative control. Histograms indicate mean of n=4 independent experiments ± StD. quantification data were normalised against a non-transfected control and analysed with an ordinary two-way ANOVA with Dunnett's correction. * p ≤ 0.05, ** p ≤ 0.01 compared to scrambled control.

It was further shown that *STUB1* and *CUL2* KD had no impact on degradation of the outer mitochondrial membrane (OMM; MFN2, TOM20) and inner mitochondrial membrane (IMM; TIM23) markers, further suggesting that these genes do not regulate the mitophagy process (data not shown). As a result, *STUB1* and *CUL2* were discarded from further investigation.

In order to better understand the biochemical processes underlying CDC37 and mitophagy, the effect of *CDC37* KD on pUb (Ser65) was further evaluated in POE SH-SY5Y mitochondria-enriched fractions, allowing for improved detection of mitochondrial markers, in particular PINK1. IB of mitochondria-enriched fractions following *CDC37* KD confirmed that pUb (Ser65) is significantly increased after 3h O/A treatment when compared to SCR control (**Figure 5.3**).

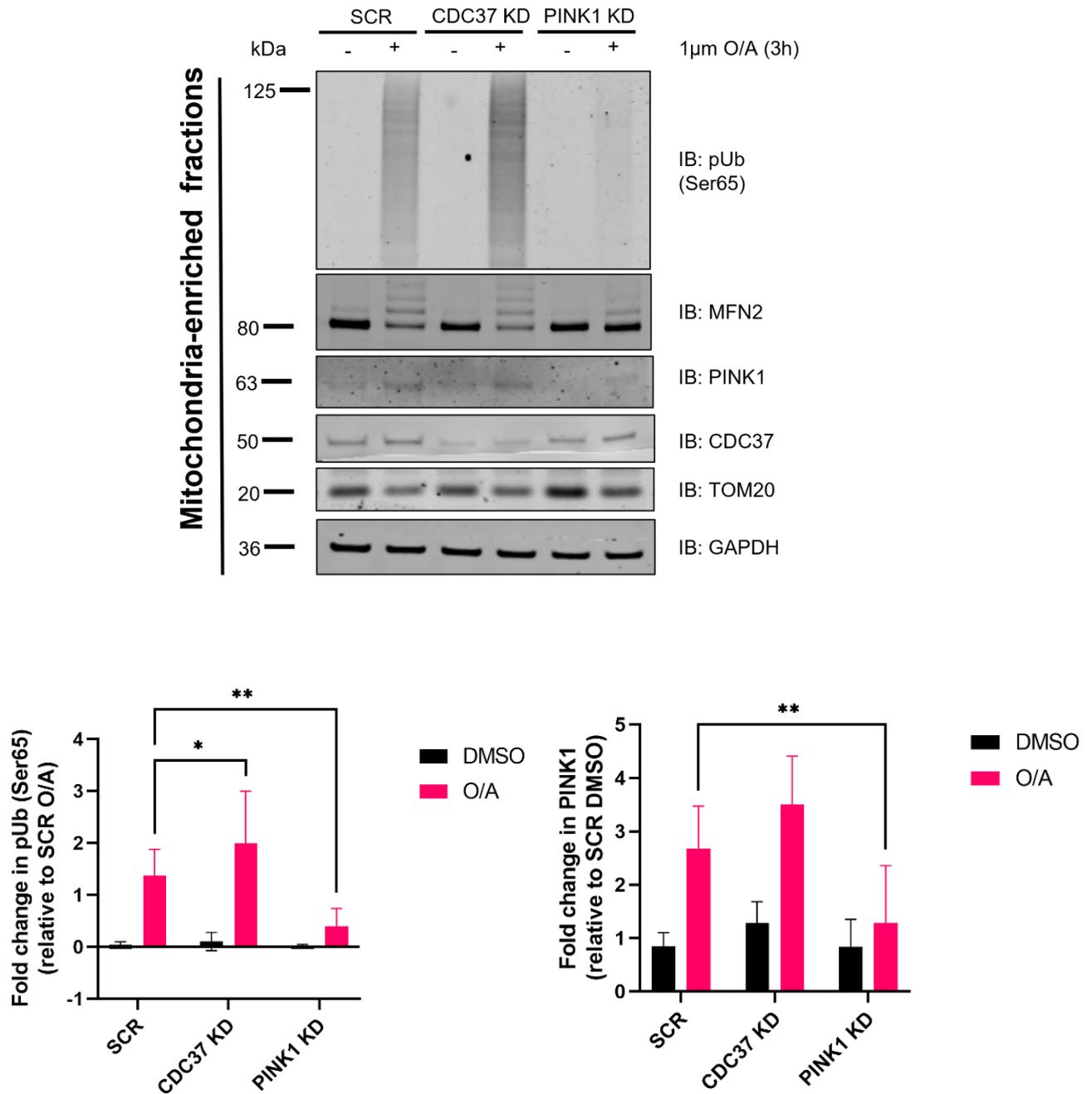


Figure 5.3. *CDC37* KD increases pUb (Ser65) accumulation in POE SH-SY5Y cells (mitochondria-enriched fractions). Representative IB of mitochondria-enriched fractions from SCR, *CDC37*, *PINK1* siRNA KDs in POE SH-SY5Y treated with 1 μM O/A for 3 h. IB for phospho-ubiquitin (pUb), mitofusin 2 (MFN2), PTEN-induced kinase 1 (PINK1), cell-division cycle 37 (CDC37), translocase of the outer membrane protein 20 (TOM20) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) control was performed. pUb protein expression upon siRNA KDs was quantified by densitometry and expressed as fold change relative to the SCR negative control. Histograms indicate mean of n=7 (pUb) and n=4 (PINK1) independent experiments ± StD. Quantification data were normalised against non-transfected control and analysed with an ordinary two-way ANOVA with Dunnett's correction. * p ≤ 0.05, ** p ≤ 0.01 compared to scrambled control.

Interestingly, PINK1 protein expression in mitochondria-enriched fractions shows an increasing trend upon *CDC37* KD, and although it does not display statistical significance ($p = 0.0674$) it might at least partially explain the increase in pUb (Ser65). With regards the remaining mitochondrial markers, both MFN2 (mitophagy marker) and TOM20 (mitochondrial load marker) showed reduced degradation upon *PINK1* silencing, confirming its successful KD, but no significant changes upon *CDC37* KD.

The effect of *CDC37* KD on pUb (Ser65) was further assessed in POE SH-SY5Y cells treated with 1 μ M O/A using immunofluorescence (IF). However, *CDC37* KD did not have any effect on pUb (Ser65) (**Figure 5.4**).

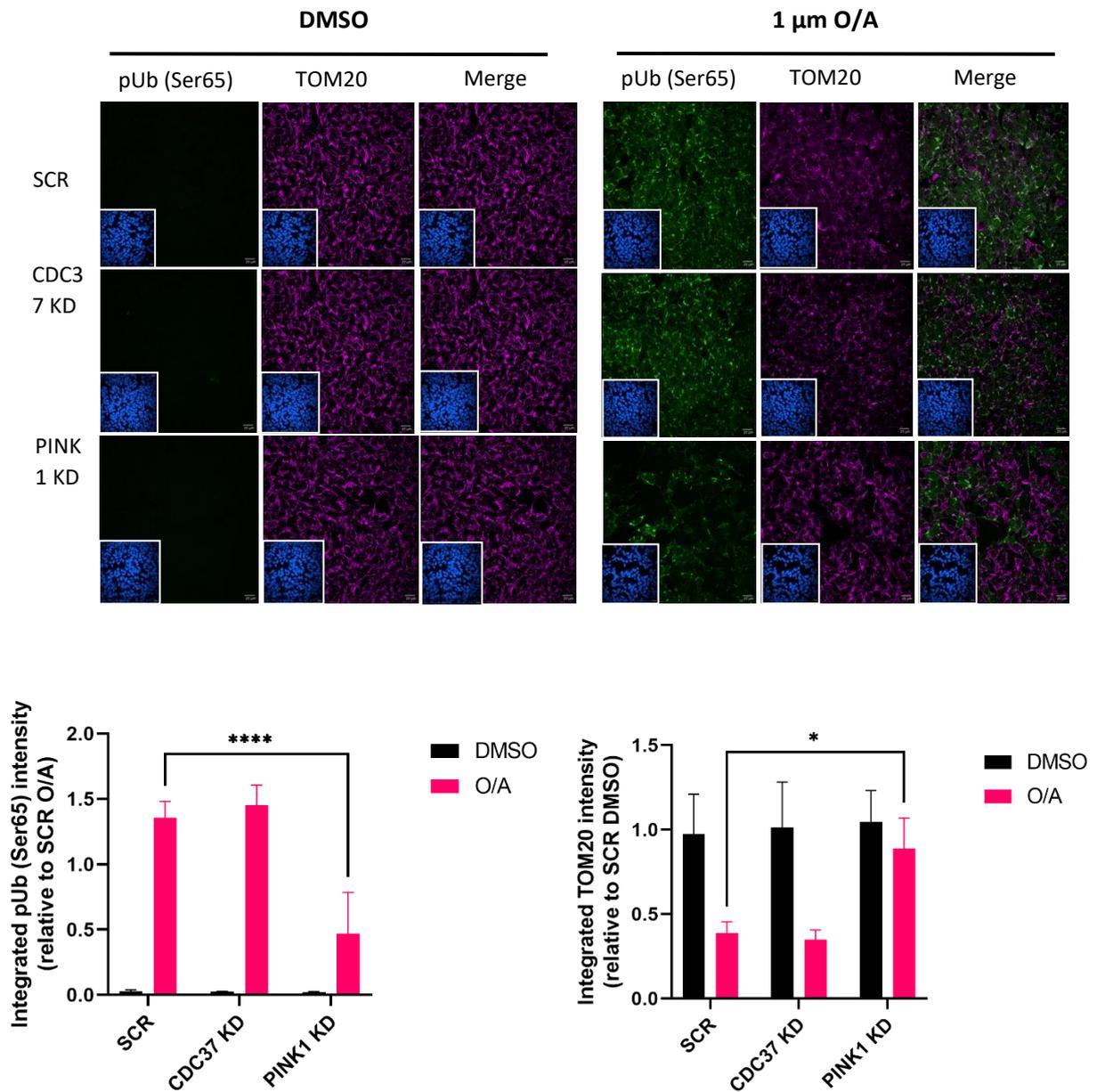


Figure 5.4. *CDC37* KD does not significantly increase pUb (Ser65) accumulation in POE SH-SY5Y cells (IF).

Representative IF images from SCR, *CDC37*, *PINK1* siRNA KDs in POE SH-SY5Y treated with 1 μM O/A for 3 h. Insets show the nuclei for the same fields, pUb signal is shown in green while TOM20 signal is displayed in magenta. pUb and TOM20 protein expression upon siRNA KDs were quantified by measuring the intensity of the fluorescence signal using the Columbus software and expressed as integrated intensity relative to the SCR negative control. Histograms indicate mean of n= 3 independent experiments ± StD. quantification data were normalised against non-transfected control and analysed with an ordinary two-way ANOVA with Dunnett's correction. * p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.0001 compared to scrambled control. Scale bar: 20 μm.

Given the decrease in *PINK1* protein expression following *CDC37* KD observed in mitochondria-enriched fractions, I further assessed whether *PINK1* mRNA levels could

be regulated by *CDC37*. To this end, I knocked down *CDC37* in POE SH-SY5Y cells before extracting RNA and performing qPCR. KD of *CDC37* significantly increased *PINK1* mRNA levels (

Figure 5.6), suggesting that *CDC37* KD may affect PINK1-mitophagy by modulating *PINK1* mRNA levels.

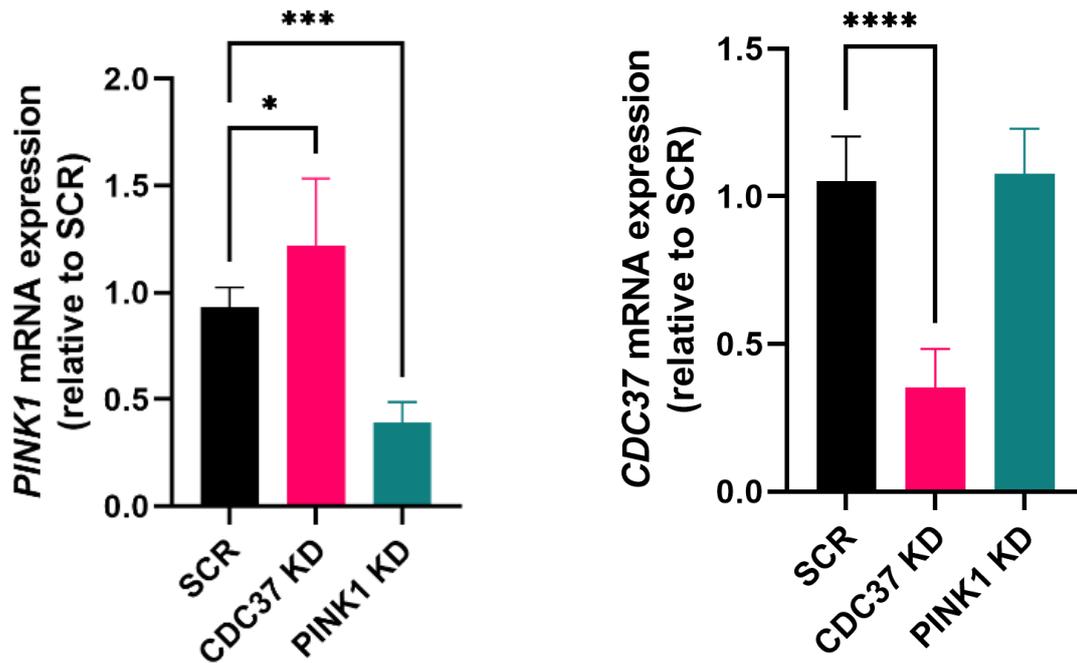


Figure 5.5. *CDC37* KD increases *PINK1* mRNA expression in POE SH-SY5Y while transcriptional levels of *CDC37* remain unchanged upon *PINK1* KD. Quantification of *PINK1* mRNA levels in SCR, *CDC37*, *PINK1* KDs. Histograms indicate mean of n=5 independent experiments \pm StD. quantification data were normalised against non-transfected control and analysed with an ordinary one-way ANOVA with Dunnett's correction. * p \leq 0.05, ** p \leq 0.01, *** p \leq 0.001 compared to scrambled control.

These qPCR data might explain the increase in pUb (Ser65) upon *CDC37* KD observed in the previous IB screening (**Figure 5.3**); an increase in *PINK1* mRNA expression following *CDC37* KD could result in an increased PINK1 protein expression, leading to increased PINK1 recruitment at mitochondria upon mitophagy initiation and, as a consequence, increased pUb (Ser65) recruitment at mitochondria, consistently with what observed in IB analyses. Additionally, analysis of *CDC37* mRNA expression upon *PINK1* KD showed no significant change, suggesting that *CDC37* is likely to act above of *PINK1* as opposed to the other way around.

5.1.2.2. *CDC37* KD does not increase pUb (Ser65) accumulation in H4 cell line

The effect of *CDC37* KD on pUb was also assessed in the human astrocytoma H4 cell line. Western blotting carried out by the HPF lab showed that the H4 cell line seems to have higher PINK1 protein availability when compared to the SH-SY5Y line, thus in this instance I was able to detect it more easily with the PINK1 antibody in WCL. No significant change in pUb levels was detected in *CDC37* KD cells when compared to SCR control in this cell model (

Figure 5.6), suggesting that increased levels of pUb following *CDC37* KD might be cell-type specific.

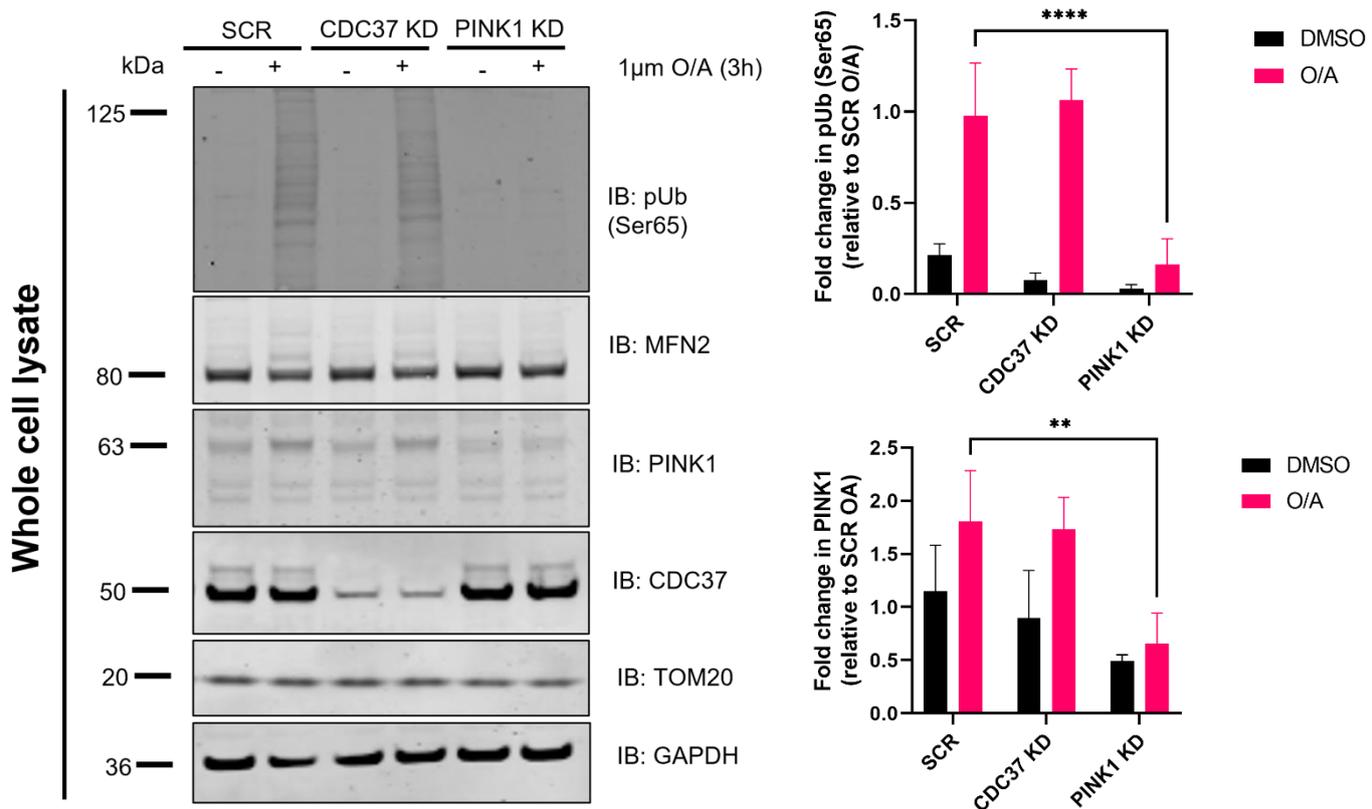


Figure 5.6. *CDC37* KD does not significantly increase pUb (Ser65) accumulation in H4 cells (WCL). Representative IB of WCL from SCR, *CDC37*, *PINK1* siRNA KDs in H4s treated with 1 µM O/A for 3 h. IB for phospho-ubiquitin (pUb), mitofusin 2 (MFN2), PTEN-induced kinase 1 (PINK1), cell-division cycle 37 (*CDC37*), translocase of the outer membrane protein 20 (TOM20) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) control was

performed. pUb protein expression upon siRNA KDs was quantified by densitometry and expressed as fold change relative to the SCR negative control. Histograms indicate mean of n=3 independent experiments \pm StD. quantification data were normalised against non-transfected control and analysed with an ordinary two-way ANOVA with Dunnett's correction. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$ compared to scrambled control.

The lack of effect of *CDC37* KD on pUb (Ser65) was mirrored by PINK1 IB, which showed no significant changes when compared to scrambled control.

5.1.2.3. *CDC37* KD increases mitochondrial clearance after 12h in mt-Keima but not in WB

In order to further explore the role of *CDC37* in the mitophagy pathway, mitochondrial clearance upon *CDC37* KD was assessed in live POE SH-SY5Y cells expressing the mitophagy reporter mt-Keima, a pH-sensitive, fluorescent protein that exhibits resistance to lysosomal degradation and that can be excited by two different wavelengths. At the physiological pH of the mitochondria (pH 8.0), the shorter-wavelength excitation predominates and is displayed in blue. Within the acidic lysosome when the pH lowers (pH 4.5) after mitophagy initiation, mt-Keima undergoes a gradual shift to the longer-wavelength excitation, which is displayed in red²⁴⁹. Therefore, mt-keima provides a measure of the mitophagic state at a single-cell level by comparing levels of free cytosolic mitochondria versus those in acidic lysosomes²⁵⁰. The rate of mitochondrial clearance for the different siRNA KDs was measured at five different timepoints (0, 3, 6, 9, 12h) of O/A treatment. In order to quantify mitochondrial clearance, the mitophagy index, namely the ratio between lysosomal and cytoplasmic mitochondria, was calculated across the five different timepoints (**Figure 5.7**).

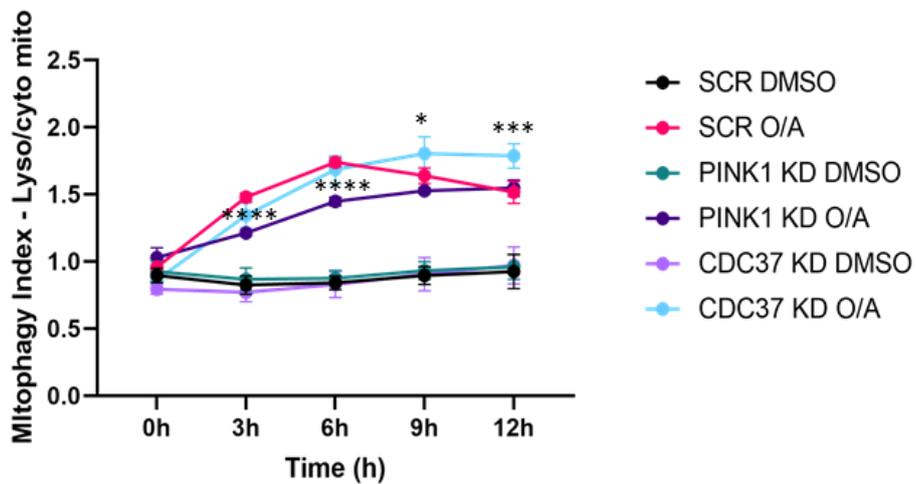
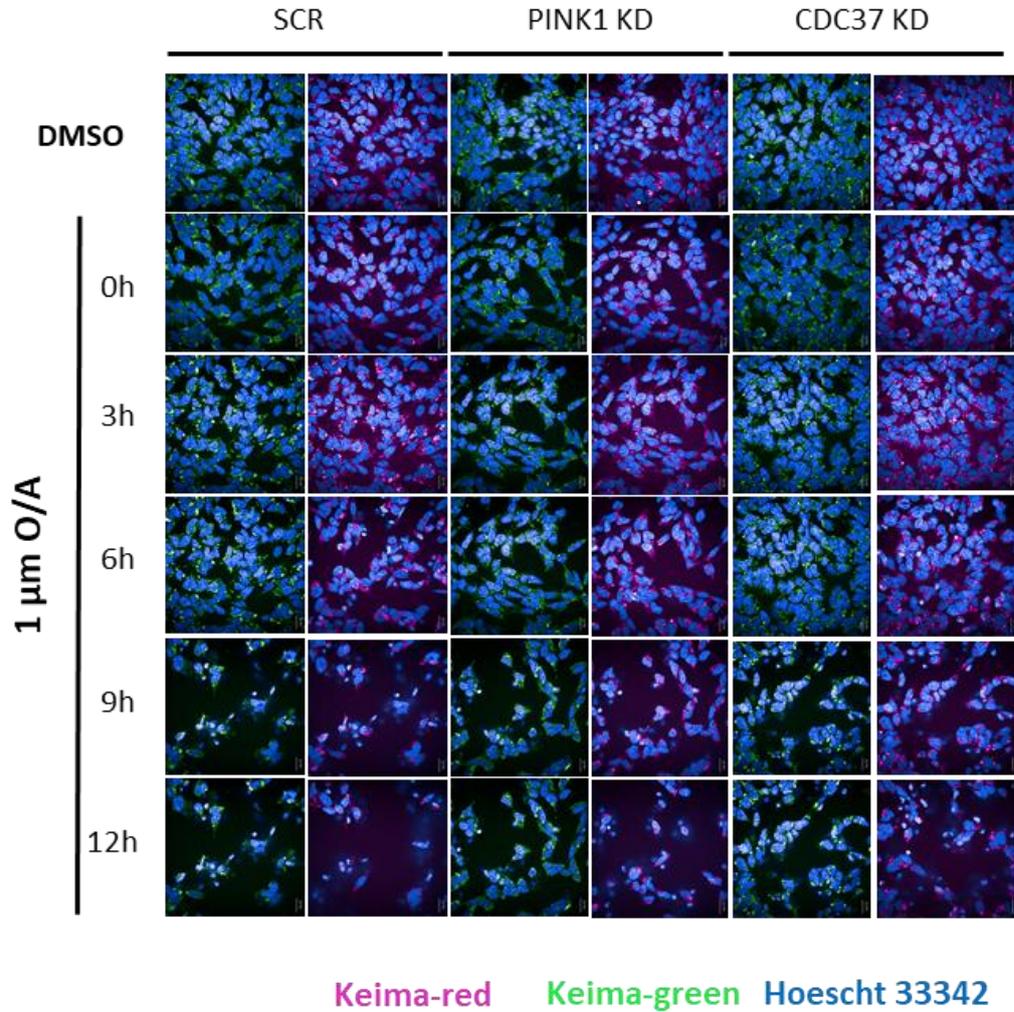


Figure 5.7. *CDC37* KD increases mitochondrial clearance. Representative images of Keima-green, Keima-red, and cell nuclei (blue) following siRNA KD of SCR, *PINK1*, *CDC37* KDs in POE keima-expressing SH-SY5Y cells treated with 1 μ M O/A for 0, 3, 6, 9, 12 h. Keima-green signal is shown in green while keima-red signal is displayed in magenta. The mitophagy index for each siRNA KD was quantified on the basis of the signal intensity and compared against SCR O/A. the plot indicates mean of n=5 independent experiments \pm StD. Quantification data were normalised against non-transfected control and analysed with an ordinary two-way ANOVA with Dunnett's correction. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$ compared to scrambled control. Scale bar: 20 μ m.

Mt-keima data showed a significant decrease of the mitophagy index after 3 and 6h O/A treatment upon *PINK1* KD, my positive control. The data also show that after 9 and 12h there is a significant difference in mitochondrial clearance between SCR and *CDC37* KD, further supporting the data previously obtained through IB and qPCR and reinforcing a potential role of *CDC37* in the regulation of *PINK1*-dependent mitochondrial clearance.

Since the increase in pUb (Ser65) was assessed only after 3h O/A treatment with IB, and the increase in mitochondrial clearance was observed after 9 and 12h only in live cells mt-Keima, I sought to investigate whether a significant increase in pUb (Ser65) levels could be also detected at different timepoints with IB. To this end, POE SH-SY5Y cells were transfected with *CDC37* siRNA and then treated with O/A for 3, 6, 9, 12h (**Figure 5.8**).

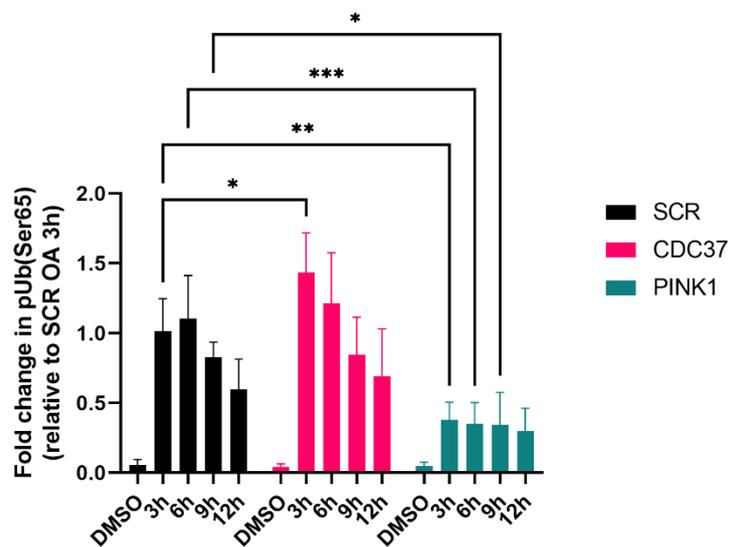
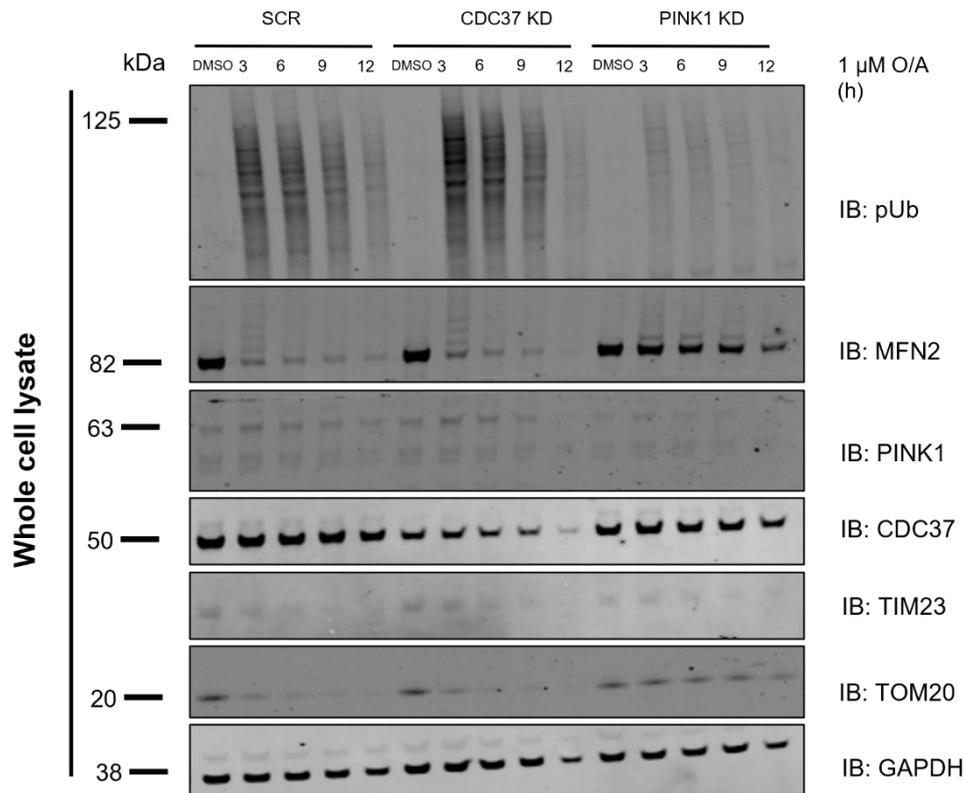
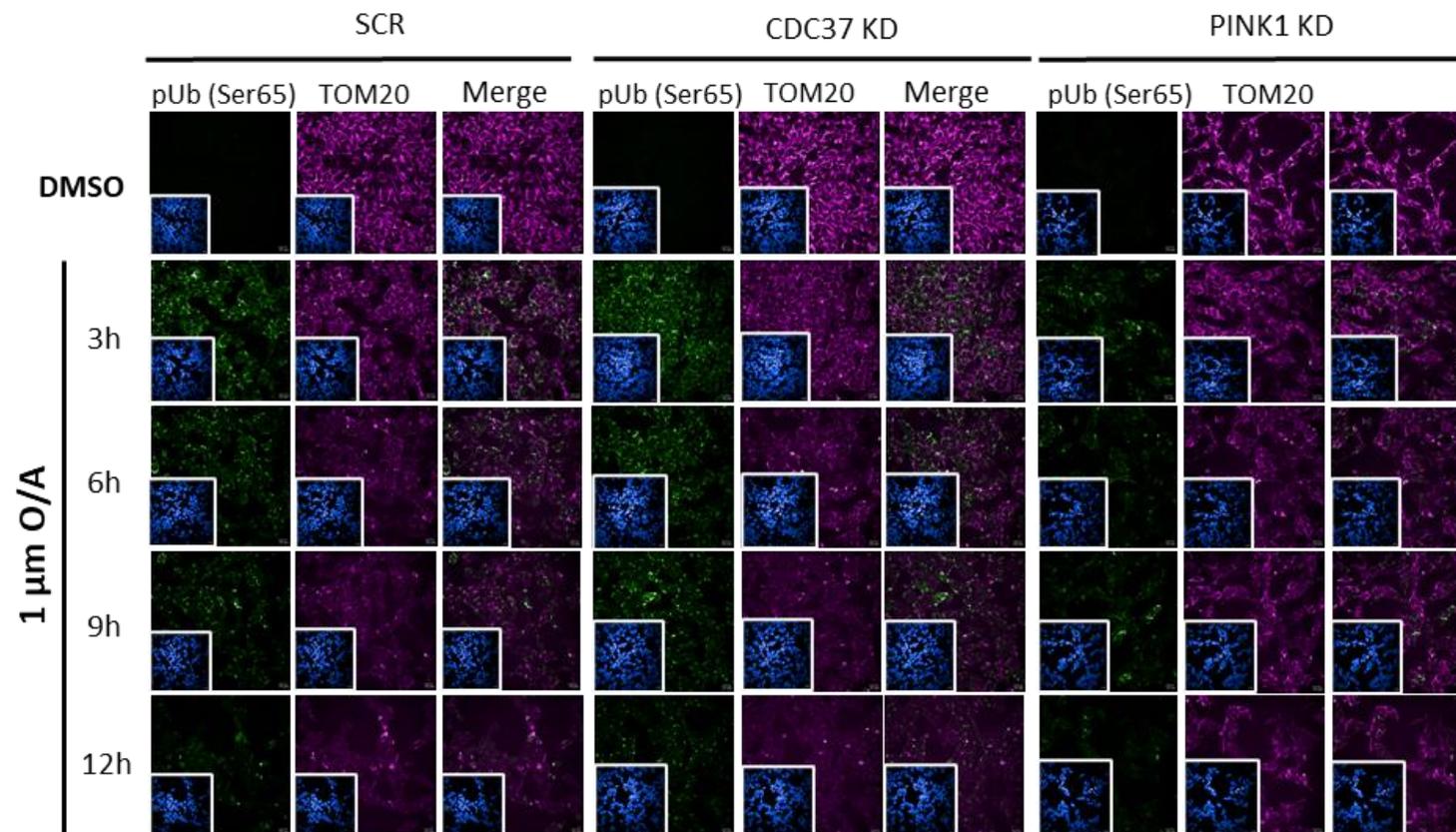


Figure 5.8. *CDC37* KD does not significantly increase pUb (Ser65) accumulation in POE SH-SY5Ys at different timepoints (WCL). Representative IB of WCL from SCR, *CDC37*, *PINK1* siRNA KDs in POE SH-SY5Y cells treated with 1 μ M O/A for 3 h. IB for phospho-ubiquitin (pUb), mitofusin 2 (MFN2), PTEN-induced kinase 1 (PINK1), cell-division cycle 37 (*CDC37*), translocase of the outer membrane protein 20 (TOM20) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) control was performed. pUb protein expression upon siRNA KDs was quantified by densitometry and expressed as fold change relative to the 3h SCR O/A negative control. Histograms indicate mean of n=3 independent experiments \pm StD. Quantification data were normalised against non-transfected control and analysed with an ordinary two-way ANOVA with Dunnett's correction. * p < 0.05, ** p < 0.01, *** p \leq 0.001 compared to scrambled control.

pUb (Ser65) IB showed significant increase in pUb (Ser65) after 3h O/A treatment, as previously shown in **Figure 5.2**, but was not significantly different at following timepoints.

I then sought to investigate whether a significant increase in pUb (Ser65) levels could be detected at different timepoints with IF. For this purpose, POE SH-SY5Y cells were transfected with *CDC37* siRNA and imaged following treatment with O/A at 3, 6, 9, 12h (**Figure 5.9**).



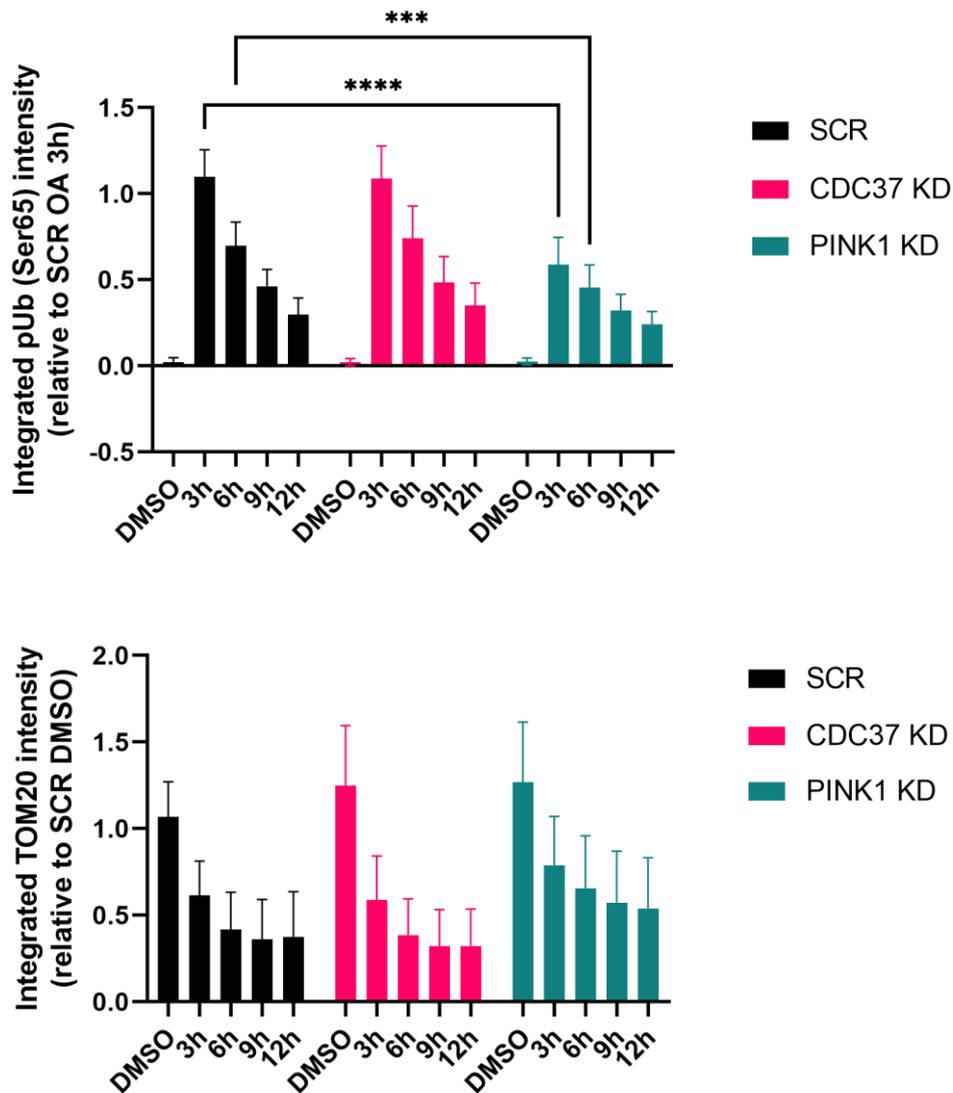


Figure 5.9. *CDC37* KD does not significantly increase pUb (Ser65) accumulation in POE SH-SY5Ys at different timepoints (IF). Representative IF images from SCR, *CDC37*, *PINK1* siRNA KDs in POE SH-SY5Y treated with 1 μ M O/A for 3, 3, 9, 12h. Insets show the nuclei for the same fields, pUb signal is shown in green while TOM20 signal is displayed in magenta. pUb and TOM20 protein expression upon siRNA KDs were quantified by measuring the intensity of the fluorescence signal using the Columbus software and expressed as integrated intensity relative to 3h SCR O/A. Histograms indicate mean of n= 3 independent experiments \pm StD. quantification data were normalised against non-transfected control and analysed with an ordinary two-way ANOVA with Dunnett's correction. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$ compared to scrambled control. Scale bar: 20 μ m.

Similarly to what observed in the previous IF experiment (**Figure 5.4**), no significant increase in pUb (Ser65) upon *CDC37* KD could be detected with IF at any timepoint (*PINK1* KD was used as a positive control) after 3h O/A induction. Collectively, these data suggest that *CDC37* KD leads to an increase in mitophagy initiation (pUb [Ser65]) after 3h

and a subsequent increase of mitochondrial clearance at 9 and 12h following O/A treatment. This is likely due to the fact that phosphorylation of ubiquitin at Ser65 occurs upstream of mitochondrial clearance, therefore implying that a phenotype observed upon mitophagy initiation is not necessarily mirrored upon mitochondrial clearance.

5.1.2.4. *CDC37* KD increases *GRN* mRNA expression but decreases its protein expression in FTLD iPSC model.

After having prioritised *CDC37* through cellular work as the most promising gene/protein candidate, I sought to assess its relevance in an FTLD context. Given the complexity of the genetic landscape of FTLD, generating or choosing relevant *in vitro* models of genetic FTLD is very challenging. In this specific instance, I selected *GRN* as likely relevant FTLD-risk gene to be modelled *in vitro* in combination with *CDC37* for a number of reasons: i) *GRN* was reported in literature to be implicated in the autophagy-lysosome pathway (ALP)^{32,79,81,91} as well as by unpublished studies to modulate mitophagy (Jacqueline Casey, UCL Institute of Neurology); ii) previous analyses performed on GCNs (see **Discussion** section in **Chapter 4**) indicated *CDC37* to be co-expressed with both *MAPT* and *GRN*, two of the most common Mendelian genes associated with FTLD pathology, making them interesting candidates to study *CDC37* implications in mitophagy in a FTLD-relevant model; iii) cortical neurons induced from an isogenic CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) series carrying a R493X *GRN* mutation were available to me thanks to the Wray lab (particularly in collaboration with Selina Wray and Jacqueline Casey, UCL), providing me with a comprehensive set of isogenic cell lines for FTLD investigation.

I performed individual siRNA KD of *GRN* and *CDC37* in POE SH-SY5Y cells and measured their relative mRNA expression. qPCR data showed that *GRN* mRNA expression significantly increases when *CDC37* is KD (**Figure 5.10**), suggesting that *CDC37* might play a role in regulating *GRN* transcriptional expression.

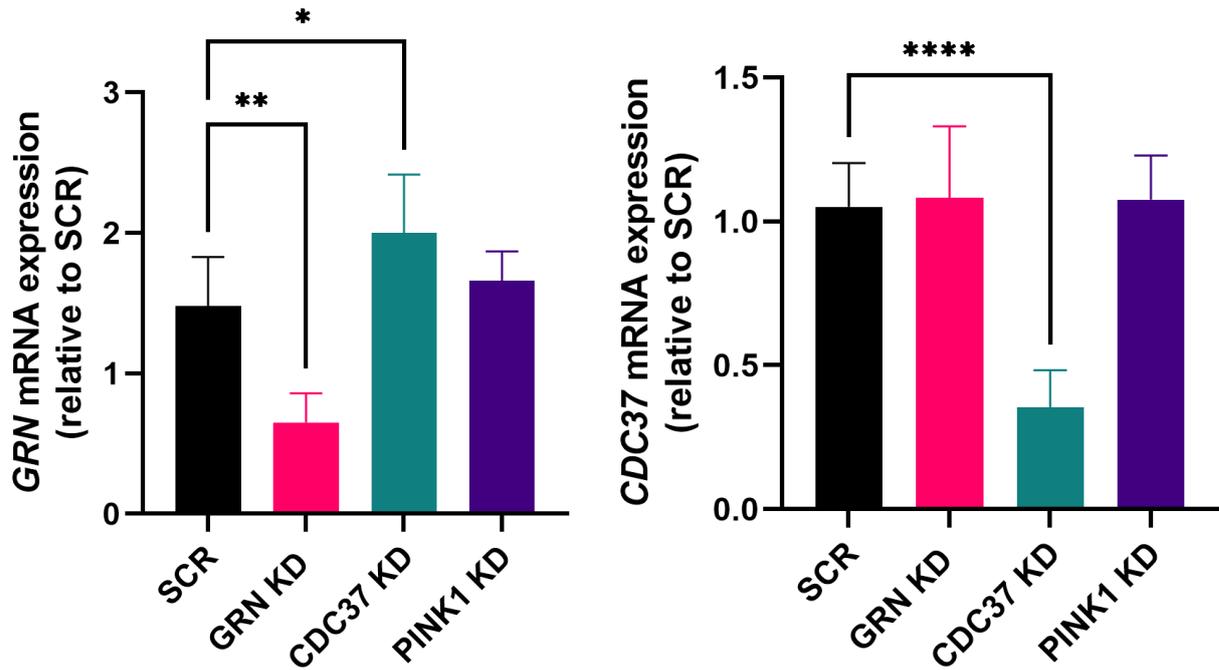


Figure 5.10. *CDC37* KD increases *GRN* mRNA expression in POE SH-SY5Y. Quantification of *GRN* mRNA levels in SCR, *CDC37*, *PINK1*, *GRN* KDs. Histograms indicate mean of n=5 independent experiments \pm StD. Quantification data were normalised against non-transfected control and analysed with a one-way ANOVA with Dunnett's correction. * p < 0.05, ** p < 0.01 compared to scrambled control.

By contrast, *CDC37* mRNA expression seems to not to be influenced by *GRN* silencing, identifying *CDC37* as a potential novel regulator of *GRN* mRNA.

I then sought to investigate PGRN involvement in PINK1-dependent mitophagy. To this end, I performed a *GRN* siRNA KD in mt-Keima-expressing POE SH-SY5Y cells to monitor the rate of mitochondrial clearance upon *GRN* deficiency (**Figure 5.11**).

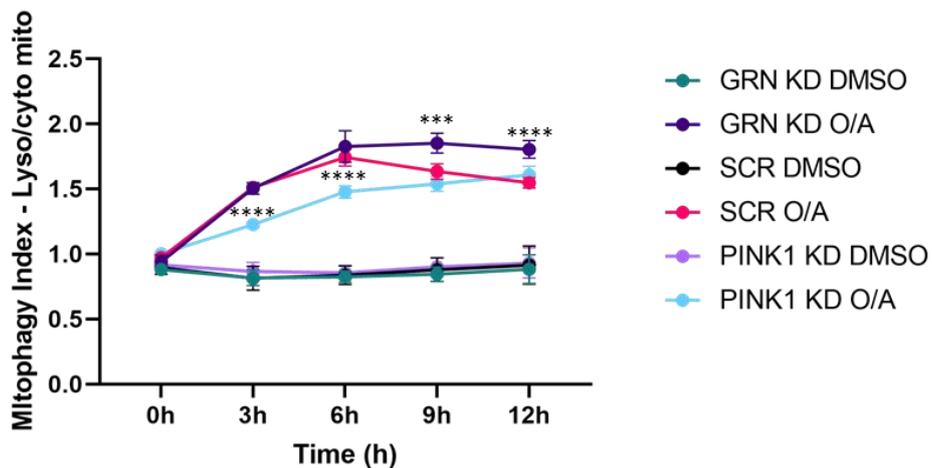
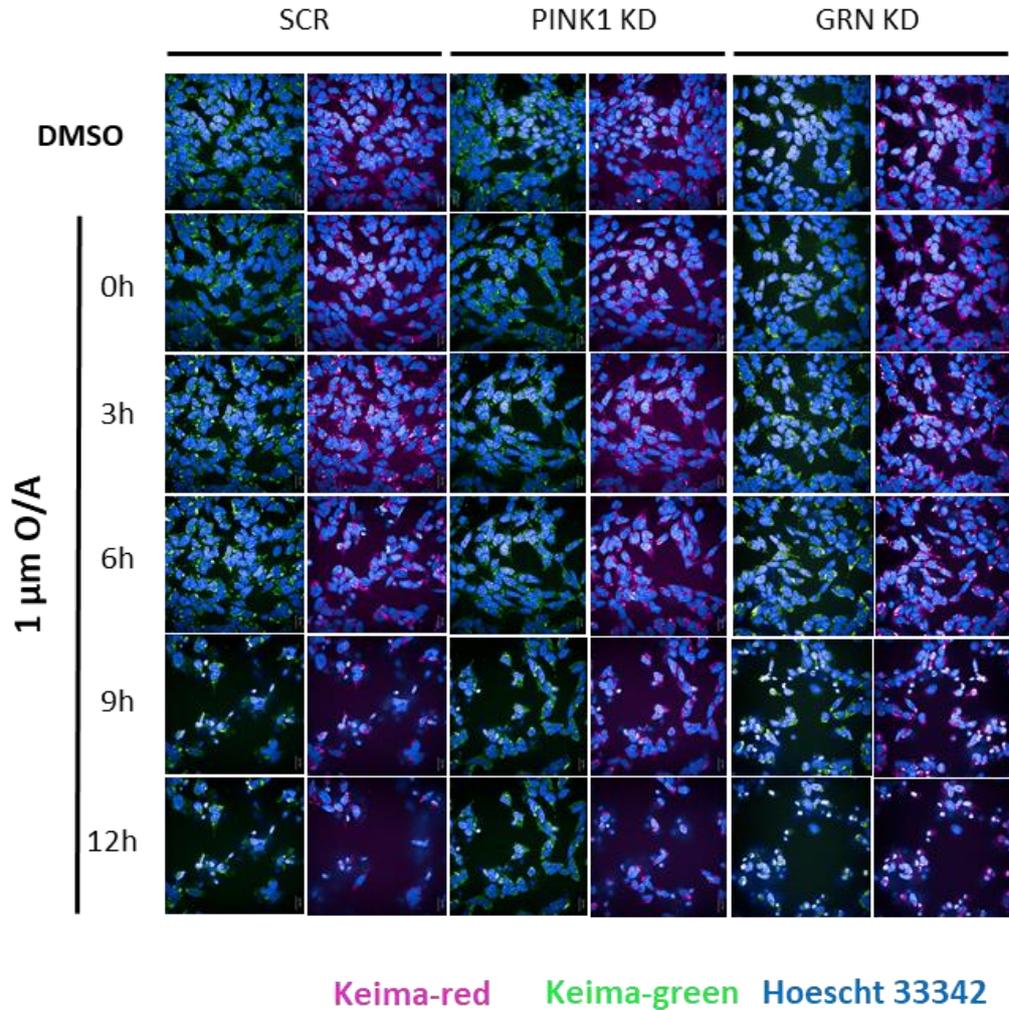


Figure 5.11. *GRN* KD increases mitochondrial clearance. Representative images of Keima-green, Keima-red, and cell nuclei (blue) following siRNA KD of *SCR*, *PINK1*, *GRN* KDs in POE keima-expressing SH-SY5Y cells treated with 1 μ M O/A for 0, 3, 6, 9, 12 h. Keima-green signal is shown in green while keima-red signal is displayed in magenta. The mitophagy index for each siRNA KD was quantified on the basis of the signal intensity and compared against *SCR* O/A. The plot indicates mean of $n=4$ independent experiments \pm StD. Quantification data were normalised against non-transfected control and analysed with a two-way ANOVA with Dunnett's correction. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$ compared to scrambled control. Scale bar: 20 μ m.

Again, mitochondrial clearance was significantly reduced after 3 and 6h O/A treatment upon *PINK1* KD, confirming successful KD of *PINK1*. I also observed that mitochondrial clearance upon *GRN* KD was significantly increased after 9 and 12h O/A treatment when compared to SCR O/A, implying that *GRN* may be involved in the selective clearance of dysfunctional mitochondria.

In order to further unravel the relationship between *CDC37* and *GRN* in an FTLD-relevant model, I used iPSC-derived cortical neuron derived from patients carrying *GRN* mutations. An isogenic CRISPR series (control, heterozygous and homozygous *GRN* R493X mutation) from the human iPSC Neurodegenerative Disease Initiative (iNDI) was available to me thanks to the work of Jacqueline Casey in the Wray's lab and was used in this study as an FTLD model. Neuronal differentiation was carried out following Shi et al.'s¹⁸² (see **Chapter 2** section **2.3.3.3** for more detail) dual SMAD (small mother against decapentaplegic) inhibition protocol. Cortical neuron characterisation was carried out by assessing the appearance of synapses by microscopy as described in the protocol. The R493X *GRN* mutation causes progranulin haploinsufficiency due to nonsense-mediated decay of mRNA, resulting in lysosomal dysfunction, increases in lipofuscin and thalamic neurodegeneration²⁵¹. The homozygous line has a bi-allelic R493X mutation, that should result in a complete knock-out (KO) of the progranulin protein.

The following work was performed by Jacqueline Casey and is still ongoing in the Wray lab, therefore quantification is not yet available for the IB of mitochondrial markers in iNDI iPSC neurons and in the H4 cell line. Data must therefore be interpreted cautiously. *CDC37* and mitophagy levels were assessed in iPSC-derived, *GRN*-deficient neurons using western blotting. To assess mitophagy, mitochondrial depolarisation was induced with 1 μ M O/A for 6, 12, 24h at 80 DIV and levels of *PINK1*, *MFN2* and *TIM23* were measured (**Figure 5.12**). Specifically, *MFN2* top bands show *PINK1*-dependent ubiquitination while *TIM23* is used as a mitochondrial marker (not yet degraded after 24h mitophagy induction); in this instance, β -actin was used as WCL marker instead of *GAPDH*.

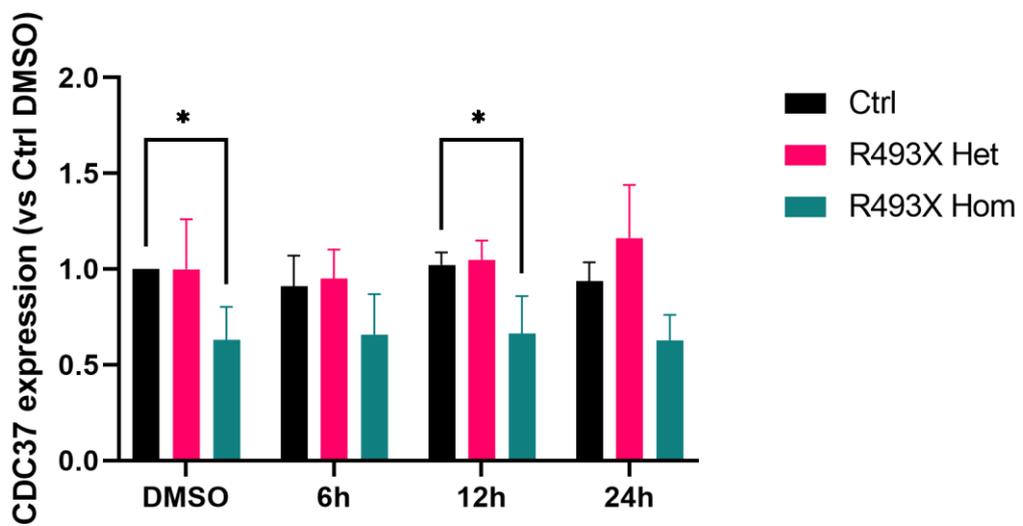
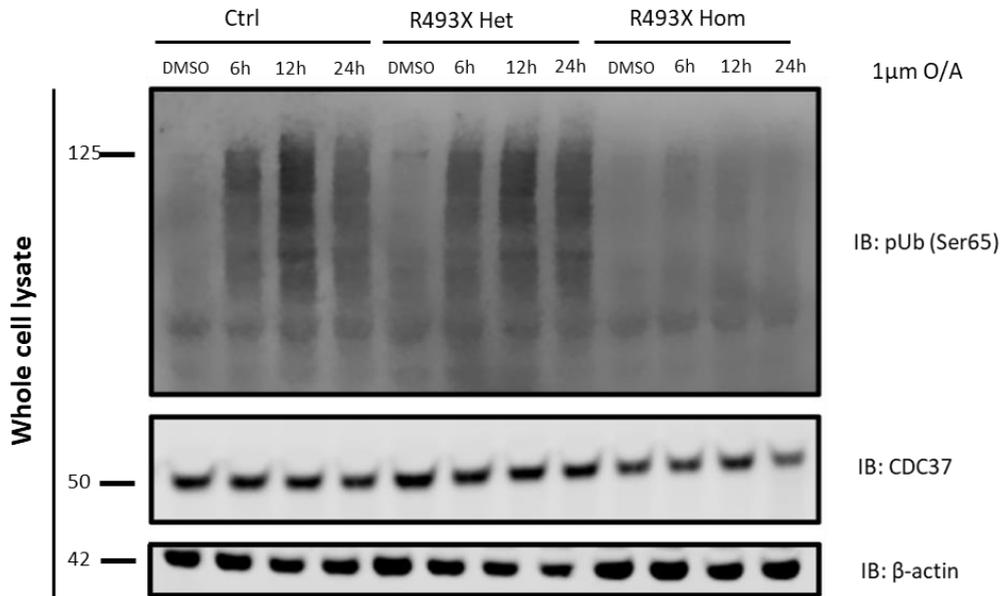
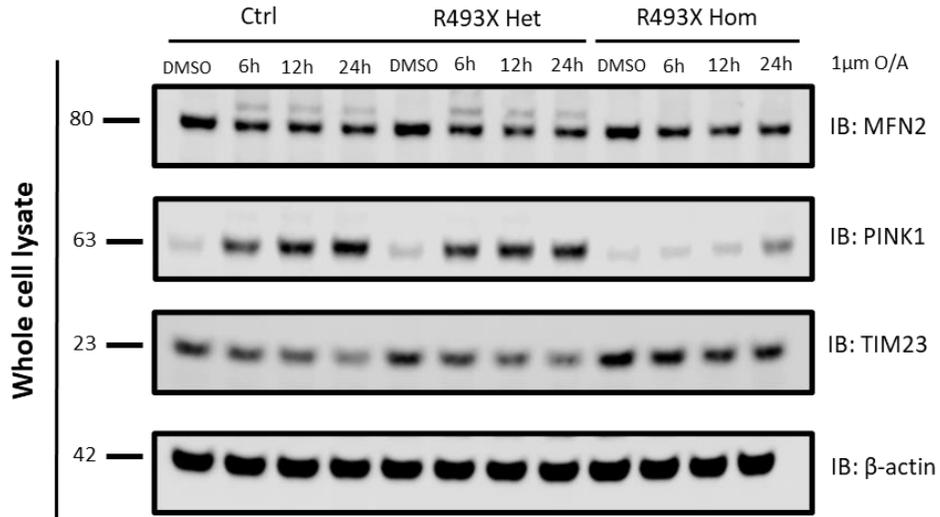


Figure 5.12. Mitophagy and CDC37 expression decrease in *GRN*^{-/-} iPSC-derived neurons (WCL). Representative IB of WCL from an isogenic *GRN* mutation CRISPR series from the iNDI (control, heterozygous and homozygous *GRN* R493X mutation). IB was performed for the mitochondrial markers mitofusin 2 (MFN2), PTEN-induced kinase 1 (PINK1), translocase of the inner membrane 23 (TIM23), phospho-ubiquitin (pUb, Ser65) and for the candidate gene/protein cell-division cycle 37 (CDC37). CDC37 protein expression was quantified by densitometry and expressed as fold change relative to the control DMSO. Histograms indicate mean of n=3 independent experiments ± StD. Quantification data were normalised against the control and analysed with a two-way ANOVA with Dunnett's correction. * p < 0.05 compared to control DMSO. These data were generated by Jacqueline Casey (Wray lab, UCL) and quantification for the mitochondrial markers was not available.

A significant reduction in PINK1 and pUb (Ser65) accumulation and decreased ubiquitination and degradation of MFN2 were observed in the homozygous *GRN*^{R493X/R493X} neurons with reduced progranulin, implying a downregulation of PINK1-dependent mitophagy upon *GRN* complete loss, as previously suggested by Casey's unpublished data. No significant difference in mitophagy regulation was detected between control *GRN*^{+/+} and heterozygous *GRN*^{+/R493X} iNDI iPSC neurons, which may be due to the upregulation of progranulin protein from the wild-type allele. These results suggest that PGRN might play a role in mitophagy by regulating stability and/or activity of PINK1. CDC37 IB revealed a significant reduction in its expression in the homozygous *GRN*^{R493X/R493X} neurons, while no significant change was observed in the heterozygous mutants when compared to control. As previously discussed in section 5.1.1.1 of this Chapter, CDC37 has been shown to influence the maturation and subcellular distribution of PINK1^{215,216}. With this knowledge in mind, these results might imply that PGRN plays a role in mitophagy by regulating the trafficking and stability of PINK1 at mitochondria via CDC37.

Interestingly, similar levels of expression in control and heterozygous *GRN* mutants (i.e., in mitophagy markers PINK1, pUb as well as CDC37) suggest that the heterozygous R493X iNDI line could have a potentially compensatory upregulation of progranulin protein levels despite reduced *GRN* mRNA levels.

The relationship between mitophagy, PGRN and CDC37 accumulation was further explored in an H4 *GRN* siRNA KD astrocytic model, also available to me thanks to Jacqueline Casey. H4s were transfected with *GRN* siRNA to perform its KD and they were probed for mitochondrial markers and CDC37 (**Figure 5.13**).

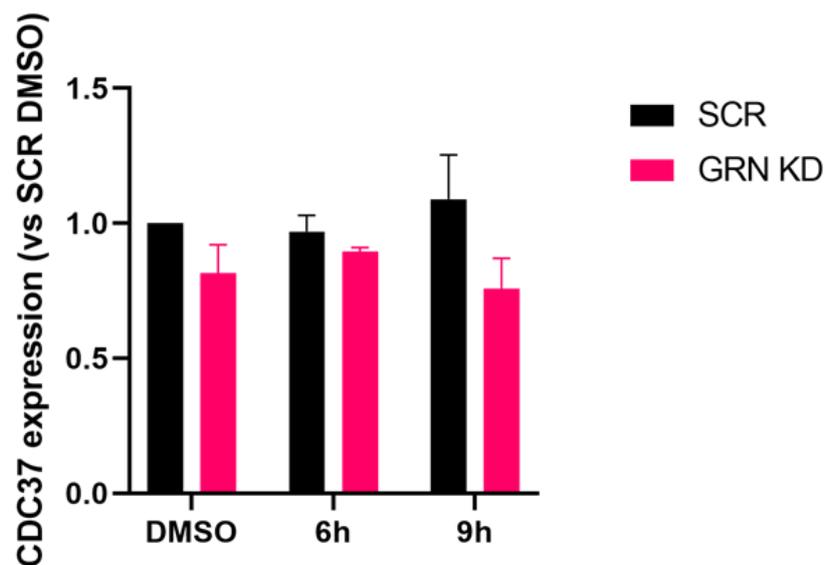
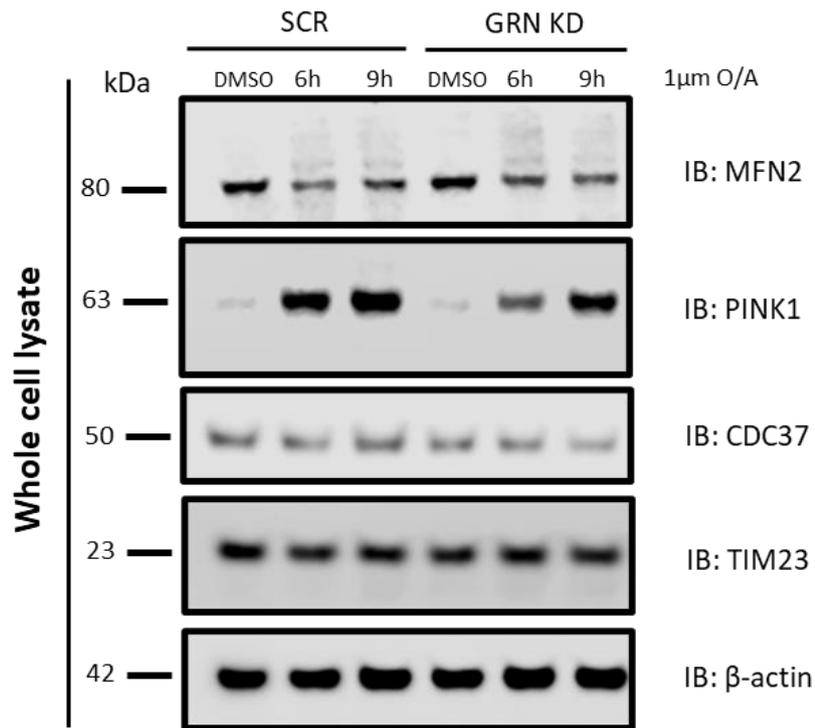


Figure 5.13. CDC37 expression is not significantly decreased in *GRN* siRNA KD H4s (WCL). Representative IB of WCL from H4s treated with *GRN* siRNA. IB was performed for the mitochondrial mitofusin 2 (MFN2), PTEN-induced kinase 1 (PINK1), translocase of the inner membrane 23 (TIM23), phospho-ubiquitin (pUb, Ser65) and for the candidate gene/protein cell-division cycle 37 (CDC37). CDC37 protein expression was quantified by densitometry and expressed as fold change relative to the scrambled DMSO negative control. Histograms indicate mean of n=3 independent experiments ± StD. Quantification data were normalised against the scrambled DMSO and analysed with a two-way ANOVA with Dunnett's correction. These data were generated by Jacqueline Casey (Wray lab, UCL) and quantification for the mitochondrial markers was not available.

A decrease in PINK1 accumulation and increased ubiquitination and degradation of MFN2 were observed in the H4 model as well, suggesting an overall decrease in mitophagy. Although not significant, a reduction in CDC37 was also observed in H4s upon *GRN* siRNA KD ($p=0.0865$), further suggesting a role for *GRN* in regulating mitophagy and CDC37 expression.

5.1.3. Discussion

In this study, I implemented a novel bioinformatics pipeline to prioritise FTLD-risk candidate genes and molecular pathways. In order to identify novel risk genes for FTLD, I used a combination of multi-omics approaches based on a selection of FTLD Mendelian, FTLD-ALS spectrum and GWAS candidate genes: i) WPPINA and ii) WGCNA. 13% of these genes were selected through multiple techniques, and three out of 157 genes/proteins (STUB1, CUL2, CDC37) were identified through a prioritization pipeline (see **Chapter 4 Results** section). These three genes were taken forward for functional analysis. BPs associated with *waste disposal*, in particular autophagy/mitophagy”, were enriched in FTLD-associated PPI and GC networks

Damaged organelle turnover (e.g., autophagy) and mitochondrial integrity are critical for cellular health and maintenance, and as a result their impairment underlies the onset of several neurodegenerative diseases (NDs). Among others, oxidative stress has been shown to be strongly implicated in PD, AD, ALS and FTLD pathogenesis, and mitochondrial damage and/or dysfunction represent major sources of such stress²⁵²⁻²⁵⁴. Several studies have shown that alterations in the mitochondrial protein PINK1 compromise the removal of mitochondria by mitophagy and are associated to recessive PD^{87,127,134,138,255}. In this scenario, while the PINK1-dependent mitophagy pathway has been well characterised, its transcriptional and molecular regulation as well as its implications in other NDs (e.g., FTLD) are still widely unexplored.

On the basis of the bioinformatics results obtained in this study, I sought to characterise the relationship between the three candidate genes/proteins and PINK1-dependent mitophagy. These data show that out of the three candidate genes/proteins, *CDC37* in particular modulates *PINK1* nuclear transcription and subsequent translation, suggesting its role as potential regulator of PINK1-dependent mitophagy. I also showed that

depletion of *CDC37* causes significant increase in the accumulation of the mitochondrial marker phospho-ubiquitin (pUb; Ser65) under mitochondrial stress conditions, ultimately improving mitochondrial clearance. Finally, both a transcriptional and a functional link were identified between *CDC37* and the lysosomal protein *PGRN*, a known FTLD Mendelian gene, providing a potential mechanistic insight on the molecular underpinnings of FTLD pathogenesis.

These data raise several intriguing questions. How do we explain pUb (Ser65) increase when *CDC37* is depleted? What is the relationship between *CDC37* and *PINK1* in the context of mitophagy? What is the mechanism through which *CDC37* and *GRN* could lead to neurodegeneration and FTLD pathology?

qPCR data suggest that transcriptional regulation of *PINK1* by *CDC37* is a possibility, implying a potential mechanism for regulation of *PINK1* trafficking by *CDC37*. This hypothesis is reinforced and cross-supported by gene co-expression network analysis showing *PINK1* and *CDC37* sharing transcriptional and thus functional overlap (refer to section 4.2.5) and by several studies that highlight *PINK1* as a client kinase of the Hsp90/*CDC37* complex^{215,216,256}. Since the identification of the serine/threonine kinase *PINK1* as a Parkinson's disease (PD)-associated gene in 2004, a lot of efforts have been directed towards identifying its biological role in health and disease to aid therapeutic advances for the treatment of PD¹³⁸. Several studies revealed *PINK1*'s pivotal role in mitochondrial removal via autophagy and showed that its interaction with a number of chaperone proteins is key for its stability and proper functioning¹³⁸. Chaperone proteins are the first line of defence against protein misfolding and degradation and therefore in this context are likely to be key players in disease pathogenesis¹³⁸. Moriwaki et al. and others^{138,215,216} reported that the interaction between *PINK1* and the Hsp90/*Cdc37* molecular chaperones complex is required to: i) stabilise the full-length *PINK1* at the OMM and the 42 kDa cleavage product; ii) drive the subcellular distribution of *PINK1* and ii) regulate the ratio of full-length to cleaved protein (63/55 kDa ratio). Therefore, it is possible that *PINK1* dependency on the interaction with the Hsp90/*Cdc37* complex could explain *PINK1* and pUb accumulation at mitochondria, as observed in the present study, and as a result enhanced mitophagy. Surprisingly, IF assays did not confirm IB findings: future studies will have to be designed to validate the accumulation of *PINK1* and pUb observed through IB. Interestingly, Watzlawik et al.²⁵⁷, recently developed a sandwich

enzyme-linked immunosorbent assay (ELISA) targeting pUb (Ser65) specifically, which is able to assess very low baseline mitophagy levels and that could be useful to confirm pUb (Ser65) changes upon *CDC37* KD.

Finally, I sought to confirm whether *CDC37*-dependent accumulation of PINK1 and pUb at damaged mitochondria might be specific to neuronal populations or could be observed in different cell types, such as astrocytes. For this reason, I used an H4 astrocytic model to investigate *CDC37* implications in PINK1-dependent mitophagy: interestingly, no significant changes in either PINK1 or pUb (Ser65) could be detected in this instance, suggesting this phenotype to be cell-type specific. Future studies will have to determine whether this phenotype is specific to neurons by looking at, for instance, primary neurons from *GRN* KO mice or CRISPR neurons; additionally, other iPSC neuronal types should be evaluated to confirm whether this phenotype can be observed in cortical neurons only.

Furthermore, since mitophagy is a multistep process, I looked towards the progressive effects of mitochondrial toxicity and subsequent degradation in live cells, to evaluate mitochondrial clearance (mitophagy endpoint) at later timepoints. Live cell imaging with mt-keima reporter displayed significantly increased mitochondrial clearance upon *CDC37* silencing after 9 and 12h mitophagy induction. The greater mitochondrial turnover observed in live cells after 9 and 12h can be interpreted as a direct consequence of increased accumulation of pUb at mitochondria after 3h, implying that mitochondrial clearance can be observed following several hours of O/A induction. Interestingly, Evans et al.²⁵⁸ reported that in basal and induced conditions, engulfed mitochondria tend to remain in non-acidified organelles for hours to days, illustrating fast and efficient autophagosome sequestration but delayed lysosomal fusion or acidification, which might explain why mitochondrial clearance can only be observed several hours after mitophagy initiation.

Notably, most of these data were obtained from POE SH-SY5Y, which was chosen because of its suitability to study PINK1-dependent mitophagy. In fact, endogenous Parkin expression is generally quite low: extensive western blotting carried out by the HPF lab showed that assessing downstream steps of mitophagy process (e.g., ubiquitination of proteins, degradation of mitochondria) is very challenging in wild-type SH-SY5Ys. However it is important to note that protein over-expression can often generate artifacts: POE will likely 'prime' the cells towards preferentially utilising the PINK1/Parkin

pathway of mitochondrial degradation, thus potentially masking any impairments in alternate pathways of mitochondrial quality control towards Parkin-dependent processes²⁴⁴. Therefore, these findings will need to be replicated in cell lines expressing endogenous Parkin as well as in other cell types (e.g., neurons, glia).

Since an interesting CDC37-associated phenotype was found in a set of immortalised cell lines (i.e., POE SH-SY5Y and H4 lines), I sought to confirm these findings in an experimental system that recapitulates FTLD genetic and pathological signatures. The *GRN* gene was chosen as candidate FTLD Mendelian gene to be modelled *in vitro* based on literature and on hypothesis-driven bioinformatics analyses performed in this study. Growing evidence suggests a role for PGRN in the lysosome^{82,121,259}, which includes regulation of lysosomal trafficking pathways and both direct and indirect regulation of lysosomal hydrolases biosynthesis, globally being responsible of lysosome-related mechanistic dynamics⁸². Following an assessment of transcriptional expression of *GRN*, qPCR data showed that *GRN* mRNA expression is significantly upregulated upon *CDC37* siRNA KD in POE SH-SY5Y cells. These data imply that not only *CDC37* is involved in *PINK1* transcriptional regulation but also in *GRN*'s, making *CDC37* a potential master regulator of the ALP by modulating the rate of mitochondrial degradation at different stages in the pathway. Additionally, assessment of mitochondrial clearance in live cells showed that *GRN* KD increased mitochondrial turnover at both 9 and 12h timepoints after mitophagy induction, suggesting an increased ratio of mitochondria in acidic lysosomes. This phenotype could reflect an increase in lysosomal clearance for the end stage of mitophagy.

However, how PGRN loss alters lysosomal functions it is still unclear. Given the importance of PGRN in the maintenance of healthy functional lysosomes as well as its role in regulating differential expression of lysosomal hydrolases, one could argue that *GRN* deficiency might impair effective intralysosomal processing of digested material, such as in the case of many lysosomal storage disorders (LSDs), resulting in aberrant accumulation of undegraded substrate within the lysosome²⁶⁰⁻²⁶². As a result, the increase in lysosomal acidification caused by *GRN* deficiency at late timepoints might take place concurrently to impaired lysosomal enzymatic function, leading to suppressed digestion in and efflux from lysosomes²⁶².

Intriguingly, a study by Beel et al.²⁶³ showed that PGRN may directly bind to and modulate lysosomal enzymes, such as cathepsin D, thus suggesting that PGRN might regulate the rate of lysosomal digestion of cargo⁸¹. This finding could be suggestive of a time-dependent inability of the cell to efficiently recycle the dysfunctional mitochondria upon *GRN* loss following continued exposure to mitochondrial toxin O/A. From a global disease-oriented perspective, it would appear that mutations or removal of *GRN* might be responsible for exacerbating the slow kinetics of neuronal mitophagy, leading to disrupted mitochondrial network integrity, oxidative stress and eventually neurodegeneration. This hypothesis is in line with previous research which showed *GRN* deficiency to alter neuronal lysosome abundance, morphology and functionality by impairing lysosomal lipid metabolism²⁶⁰. Importantly, further work is needed to confirm impairment of intralysosomal processing, for instance by evaluating later timepoints in mt-keima (e.g., 24 and 36h). Altogether, these results strongly support the role of PGRN in mitophagy regulation and future studies should aim to understand whether cellular homeostasis is able to rescue PGRN deficit and carry out efficient lysosomal digestion.

Mitophagy and CDC37 accumulation were then examined in an H4 *GRN* KD model and in an isogenic CRISPR series carrying the *GRN*-associated R493X mutation. CDC37 protein levels were found to be significantly reduced during mitophagy in R493X homozygous mutation neurons, while no effect was observed in H4 astrocytomas. These results have a number of implications: i) PGRN plays a role in neuronal mitophagy, in particular PGRN complete deletion seems to downregulate the mitophagy pathway by decreasing PINK1 and subsequent pUb (Ser65) accumulation at mitochondria; ii) CDC37 expression is also regulated by PGRN, as its expression decreases upon *GRN* deficiency; iii) collectively, these findings seem to point towards a shared neuron-specific mechanism between PGRN, PINK1 and CDC37, where *GRN* modulation might be regulating the trafficking and stability of PINK1 via CDC37. Recent studies had already pointed at a potential involvement of PGRN in mitophagy, but a solid link between the two has not been established yet^{264,265}. Zhou et al.²⁶⁴ found that PGRN deficiency exacerbated mitochondrial damage and dysfunction in a murine diabetic nephropathy model, showing that basal PGRN levels have protective effects by attenuating mitochondrial dysfunction and enhancing mitochondrial biogenesis; Chang et al.²⁶⁵ showed that *GRN* KO mice exhibit reduced xenophagy (i.e., the selective clearance of intracellular pathogens) which notably relies on a set of proteins and chaperones that also regulate mitophagy (e.g., OPTN, Rab7,

TBK1 etc.). The present study is among the first studies showing PGRN implications in neuronal PINK1-dependent mitophagy and suggesting a molecular mechanism via CDC37.

These results support a role for *CDC37* as a gene regulating the activity of both *PINK1* and *GRN* at the transcriptome level as well as PINK1 stability and trafficking at mitochondria in neurons.

Further studies should aim at further characterising PGRN implications in a *CDC37* KD model to confirm transcriptional data obtained through qPCR and at further dissecting cell-type specific contributions of PGRN to mitophagy in iPSC-derived neurons, astrocytes and microglia. Additionally, as *CDC37* KD seems to enhance mitophagy, future studies will need to determine whether CDC37-dependent modulation of PINK1-dependent mitophagy is able to rescue mitophagy-deprived disease phenotypes. More work is needed to explain why no effect on pUb (Ser65) accumulation was shown in IF experiments upon *CDC37* KD; in this instance, it might be useful to: i) evaluate more timepoints; ii) increase the number of biological replicates to increase the statistical power of the analysis; iii) confirm *CDC37* KD in a neuronal model (e.g., CRISPRi, etc.).

In the past ten years, numerous studies have identified pathogenic mitochondrial pathways in FTLN-ALS and have suggested that mitochondrial defects could be one of the primary causes of FTLN, ALS and related diseases; in particular, mutations in the *TBK1*, *CHCHD10* and *OPTN* genes have shown to prevent efficient mitochondrial engulfment and translocation^{88,90,266,267}, making them suitable candidates for investigation in the context of FTLN pathogenesis. Clearly to address these gaps there is still need for a better molecular understanding of mitochondrial quality control and more precise clinical phenotyping.

To summarise, these findings implicate mitophagy and *CDC37* as novel therapeutic targets for *GRN*-associated FTLN and highlight the emerging theme of defective autophagy in the broader FTLN-ALS spectrum of neurodegenerative disease.

Chapter 6

6. Discussion and future directions

After over a century since FTLN was first described, our understanding of the pathogenic mechanisms that underlie the disease has substantially improved but it has failed to translate into effective therapeutic strategies. As discussed in **Chapter 1**, FTLN appears to result from a combination of defective pathways arising from both genetic and environmental factors. *Waste disposal*, and mitochondrial dysfunction as previously discussed, is by no means the only biological process impacted in disease pathogenesis (see section **1.1.4**). However, it has been strongly implicated in both familial and sporadic FTLN and as such has been previously suggested as a potential therapeutic target^{9,79,97}. Looking to the future, it is useful to consider which questions still need to be answered, and how existing knowledge may be applied to treating the disease. In this thesis, on the basis of known Mendelian genes, I presented a bioinformatics multi-omics pipeline aimed at defining and prioritising biological processes and candidate genes/proteins impacted in FTLN on the basis of known Mendelian genes, to be carried forward for functional validation in disease-relevant models. This final chapter will be considering major implications, limitations and future directions to be undertaken following what has been shown in the present work.

6.1. General limitations

The current study presents a newly developed approach providing a novel insight on the biological processes and molecular drivers underlying a multifactorial disease (i.e., FTLN). Despite the promise of this approach, it is important to mention that several limitations apply.

Time always is a major limitation when experimental biology is involved. While the bioinformatics pipeline was successful in identifying candidate biological processes and genes/proteins impacted in FTLN, experimental validation was conducted only in immortalised cell lines (e.g., SH-SY5Y and H4s), which do not recapitulate FTLN markers and pathological features. In this instance, validation in a disease-relevant model would be essential for an adequate evaluation of the pipeline performance. While the work conducted by J. Casey has been helpful to further define the relationship between CDC37

(candidate gene/protein), GRN (FTLD) and PINK1 (mitophagy), further experiments conducted in disease-relevant models would be required to better characterise the link between the prioritised gene/protein and mitophagy, as well as their implications in FTLD.

Another important limitation applying to my approach is represented by the controversial definition of genes 'associated to FTLD'. Defining what genes can be considered Mendelian (outside of the three most frequently mutated genes i.e., *MAPT*, *GRN* and *C9ORF72*) is very challenging, especially for such a genetically, pathologically, clinically heterogenous disease. In order to attempt to overcome this issue, I used the most recent FTLD review at the time (i.e., February 2019; Ferrari et al.⁹) to select the most up-to-date FTLD-risk gene list known to date. Notably, these bottlenecks can be partially considered as a direct consequence of the current degree of characterisation of FTLD etiology, and thus potentially inherent to FTLD itself. In fact, FTLD being a rarer condition when compared to other neurodegenerative diseases, such as AD and PD, proportionately less efforts and resources have been directed towards elucidating the pathogenic mechanisms underlying disease and identifying candidate therapeutic targets. It is notable that only five FTLD GWASs⁶⁴⁻⁶⁸ have been conducted in the past two decades (see section 1.1.3.2), which generally included hundreds to few thousands of cases-controls, against over 20 GWAS studies for AD²⁶⁸⁻²⁷⁴ and nearly ten for PD²⁷⁵⁻²⁸², which by contrast present with several thousands to a million cases each. Additionally, availability of FTLD RNA expression datasets is very limited, making it increasingly difficult to investigate variability in gene expression in health and disease as well as to determine genes and pathways involved in the disease process. Finally, the considerable size of the FTLD clinical spectrum as well as its significant overlap with ALS add an extra level of complexity when trying to better define and characterise a single disease.

Altogether, these factors contribute to the several challenges encountered when approaching and investigating FTLD and generally neurodegeneration. Fortunately, the rise of new sequencing and omics technologies as well as the creation of larger sample sets will increasingly overcome many of these limitations, allowing for a better definition of the genetic and molecular components of disease.

6.2. Future directions in FTLD research

Following directly from this thesis, the directions for future research are multi-faceted. From a bioinformatics perspective, the multi-omics approach used in the present study has proven to be particularly useful in dissecting complex traits such as neurodegenerative diseases, that present with very heterogeneous clinical and molecular phenotypes. Importantly, it paves the way to novel pipelines that aim at integrating multidisciplinary fields and orthogonal data to identify disease candidate markers and therapeutic targets by evaluating the global aspects of the disease. Fine-tuning of the pipeline to increase its power to identify true disease markers will be essential by: i) integrating additional layers of information (e.g., evaluation of cell-type specificity expression, eQTLs etc.) to increase the resolution and accuracy of candidate gene-protein selection as well as to better characterise the cellular and molecular context they are involved in to aid experimental validation; ii) focusing on characterising the complexity and connectivity of biological processes associated to apparently distinct pathophenotypes which likely share common mechanisms underpinning disease.

Indeed, it is hard to miss that FTLD bears numerous similarities to the other common neurodegenerative disorders, such as AD, PD and ALS. For instance, all four diseases commonly occur in later life, indicating a slow disease progression prior to onset of symptoms; all have strong hereditary components, at least in some forms of the disease²⁸³; all are characterised pathologically by aberrant intracellular protein inclusions²⁸⁴; many of them are found on similar or even the very same clinical and pathological continuum^{283,285}; and all have been associated with alterations in *waste disposal* and mitochondrial dysfunction²⁸⁶. Furthermore, the highly interconnected nature of the proteins and genes interactomes means that, at the molecular level, it is difficult to consider diseases as being independent of one another. Interestingly, the mapping of network-based dependencies between pathophenotypes has culminated in the concept of the *diseasome*, which represents disease networks where nodes represent diseases and edges represent various molecular relationships between the disease-associated cellular components, collectively encoding for the intrinsic features of diseases and cognate 'disease' genes^{4,287}. It is therefore evident that an integrated understanding of the interactions among the genome, the transcriptome, the proteome, the epigenome and the pathophenome, mediated by the underlying genetic and molecular network, offers a basis for future advancements in our understanding of the

structure and workings of global polygenic diseases and in elucidating the intrinsic complexity of genotype-phenotype associations related with human disease^{4,287}.

With regards to the experimental validation, further investigation of the mitophagy pathway in a disease-relevant setting is of paramount importance. The discovery that CDC37 participates in this pathway requires further study along the lines previously discussed in section 5.1.3 in order to fully elucidate its role along the different steps of the ALP and its relationship with mitophagy molecular components. Additionally, CDC37 implications with PGRN provide a potential insight into the disease mechanism where CDC37 is likely to be involved in and, as a result, it gives an indication of the genetic features that a suitable disease-relevant model should carry in this instance (i.e., GRN-linked FTLN model). However, genetic signatures are not the only features that need to be recapitulated in a relevant model: selective cell-type vulnerability in disease as well as age-associated markers of mature human neurons are only few of the features to be evaluated in a relevant NDs model.

In view of future studies carrying on investigating these findings, I propose using the CRISPR interference (CRISPRi) platform developed by Tian and co-authors²⁸⁸ to further elucidate CDC37 implications in mitophagy and *GRN*-linked FTLN. Briefly, a human iPSC line is stably transduced with a constitutively expressed dead Cas9 (dCas9) nuclease conjugated to the transcriptional repressor KRAB. Following introduction of a single guide RNA (sgRNA) targeting promoter regions of a gene of interest, expression of such gene is repressed by the dCas9 (KD), virtually allowing for effective and controlled KD of any gene. Additionally, transduction with a dox-inducible system forces the over-expression of the neuronal differentiation master regulator neurogenin 2 (NGN2), thus forcing the pluripotent cells to rewire their transcriptional program to that of a glutamatergic neuron following doxycycline treatment. This protocol allows for a neuronal model with KD of one or more genes of interest and presents with numerous advantages: i) provides rapid yield of large numbers of homogeneous neurons; ii) it does not cause p53-mediate toxicity and DNA damage; iii) it is inducible and reversible; iv) it has the potential of uncovering novel cell-type-specific gene functions; v) it has the potential of using additional combinations of transcription factors driving specific neuronal fates, allowing for dissection of neuronal subtype-specific gene function. By using these CRISPR-inducible neurons, it will be possible to assess PGRN expression and

mitophagy upregulation upon *CDC37* KD in a neuronal model. Validation of the findings obtained by J. Casey (*CDC37* expression reduction upon *GRN* R493X homozygous mutation) is also warranted, eventually followed by double KD of *GRN* and *CDC37* to assess whether *CDC37* suppression might contribute to rescue FTLN-linked *GRN* deficiency, making it a valuable therapeutic target for GRN-linked FTLN.

Finally, thanks to the continuous advances in the development of 3D cell culture models, the gap between pre-clinical animal and human studies is increasingly being bridged, providing a new paradigm for investigating the specific biology of complex traits and allowing for more accurate predictions of the effects of new therapies in models that replicate human pathophysiology as well as some aspects of the brain microenvironment, such as neurons and glia cross-talk and vascularisation of the tissue¹⁴². In the event that *CDC37* modulation yielded promising results in the CRISPRi model, organ-like cultures might be useful to target *CDC37* with an inhibitor and perform drug testing in a more complex high-throughput 3D system.

Notably, the present *in silico* approach had identified not only the autophagy/mitophagy pathway but also the UPS as candidate FTLN-risk pathway: further validation of the three candidate genes/proteins should be carried out in the context of the UPS to confirm their functional relevance. It goes without saying that all of the top biological processes and their top-ranked, associated proteins identified during the functional enrichment analysis are likely to hold variable degrees of functional relevance within FTLN and should be carried forward for validation. Future studies should aim at dissecting the functional implications of these pathways guided by the prioritisation pipeline.

Whilst *in silico* analyses can imply to the structural significance of the prioritised genes/proteins, genetic analysis might shed light on their possible involvement in the genetic risk (or components of the genetic heritability) of the FTLN trait. Future studies should use candidate genes/proteins to prioritise genetic markers in order to: i) gain genetic evidence on whether the identified genes/proteins are implicated in the FTLN trait, as they are closely linked to many seeds within relevant pathways; ii) provide cross-supporting data for *in silico* analyses and reinforce the validity of the present approach. Thanks to the IFGC initiative, an exceptionally large cohort is available to the public, with more than 6000 FTLN cases, for which genetic data is available for genetic analyses. Additionally, the three genes/proteins may potentially harbour genetic burden associated with an increased risk of developing FTLN; future studies should be

performing exploratory genetic analyses (i.e., rare variant discovery, burden analyses) to identify novel genetic risk factors which are contributing to disease development and progression. Identification of FTLD-specific and disease-associated rare variants might imply that the candidate genes/proteins tend to carry a burden of coding markers that could constitutively predispose patients to develop FTLD. All this taken together, would support the candidate genes/proteins importance in FTLD phenotype and would generate sufficient evidence to suggest that this approach is a valuable tool to guide lab-based hypothesis-driven experimental work, as well as genetic discovery analyses.

6.3. Current therapeutic approaches to treat FTLD and the *waste disposal* pathway as a therapeutic target

The clinical and molecular heterogeneity of FTLD has always represented a significant challenge for the development of effective treatments. To date, most treatments available for FTLD are purely symptomatic and have shown to have very little efficacy, sometimes even exacerbating behavioural symptoms as in the case of anticholinesterases^{289,290}. Similarly, selective serotonin reuptake inhibitors and antipsychotic therapies are generally considered helpful for mood and behaviour management in individual patients, despite lack of evidence²⁸⁹⁻²⁹².

It is worth mentioning that given their tightly related clinical and molecular phenotypes, often the therapeutic interventions of FTLD and ALS share common targets^{289,290}. In this scenario, therapies targeting the expression of heat shock proteins Hsp70 and Hsp90, which help newly synthesised proteins to properly fold, have long held potential for the treatment of neurodegenerative diseases, with the small molecule arimoclomol showing encouraging results at phase II clinical trials in reducing the levels of protein aggregates in motor nerves, a possible cause for ALS and FTLD-MND^{284,291}. In addition, arimoclomol was recently found to induce a HSF1-mediated reduction of the TDP-43 aggregates, a common pathological signature of both FTLD and ALS²⁹³. A phase III randomized, double blind, placebo-controlled trial is currently underway in order to further evaluate the efficacy and safety of arimoclomol as well as the therapeutic benefit of this molecule²⁹⁰.

In the last decade, tauopathy has also become a target for novel disease-modifying treatments for FTLD: the Methylene Blue drug is currently under evaluation in a phase III

clinical trial after it was shown to be effective in arresting age-related cognitive decline of tau-transgenic mice²⁹⁴. Of interest, a number of therapeutic approaches has emerged on the basis of other genetic factors contributing to FTLD pathology, where the identification of dominant mutations in *TARDBP* (TDP-43), *C9ORF72*, *FUS* and *GRN* represented the starting point to highlight alterations in biological processes underlying disease pathology.

A major bottleneck hampering the use of TDP-43 as a therapeutic target is that it is still not clear whether its pathology is driven by a mechanism of loss- or gain-of-function, and most importantly that TDP-43 is a ubiquitous protein that plays several roles within the cell^{295,296}. As a result, it may be difficult to target this protein in a generalised manner, although it might be possible to target the regions responsible for triggering the aggregation process, such as RNA Recognition Motif 1 (RRM-1)²⁹⁰. An alternative therapeutic option could consist of targeting eventual mutations in the protein sequence, which has been suggested to inhibit stress granule formation and thus suppress TDP-43-mediated toxicity.

By contrast, *C9ORF72* haploinsufficiency might be addressed by gene therapy approaches and particularly by using oligonucleotide-based antisense aimed at decreasing aberrant RNA expression, which to date has been shown to be one of the more successful approaches for the treatment of various NDS²⁸⁹.

Interestingly, autophagy enhancement using small molecules has been shown to successfully reduce cytoplasmic FUS inclusions, to restore mRBP homeostasis and rescue motor function *in vivo*²⁹⁰. Finally, gene therapy has been suggested as an obvious therapeutic target for *GRN* haploinsufficiency to restore proper *GRN* expression levels. Studies have shown promising results using adeno associated vectors (AAV) capable of expressing PGRN in the brain of *GRN*^{-/-} mice, thus improving lysosomal dysfunction and microglial pathology^{290,297,298}. Interestingly, Elia et al.⁸¹ found that a number of potential genetic modifiers of *GRN* seem to increase intracellular PGRN levels by either regulating PGRN expression at the transcriptional or at a post-translational level, suggesting that PGRN levels are regulated by multiple mechanisms. Elia and colleagues tested two small molecule drug inhibitors targeting *Foxo1* (PsammaphyseneA [PSA]²⁹⁹, AS1842856³⁰⁰) for efficacy in increasing PGRN levels in neurons, and found that not only they increased PGRN intracellular levels but also suppress lysosome enlargement-induced

neuropathology caused by PGRN deficiency, showing that their administration would be sufficient to reverse PGRN haploinsufficiency and thus that they could be of therapeutic utility. TRAP1, a member of the HSP90 family of molecular chaperones, was also shown to regulate proinflammatory cytokine TNF α -mediated neurodegeneration in *GRN*-linked FTL³⁰¹. As a result, two HSP90 inhibitors, currently used as anti-cancer agents in multiple phase I and II clinical trials (17AAG and AUY922^{32,302-304}), were used to inhibit TRAP1 and both showed dose-dependent increases in PGRN³², implying that they could eventually be transitioned for testing in humans to treat FTL.

Considering the findings of the present study, modulation of *CDC37* might also represent a potential therapeutic target able to reverse the loss of PGRN in disease. Inhibitors of *CDC37* might effectively upregulate *GRN* mRNA expression as well as *GRN* deficiency-mediated mitochondrial dysfunction, thereby partially restoring PGRN levels and boosting mitochondrial clearance, both of which are likely to reverse the loss of PGRN and selective functional deficits in the ALP in *GRN*-linked FTL.

While the identification of a role for autophagic and proteostasis dysregulation in the pathogenesis of neurodegenerative diseases is well established and has provided a rationale for interventions aiming to enhance autophagy, therapeutic strategies based on autophagy modulation have not been systematically explored yet for FTL. Importantly, the development of therapeutic strategies that directly target the ALP presents with a number of major challenges: i) establishing whether upregulation or repair of the autophagy pathway is needed to restore autophagy function and reinstate neuronal proteostasis; ii) examining potential detrimental effects resulting from exogenous autophagy induction; iii) the difficulty to dynamically evaluate autophagy *in vivo*³⁰⁵. Unfortunately, in many experimental models, the non-specific nature of the agents used to stimulate macroautophagy makes it difficult to directly link the toxic effect to pharmacological autophagy enhancement³⁰⁵. Determining at which point the autophagy pathway fails and focusing on pathways-specific activation of autophagy will be essential for therapy development for neurodegenerative disorders³⁰⁶.

6.4. Concluding remarks

Despite the established role of defective autophagy in the pathogenesis and progression of NDs and the promise of autophagy-based therapies, clinical translation of this wealth of knowledge is still at an early stage. Importantly, as mentioned in the previous section, several methodological limitations apply to such therapies, which become relevant both when diagnosing patients and monitoring efficiency of any autophagy-based intervention. Additionally, capturing and addressing early phases of pathophysiological molecular cascades is crucial for successful therapeutics: the fact that clinical presentation of disease is not evident until a large percentage of the frontal and temporal cortex is already lost represents a huge step-back in the effective implementation of such interventions, as they will not be able to bring those neurons back³⁰⁸. For this reason, it is essential to identify the molecular signatures of defective autophagy through fluid biomarkers (i.e., cerebrospinal fluid [CSF] or serum) to allow earlier diagnosis, constant monitoring and effective management of disease³⁰⁷. Developing effective autophagy biomarkers will be of paramount importance for early diagnosis, as autophagy failure might precede clinical presentation of disease. These might also include non-invasive, image-based methods for dynamic measurement of CNS autophagy (e.g., autophagic flux)³⁰⁵. To conclude, a fuller understanding and identification of the molecular factors and biomarkers underlying autophagy as well as FTLN pathophysiology at different stages are among the most urgent challenges that should be overcome in the next years, so that therapies based on autophagy modulation can be implemented to treat NDs and improve quality of life for a large number of people.

References

1. Munafò, M. R. & Davey Smith, G. Robust research needs many lines of evidence. *Nat. 2021 5537689* **553**, 399–401 (2018).
2. Baker, M. & Penny, D. Is there a reproducibility crisis? *Nature* **533**, 452–454 (2016).
3. Paci, P. *et al.* Gene co-expression in the interactome: moving from correlation toward causation via an integrated approach to disease module discovery. *npj Syst. Biol. Appl. 2021 71* **7**, 1–11 (2021).
4. Barabási, A. L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet. 2011 121* **12**, 56–68 (2010).
5. Warren, J. D., Rohrer, J. D. & Rossor, M. N. Frontotemporal dementia. *BMJ (Online)* vol. 347 (2013).
6. Clarimon, J. *et al.* Genetic architecture of neurodegenerative dementias. *Neuropharmacology* vol. 168 108014 (2020).
7. Spatt, J. Arnold Pick's concept of dementia. *Cortex* **39**, 525–531 (2003).
8. Rabinovici, G. D. & Miller, B. L. Frontotemporal lobar degeneration: epidemiology, pathophysiology, diagnosis and management. *CNS Drugs* **24**, 375–398 (2010).
9. Ferrari, R., Manzoni, C. & Hardy, J. Genetics and molecular mechanisms of frontotemporal lobar degeneration: an update and future avenues. *Neurobiol. Aging* **78**, 98–110 (2019).
10. Ferrari, R. *et al.* Frontotemporal dementia: insights into the biological underpinnings of disease through gene co-expression network analysis. *Mol Neurodegener* **11**, 21 (2016).
11. Young, J. J., Lavakumar, M., Tampi, D., Balachandran, S. & Tampi, R. R. Frontotemporal dementia: latest evidence and clinical implications. *Ther. Adv. Psychopharmacol.* **8**, 33–48 (2018).
12. Josephs, K. A. Frontotemporal dementia and related disorders: deciphering the enigma. *Ann Neurol* **64**, 4–14 (2008).
13. Boeve, B. F., Boxer, A. L., Kumfor, F., Pijnenburg, Y. & Rohrer, J. D. Advances and controversies in frontotemporal dementia: diagnosis, biomarkers, and therapeutic considerations. *Lancet Neurol.* **21**, 258–272 (2022).
14. Seelaar, H., Rohrer, J. D., Pijnenburg, Y. A., Fox, N. C. & van Swieten, J. C. Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. *J*

- Neurol Neurosurg Psychiatry* **82**, 476–486 (2011).
15. Neary, D. *et al.* Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* **51**, 1546–1554 (1998).
 16. O'Rourke, J. G. *et al.* C9orf72 is required for proper macrophage and microglial function in mice. *Science (80-.)*. **351**, 1324–1329 (2016).
 17. Lattante, S., Ciura, S., Rouleau, G. A. & Kabashi, E. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends in Genetics* vol. 31 263–273 (2015).
 18. Kurz, A. & Perneczky, R. Neurobiology of cognitive disorders. *Curr Opin Psychiatry* **22**, 546–551 (2009).
 19. Ferrari, R. *et al.* Screening for C9ORF72 repeat expansion in FTLD. *Neurobiol Aging* **33**, 1850 e1–11 (2012).
 20. Mackenzie, I. R. & Neumann, M. Molecular neuropathology of frontotemporal dementia: insights into disease mechanisms from postmortem studies. *J Neurochem* **138 Suppl**, 54–70 (2016).
 21. Ferrari R Momeni P, T. A. Molecular Genetics of Frontotemporal Dementia. *eLS* (2013) doi:10.1002/9780470015902.a0024457.
 22. Raz, L., Knoefel, J. & Bhaskar, K. The neuropathology and cerebrovascular mechanisms of dementia. *Journal of Cerebral Blood Flow and Metabolism* vol. 36 172–186 (2016).
 23. van Swieten, J. C. & Heutink, P. Mutations in progranulin (GRN) within the spectrum of clinical and pathological phenotypes of frontotemporal dementia. *Lancet Neurol* **7**, 965–974 (2008).
 24. Jiang, Y. X. *et al.* Amyloid fibrils in FTLD-TDP are composed of TMEM106B and not TDP-43. *Nat. 2022 6057909* **605**, 304–309 (2022).
 25. Hardy, J. & Singleton, A. Genomewide association studies and human disease. *N Engl J Med* **360**, 1759–1768 (2009).
 26. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
 27. Blauwendraat, C. *et al.* The wide genetic landscape of clinical frontotemporal dementia: Systematic combined sequencing of 121 consecutive subjects. *Genet. Med.* **20**, 240–249 (2018).
 28. Snowden, J. S., Neary, D. & Mann, D. M. Frontotemporal dementia. *Br J Psychiatry*

- 180**, 140–143 (2002).
29. Turner, M. R. *et al.* Genetic screening in sporadic ALS and FTD. *J Neurol Neurosurg Psychiatry* (2017) doi:10.1136/jnnp-2017-315995.
 30. Freibaum, B. D. & Taylor, J. P. The role of dipeptide repeats in C9ORF72-related ALS-FTD. *Frontiers in Molecular Neuroscience* vol. 10 (2017).
 31. Gijssels, I. *et al.* Progranulin locus deletion in frontotemporal dementia. *Hum Mutat* **29**, 53–58 (2008).
 32. Elia, L. P., Mason, A. R., Alijagic, A. & Finkbeiner, S. Genetic Regulation of Neuronal Progranulin Reveals a Critical Role for the Autophagy-Lysosome Pathway. **39**, 3332–3344 (2019).
 33. Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
 34. Majounie, E. *et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol* **11**, 323–330 (2012).
 35. van der Zee, J. *et al.* A pan-European study of the C9orf72 repeat associated with FTLT: geographic prevalence, genomic instability, and intermediate repeats. *Hum Mutat* **34**, 363–373 (2013).
 36. Al-Chalabi, A., van den Berg, L. H. & Veldink, J. Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nat Rev Neurol* **13**, 96–104 (2017).
 37. Cooper-Knock, J., Shaw, P. J. & Kirby, J. The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. *Acta Neuropathol* **127**, 333–345 (2014).
 38. Galimberti, D. *et al.* Incomplete penetrance of the C9ORF72 hexanucleotide repeat expansions: frequency in a cohort of geriatric non-demented subjects. *J Alzheimers Dis* **39**, 19–22 (2014).
 39. Hensman Moss, D. J. *et al.* C9orf72 expansions are the most common genetic cause of Huntington disease phenocopies. *Neurology* **82**, 292–299 (2014).
 40. Lindquist, S. G. *et al.* Corticobasal and ataxia syndromes widen the spectrum of C9ORF72 hexanucleotide expansion disease. *Clin Genet* **83**, 279–283 (2013).
 41. Majounie, E. *et al.* Large C9orf72 repeat expansions are not a common cause of Parkinson's disease. *Neurobiol Aging* **33**, 2527 e1–2 (2012).

42. Simon-Sanchez, J. *et al.* The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. *Brain* **135**, 723–735 (2012).
43. Smith, B. N. *et al.* The C9ORF72 expansion mutation is a common cause of ALS+/- FTD in Europe and has a single founder. *Eur J Hum Genet* **21**, 102–108 (2013).
44. van der Zee, J. *et al.* CHMP2B C-truncating mutations in frontotemporal lobar degeneration are associated with an aberrant endosomal phenotype in vitro. *Hum Mol Genet* **17**, 313–322 (2008).
45. Skibinski, G. *et al.* Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nat Genet* **37**, 806–808 (2005).
46. Le Ber, I. *et al.* SQSTM1 mutations in French patients with frontotemporal dementia or frontotemporal dementia with amyotrophic lateral sclerosis. *JAMA Neurol* **70**, 1403–1410 (2013).
47. Synofzik, M. *et al.* Screening in ALS and FTD patients reveals 3 novel UBQLN2 mutations outside the PXX domain and a pure FTD phenotype. *Neurobiol Aging* **33**, 2949 e13–7 (2012).
48. Pottier, C. *et al.* Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta Neuropathol* **130**, 77–92 (2015).
49. Bannwarth, S. *et al.* A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain* **137**, 2329–2345 (2014).
50. Gijssels, I. *et al.* Loss of TBK1 is a frequent cause of frontotemporal dementia in a Belgian cohort. *Neurology* **85**, 2116–2125 (2015).
51. Momeni, P. *et al.* Analysis of IFT74 as a candidate gene for chromosome 9p-linked ALS-FTD. *BMC Neurol* **6**, 44 (2006).
52. Münch, C. *et al.* Heterozygous R1101K mutation of the DCTN1 gene in a family with ALS and FTD. *Ann. Neurol.* **58**, 777–780 (2005).
53. Mackenzie, I. R. *et al.* TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron* **95**, 808–816 e9 (2017).
54. Hardy, J. & Rogaeva, E. Motor neuron disease and frontotemporal dementia: sometimes related, sometimes not. *Exp Neurol* **262 Pt B**, 75–83 (2014).
55. Van Mossevelde, S., Engelborghs, S., Van Der Zee, J. & Van Broeckhoven, C.

- Genotype-phenotype links in frontotemporal lobar degeneration. *Nature Reviews Neurology* vol. 14 363–378 (2018).
56. Watts, G. D. *et al.* Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet* **36**, 377–381 (2004).
 57. Koppers, M. *et al.* Screening for rare variants in the coding region of ALS-associated genes at 9p21.2 and 19p13.3. *Neurobiol Aging* **34**, 1518 e5–7 (2013).
 58. Benajiba, L. *et al.* TARDBP mutations in motoneuron disease with frontotemporal lobar degeneration. *Ann Neurol* **65**, 470–473 (2009).
 59. Borroni, B. *et al.* Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. *Hum Mutat* **30**, E974-83 (2009).
 60. Borroni, B. *et al.* TARDBP mutations in frontotemporal lobar degeneration: frequency, clinical features, and disease course. *Rejuvenation Res* **13**, 509–517 (2010).
 61. Sellami, L., Saracino, D. & Le Ber, I. Genetic forms of frontotemporal lobar degeneration: Current diagnostic approach and new directions in therapeutic strategies. *Rev. Neurol. (Paris)*. **176**, 571–581 (2020).
 62. Pottier, C., Ravenscroft, T. A., Sanchez-Contreras, M. & Rademakers, R. Genetics of FTLTD: overview and what else we can expect from genetic studies. *J Neurochem* **138 Suppl**, 32–53 (2016).
 63. LT, T. The Genetics of Monogenic Frontotemporal Dementia. *Dement Neuropsychol* 219–229 (2015).
 64. Ferrari, R. *et al.* Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet. Neurol.* **13**, 686–699 (2014).
 65. Ferrari, R. *et al.* A genome-wide screening and SNPs-to-genes approach to identify novel genetic risk factors associated with frontotemporal dementia. *Neurobiol Aging* **36**, 2904 e13–26 (2015).
 66. Reus, L. M. *et al.* Genome-wide association study of frontotemporal dementia identifies a C9ORF72 haplotype with a median of 12-G4C2 repeats that predisposes to pathological repeat expansions. *Transl. Psychiatry* 2021 **11**, 1–8 (2021).
 67. Van Deerlin, V. M. *et al.* Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* **42**, 234–

- 239 (2010).
68. Pottier, C. *et al.* Potential genetic modifiers of disease risk and age at onset in patients with frontotemporal lobar degeneration and GRN mutations: a genome-wide association study. *Lancet Neurol.* **17**, 548–558 (2018).
 69. Cruchaga, C. *et al.* Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. *Arch Neurol* **68**, 581–586 (2011).
 70. Finch, N. *et al.* TMEM106B regulates progranulin levels and the penetrance of FTLN in GRN mutation carriers. *Neurology* **76**, 467–474 (2011).
 71. Vass, R. *et al.* Risk genotypes at TMEM106B are associated with cognitive impairment in amyotrophic lateral sclerosis. *Acta Neuropathol* **121**, 373–380 (2011).
 72. Lattante, S. *et al.* Defining the association of TMEM106B variants among frontotemporal lobar degeneration patients with GRN mutations and C9orf72 repeat expansions. *Neurobiol Aging* **35**, 2658 e1-2658 e5 (2014).
 73. Takahashi, M. The GDNF/RET signaling pathway and human diseases. *Cytokine Growth Factor Rev.* **12**, 361–373 (2001).
 74. Costa, B. *et al.* C9orf72, age at onset, and ancestry help discriminate behavioral from language variants in FTLN cohorts. *Neurology* **95**, E3288–E3302 (2020).
 75. Mok, K. *et al.* Chromosome 9 ALS and FTD locus is probably derived from a single founder. *Neurobiol Aging* **33**, 209 e3–8 (2012).
 76. Zhang, M., Ferrari, R., Tartaglia, M. C., Keith, J. & *et al.* C6orf10/LOC101929163 locus is associated with age of onset in C9orf72 carriers | *Brain* | Oxford Academic. <https://academic.oup.com/brain/article/141/10/2895/5106718> (2018).
 77. Karczewski, K. J. & Snyder, M. P. Integrative omics for health and disease. *Nat. Rev. Genet.* **2018 195 19**, 299–310 (2018).
 78. Menzies, F. M. *et al.* Autophagy and Neurodegeneration: Pathogenic Mechanisms and Therapeutic Opportunities. *Neuron* vol. 93 1015–1034 (2017).
 79. S, F. The Autophagy Lysosomal Pathway and Neurodegeneration. *Cold Spring Harb. Perspect. Biol.* **12**, (2020).
 80. Sirkis, D. W., Geier, E. G., Bonham, L. W., Karch, C. M. & Yokoyama, J. S. Recent advances in the genetics of frontotemporal dementia. *Curr. Genet. Med. Rep.* **7**, 41

- (2019).
81. Elia, L. P., Reisine, T., Alijagic, A. & Finkbeiner, S. Approaches to develop therapeutics to treat Frontotemporal Dementia. *Neuropharmacology* **166**, 107948 (2020).
 82. Paushter, D. H., Du, H., Feng, T. & Hu, F. The lysosomal function of progranulin, a guardian against neurodegeneration. *Acta Neuropathol.* **136**, 1 (2018).
 83. Nguyen, A. D. *et al.* Murine knockin model for progranulin-deficient frontotemporal dementia with nonsensemediated mRNA decay. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E2849–E2858 (2018).
 84. Casey, J. M. *et al.* Haploinsufficiency of progranulin causes impairments in PINK/PARKIN mitophagy. *Alzheimer's Dement.* **16**, e042104 (2020).
 85. Klein, Z. A. *et al.* Loss of TMEM106B Ameliorates Lysosomal and Frontotemporal Dementia-Related Phenotypes in Progranulin-Deficient Mice. *Neuron* **95**, 281-296 e6 (2017).
 86. Mackenzie, I. R. & Rademakers, R. The molecular genetics and neuropathology of frontotemporal lobar degeneration: recent developments. *Neurogenetics* **8**, 237–248 (2007).
 87. Bader, V. & Winklhofer, K. F. PINK1 and Parkin: Team players in stress-induced mitophagy. *Biol. Chem.* **401**, 891–899 (2020).
 88. Harding, O. *et al.* ALS- And FTD-associated missense mutations in TBK1 differentially disrupt mitophagy. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
 89. Richter, B. *et al.* Phosphorylation of OPTN by TBK1 enhances its binding to Ub chains and promotes selective autophagy of damaged mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 4039–4044 (2016).
 90. Baek, M. *et al.* TDP-43 and PINK1 mediate CHCHD10 S59L mutation-induced defects in Drosophila and in vitro. *Nat. Commun.* **12**, 1–20 (2021).
 91. Root, J., Merino, P., Nuckols, A., Johnson, M. & Kukar, T. Lysosome dysfunction as a cause of neurodegenerative diseases: Lessons from frontotemporal dementia and amyotrophic lateral sclerosis. *Neurobiol. Dis.* **154**, 105360 (2021).
 92. Tan, C. C. *et al.* Association of Frontotemporal Dementia GWAS Loci with Late-Onset Alzheimer's Disease in a Northern Han Chinese Population. *J Alzheimers Dis* **52**, 43–50 (2016).
 93. Ferrari, R. *et al.* Frontotemporal dementia: Insights into the biological

- underpinnings of disease through gene co-expression network analysis. *Mol. Neurodegener.* **11**, 21 (2016).
94. Ratti, A. & Buratti, E. Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *J. Neurochem.* **138 Suppl**, 95–111 (2016).
 95. Corrado, L. *et al.* High frequency of TARDBP gene mutations in Italian patients with amyotrophic lateral sclerosis. *Hum Mutat* **30**, 688–694 (2009).
 96. Ferrari, R., Hardy, J. & Momeni, P. Frontotemporal dementia: from Mendelian genetics towards genome wide association studies. *J Mol Neurosci* **45**, 500–515 (2011).
 97. Ferrari, R., Lovering, R. C., Hardy, J., Lewis, P. A. & Manzoni, C. Weighted Protein Interaction Network Analysis of Frontotemporal Dementia. *J Proteome Res* **16**, 999–1013 (2017).
 98. Giglia-Mari, G., Zotter, A. & Vermeulen, W. DNA Damage Response. *Cold Spring Harb. Perspect. Biol.* **3**, a000745 (2011).
 99. Klaips, C. L., Jayaraj, G. G. & Hartl, F. U. Pathways of cellular proteostasis in aging and disease. *J. Cell Biol.* **217**, 51–63 (2018).
 100. Shahheydari, H. *et al.* Protein Quality Control and the Amyotrophic Lateral Sclerosis/Frontotemporal Dementia Continuum. *Front. Mol. Neurosci.* **10**, (2017).
 101. Pickles, S., Vigié, P. & Youle, R. J. Mitophagy and Quality Control Mechanisms in Mitochondrial Maintenance. *Curr. Biol.* **28**, R170–R185 (2018).
 102. Narayan, M. & Jinwal, U. K. Cdc37: Implications in Regulation of Kinases and Proteins Linked to Neurodegenerative and Other Diseases. *Autophagy Cancer, Other Pathol. Inflammation, Immunity, Infect. Aging* 187–196 (2016)
doi:10.1016/B978-0-12-805421-5.00009-4.
 103. Watanabe, Y., Taguchi, K. & Tanaka, M. Ubiquitin, Autophagy and Neurodegenerative Diseases. *Cells* **9**, (2020).
 104. Klionsky, D. J. *et al.* Guidelines for the use and interpretation of assays for monitoring autophagy. *Autophagy* vol. 8 445–544 (2012).
 105. Tanaka, K. The proteasome: Overview of structure and functions. *Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci.* **85**, 12 (2009).
 106. Lizama, B. N. *et al.* Neuronal preconditioning requires the mitophagic activity of C-terminus of HSC70-interacting protein. *J. Neurosci.* **38**, 6825–6840 (2018).
 107. Kumar, P., Jha, N. K., Jha, S. K., Ramani, K. & Ambasta, R. K. Tau phosphorylation,

- molecular chaperones, and ubiquitin E3 Ligase: Clinical relevance in Alzheimer's disease. *Journal of Alzheimer's Disease* vol. 43 341–361 (2015).
108. Klionsky, D. J. Autophagy: from phenomenology to molecular understanding in less than a decade. *Nat. Rev. Mol. Cell Biol.* 2007 811 **8**, 931–937 (2007).
 109. Rubinsztein, D. *et al.* Mendelian neurodegenerative disease genes involved in autophagy. *Cell Discov.* 2020 61 **6**, 1–13 (2020).
 110. Li, R. *et al.* Serine/threonine Kinase Unc-51-like Kinase-1 (Ulk1) phosphorylates the co-chaperone cell division cycle protein 37 (Cdc37) and thereby disrupts the stability of Cdc37 client proteins. *J. Biol. Chem.* **292**, 2830–2841 (2017).
 111. Martinez-Vicente, M. *et al.* Dopamine-modified alpha-synuclein blocks chaperone-mediated autophagy. *J. Clin. Invest.* **118**, 777–778 (2008).
 112. Shpilka, T. & Elazar, Z. Shedding light on mammalian microautophagy. *Dev. Cell* **20**, 1–2 (2011).
 113. Wong, Y. C. & Holzbaur, E. L. F. The regulation of autophagosome dynamics by huntingtin and HAP1 is disrupted by expression of mutant huntingtin, leading to defective cargo degradation. *J. Neurosci.* **34**, 1293–1305 (2014).
 114. Manzoni, C. The LRRK2-macroautophagy axis and its relevance to Parkinson's disease. *Biochem. Soc. Trans.* **45**, 155–162 (2017).
 115. Ramirez, A. *et al.* Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. *Nat. Genet.* **38**, 1184–1191 (2006).
 116. Youle, R. J. & Narendra, D. P. Mechanisms of mitophagy. *Nat. Rev. Mol. Cell Biol.* **12**, 9 (2011).
 117. Malpartida, A. B., Williamson, M., Narendra, D. P., Wade-Martins, R. & Ryan, B. J. Mitochondrial Dysfunction and Mitophagy in Parkinson's Disease: From Mechanism to Therapy. *Trends in Biochemical Sciences* (2020) doi:10.1016/j.tibs.2020.11.007.
 118. Kim, J. & Klionsky, D. J. Autophagy, Cytoplasm-to-Vacuole Targeting Pathway, and Pexophagy in Yeast and Mammalian Cells. <http://dx.doi.org/10.1146/annurev.biochem.69.1.303> **69**, 303–342 (2003).
 119. Maday, S. & Holzbaur, E. L. F. Autophagosome biogenesis in primary neurons follows an ordered and spatially regulated pathway. *Dev. Cell* **30**, 71–85 (2014).
 120. Eden, E. *et al.* Proteome half-life dynamics in living human cells. *Science* (80-.).

- 331**, 764–768 (2011).
121. Huang, M. *et al.* Network analysis of the progranulin-deficient mouse brain proteome reveals pathogenic mechanisms shared in human frontotemporal dementia caused by GRN mutations. *Acta Neuropathol. Commun.* **8**, 163 (2020).
 122. Ondaro, J. *et al.* Defects of Nutrient Signaling and Autophagy in Neurodegeneration. *Front. Cell Dev. Biol.* **10**, 836196 (2022).
 123. Dröge, W. Autophagy and aging--importance of amino acid levels. *Mech. Ageing Dev.* **125**, 161–168 (2004).
 124. Meléndez, A. *et al.* Autophagy genes are essential for dauer development and life-span extension in *C. elegans*. *Science* **301**, 1387–1391 (2003).
 125. Carnio, S. *et al.* Autophagy impairment in muscle induces neuromuscular junction degeneration and precocious aging. *Cell Rep.* **8**, 1509–1521 (2014).
 126. Urwin, H., Ghazi-Noori, S., Collinge, J. & Isaacs, A. The role of CHMP2B in frontotemporal dementia. *Biochem Soc Trans* **37**, 208–212 (2009).
 127. Ge, P., Dawson, V. L. & Dawson, T. M. PINK1 and Parkin mitochondrial quality control: a source of regional vulnerability in Parkinson's disease. *Mol. Neurodegener.* 2020 151 **15**, 1–18 (2020).
 128. Grenier, K., McLelland, G. L. & Fon, E. A. Parkin- and PINK1-Dependent Mitophagy in Neurons: Will the Real Pathway Please Stand Up? *Front. Neurol.* **4**, (2013).
 129. Aklima, J. *et al.* Effects of Matrix pH on Spontaneous Transient Depolarization and Reactive Oxygen Species Production in Mitochondria. *Front. cell Dev. Biol.* **9**, (2021).
 130. Rossi, A., Pizzo, P. & Filadi, R. Calcium, mitochondria and cell metabolism: A functional triangle in bioenergetics. *Biochimica et Biophysica Acta - Molecular Cell Research* vol. 1866 1068–1078 (2019).
 131. Tatsuta, T. & Langer, T. Quality control of mitochondria: protection against neurodegeneration and ageing. *EMBO J.* **27**, 306 (2008).
 132. Karbowski, M. & Neutzner, A. Neurodegeneration as a consequence of failed mitochondrial maintenance. *Acta Neuropathol.* **123**, 157–171 (2012).
 133. Ng, M. Y. W., Wai, T. & Simonsen, A. Quality control of the mitochondrion. *Dev. Cell* **56**, 881–905 (2021).
 134. Deas, E., Wood, N. W. & Plun-Favreau, H. Mitophagy and Parkinson's disease: The PINK1–parkin link. *Biochim. Biophys. Acta* **1813**, 623 (2011).

135. Lemasters, J. J. Selective Mitochondrial Autophagy, or Mitophagy, as a Targeted Defense Against Oxidative Stress, Mitochondrial Dysfunction, and Aging. <https://home.liebertpub.com/rej> **8**, 3–5 (2005).
136. Kim, I., Rodriguez-Enriquez, S. & Lemasters, J. J. Selective degradation of mitochondria by mitophagy. *Arch. Biochem. Biophys.* **462**, 245–253 (2007).
137. Villa, E. *et al.* Parkin-Independent Mitophagy Controls Chemotherapeutic Response in Cancer Cells. *Cell Rep.* **20**, 2846–2859 (2017).
138. Deas, E., Plun-Favreau, H. & Wood, N. W. PINK1 function in health and disease. *EMBO Mol. Med.* **1**, 152–165 (2009).
139. Fiesel, F. C. *et al.* (Patho-)physiological relevance of PINK 1-dependent ubiquitin phosphorylation. *EMBO Rep.* **16**, 1114–1130 (2015).
140. Hou, X. *et al.* Age- and disease-dependent increase of the mitophagy marker phospho-ubiquitin in normal aging and Lewy body disease. *Autophagy* **14**, 1404–1418 (2018).
141. Lines, G., Casey, J. M., Preza, E. & Wray, S. Modelling frontotemporal dementia using patient-derived induced pluripotent stem cells. *Mol. Cell. Neurosci.* **109**, 103553 (2020).
142. Slanzi, A., Iannoto, G., Rossi, B., Zenaro, E. & Constantin, G. In vitro Models of Neurodegenerative Diseases. *Front. Cell Dev. Biol.* **8**, 328 (2020).
143. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).
144. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
145. Tan, R. H., Ke, Y. D., Ittner, L. M. & Halliday, G. M. ALS/FTLD: experimental models and reality. *Acta Neuropathol.* 2017 1332 **133**, 177–196 (2017).
146. Kühn, R., Mahajan, A., Canoll, P. & Hargus, G. Human Induced Pluripotent Stem Cell Models of Frontotemporal Dementia With Tau Pathology. *Front. Cell Dev. Biol.* **9**, 2927 (2021).
147. Leskelä, S. *et al.* FTLD Patient-Derived Fibroblasts Show Defective Mitochondrial Function and Accumulation of p62. *Mol. Neurobiol.* **58**, 5438–5458 (2021).
148. Guttikonda, S. R. *et al.* Fully defined human pluripotent stem cell-derived microglia and tri-culture system model C3 production in Alzheimer’s disease. *Nat.*

- Neurosci.* **24**, 343–354 (2021).
149. Jawaid, A., Khan, R., Polymenidou, M. & Schulz, P. E. Disease-modifying effects of metabolic perturbations in ALS/FTLD. *Mol. Neurodegener.* **13**, 1–16 (2018).
 150. Manzoni, C. *et al.* Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences. *Br. Bioinform* (2016) doi:10.1093/bib/bbw114.
 151. Ferrari, R. *et al.* Stratification of candidate genes for Parkinson’s disease using weighted protein-protein interaction network analysis. *BMC Genomics* **19**, 452 (2018).
 152. Wong, L. *Using Biological Networks in Protein Function Prediction and Gene Expression Analysis. Internet Mathematics* vol. 7.
 153. Manzoni, C., Lewis, P. A. & Ferrari, R. Network Analysis for Complex Neurodegenerative Diseases. *Curr. Genet. Med. Rep.* **8**, 17–25 (2020).
 154. Bray, D. Molecular networks: The top-down view. *Science* vol. 301 1864–1865 (2003).
 155. Parikshak, N. N., Gandal, M. J. & Geschwind, D. H. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. *Nat. Rev. Genet.* **16**, 441 (2015).
 156. Barabási, A. L. & Oltvai, Z. N. Network biology: Understanding the cell’s functional organization. *Nature Reviews Genetics* vol. 5 101–113 (2004).
 157. Snider, J. *et al.* Fundamentals of protein interaction network mapping. *Mol. Syst. Biol.* **11**, 848 (2015).
 158. Stoney, R., Robertson, D. L., Nenadic, G. & Schwartz, J. M. Mapping biological process relationships and disease perturbations within a pathway network. *npj Syst. Biol. Appl.* **4**, (2018).
 159. Dorogovtsev, S. N. & Mendes, J. F. F. *Evolution of Networks: From Biological Nets to the Internet and WWW.* (Oxford University Press, 2003). doi:10.1093/acprof:oso/9780198515906.001.0001.
 160. Taylor, I. W. & Wrana, J. L. Protein interaction networks in medicine and disease. *Proteomics* **12**, 1706–1716 (2012).
 161. Carter, H., Hofree, M. & Ideker, T. Genotype to phenotype via network analysis. *Curr. Opin. Genet. Dev.* **23**, 611–621 (2013).

162. Califano, A., Butte, A. J., Friend, S., Ideker, T. & Schadt, E. Leveraging models of cell regulation and GWAS data in integrative network-based association studies. *Nat. Genet.* 2012 448 **44**, 841–847 (2012).
163. Furlong, L. I. Human diseases through the lens of network biology. *Trends Genet* **29**, 150–159 (2013).
164. Ideker, T. & Sharan, R. Protein networks in disease. *Genome Res.* **18**, 644 (2008).
165. Serin, E. A. R., Nijveen, H., Hilhorst, H. W. M. & Ligterink, W. Learning from co-expression networks: Possibilities and challenges. *Front. Plant Sci.* **7**, 444 (2016).
166. Ma, X. *et al.* Co-expression Gene Network Analysis and Functional Module Identification in Bamboo Growth and Development. *Front. Genet.* **9**, 574 (2018).
167. van Dam, S., Vösa, U., van der Graaf, A., Franke, L. & de Magalhães, J. P. Gene co-expression analysis for functional classification and gene-disease predictions. *Br. Bioinform* **19**, 575–592 (2017).
168. Tomkins, J. E. *et al.* Comparative Protein Interaction Network Analysis Identifies Shared and Distinct Functions for the Human ROCO Proteins. *Proteomics* **18**, (2018).
169. Fionda, V. Networks in Biology. *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.* **1–3**, 915–921 (2019).
170. Silverbush, D. & Sharan, R. A systematic approach to orient the human protein–protein interaction network. *Nat. Commun.* 2019 101 **10**, 1–9 (2019).
171. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* 2016 483 **48**, 245–252 (2016).
172. Ahmadi, A., Gispert, J. D., Navarro, A., Vilor-Tejedor, N. & Sadeghi, I. Single-cell Transcriptional Changes in Neurodegenerative Diseases. *Neuroscience* **479**, 192–205 (2021).
173. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
174. Gorno-Tempini, M. L. *et al.* Classification of primary progressive aphasia and its variants. *Neurology* **76**, 1006–1014 (2011).
175. Rascovsky, K. *et al.* Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* **134**, 2456–2477 (2011).
176. Blauwendraat, C. *et al.* NeuroChip, an updated version of the NeuroX genotyping platform to rapidly screen for variants associated with neurological diseases.

- Neurobiol Aging* **57**, 247 e9-247 e13 (2017).
177. Van Mossevelde, S., van der Zee, J., Cruts, M. & Van Broeckhoven, C. Relationship between C9orf72 repeat size and clinical phenotype. *Current Opinion in Genetics and Development* vol. 44 117–124 (2017).
 178. García-Ruiz, S. *et al.* CoExp: A Web Tool for the Exploitation of Co-expression Networks. *Front. Genet.* **0**, 218 (2021).
 179. Tomkins, J. E. *et al.* PINOT: An intuitive resource for integrating protein-protein interactions. *Cell Commun. Signal.* **18**, 1–11 (2020).
 180. Bonham, L. W. *et al.* Protein network analysis reveals selectively vulnerable regions and biological processes in FTD. *Neurol. Genet.* **4**, e266 (2018).
 181. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **3**, 1101–1108 (2008).
 182. Shi, Y., Kirwan, P. & Livesey, F. J. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat. Protoc.* **2012** 710 7, 1836–1846 (2012).
 183. Harms, M. B. *et al.* Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. *Neurobiol. Aging* **34**, 2234.e13-2234.e19 (2013).
 184. Babic Leko, M. *et al.* Molecular Mechanisms of Neurodegeneration Related to C9orf72 Hexanucleotide Repeat Expansion. *Behav Neurol* **2019**, 2909168 (2019).
 185. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
 186. Beck, J. *et al.* Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *Am. J. Hum. Genet.* **92**, 345–353 (2013).
 187. Fournier, C. *et al.* Relations between C9orf72 expansion size in blood, age at onset, age at collection and transmission across generations in patients and presymptomatic carriers. *Neurobiol. Aging* **74**, 234.e1-234.e8 (2019).
 188. Gijssels, I. *et al.* The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Mol. Psychiatry* **21**, 1112–1124 (2016).
 189. van Blitterswijk, M. *et al.* Association between repeat sizes and clinical and

- pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): A cross-sectional cohort study. *Lancet Neurol.* **12**, 978–988 (2013).
190. Van Mossevelde, S. *et al.* Clinical evidence of disease anticipation in families segregating a C9orf72 repeat expansion. *JAMA Neurol.* **74**, 445–452 (2017).
 191. Benussi, L. *et al.* C9ORF72 hexanucleotide repeat number in frontotemporal lobar degeneration: A genotype-phenotype correlation study. *J. Alzheimer's Dis.* **38**, 799–808 (2014).
 192. Devenney, E. *et al.* Progression in behavioral variant frontotemporal dementia: A longitudinal study. *JAMA Neurol.* **72**, 1501–1509 (2015).
 193. Galimberti, D. *et al.* Autosomal dominant frontotemporal lobar degeneration due to the C9ORF72 hexanucleotide repeat expansion: Late-onset psychotic clinical presentation. *Biol. Psychiatry* **74**, 384–391 (2013).
 194. Ramos, E. M. *et al.* Genetic screen in a large series of patients with primary progressive aphasia. *Alzheimer's Dement.* **15**, 553–560 (2019).
 195. Ralph, P. & Coop, G. The Geography of Recent Genetic Ancestry across Europe. *PLoS Biol.* **11**, (2013).
 196. Warby, S. C. *et al.* CAG Expansion in the Huntington Disease Gene Is Associated with a Specific and Targetable Predisposing Haplogroup. *Am. J. Hum. Genet.* **84**, 351–366 (2009).
 197. Huang, T., Shu, Y. & Cai, Y. D. Genetic differences among ethnic groups. *BMC Genomics* **16**, (2015).
 198. Manzoni, Claudia; Kia, Demis A.; Vandrovcova, Jana; Hardy, John; Wood, Nicholas W.; Lewis, Patrick A.; Ferrari, R. Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences | Briefings in Bioinformatics | Oxford Academic. *Briefings in Bioinformatics* <https://academic.oup.com/bib/article/19/2/286/2562648> (2018).
 199. Ferrari, R. *et al.* Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurol* **13**, 686–699 (2014).
 200. Liu, D., Ke, Z. & Luo, J. Thiamine Deficiency and Neurodegeneration: the Interplay Among Oxidative Stress, Endoplasmic Reticulum Stress, and Autophagy. *Molecular Neurobiology* vol. 54 5440–5448 (2017).
 201. Miki, Y. *et al.* Alteration of mitochondrial protein PDHA1 in Lewy body disease and PARK14. *Biochem. Biophys. Res. Commun.* **489**, 439–444 (2017).

202. Yang, F. *et al.* Identification of Key Regulatory Genes and Pathways in Prefrontal Cortex of Alzheimer's Disease. *Interdiscip. Sci. Comput. Life Sci.* **12**, 90–98 (2020).
203. Thathiah, A. *et al.* β -Arrestin 2 regulates A β generation and γ -secretase activity in Alzheimer's disease. *Nat. Med.* **19**, 44–49 (2013).
204. Smith, A., Bourdeau, I., Wang, J. & Bondy, C. A. Expression of Catenin family members CTNNA1, CTNNA2, CTNNB1 and JUP in the primate prefrontal cortex and hippocampus. *Mol. Brain Res.* **135**, 225–231 (2005).
205. Burns, L. H. & Wang, H.-Y. Altered filamin A enables amyloid beta-induced tau hyperphosphorylation and neuroinflammation in Alzheimer's disease. *Neuroimmunol. Neuroinflammation* **4**, 263 (2017).
206. Haase, M. & Fitze, G. HSP90AB1: Helping the good and the bad. *Gene* vol. 575 171–186 (2016).
207. Mao, Y., Fisher, D. W., Yang, S., Keszycki, R. M. & Dong, H. Protein-protein interactions underlying the behavioral and psychological symptoms of dementia (BPSD) and Alzheimer's disease. *PLoS One* **15**, (2020).
208. Silva, P. N. *et al.* Analysis of HSPA8 and HSPA9 mRNA expression and promoter methylation in the brain and blood of Alzheimer's disease patients. *J. Alzheimer's Dis.* **38**, 165–170 (2014).
209. Conway, M., Nafar, F., Straka, T. & Mearow, K. Modulation of amyloid- β protein precursor expression by HspB1. *J. Alzheimer's Dis.* **42**, 435–450 (2014).
210. Yip, A. M. & Horvath, S. Gene network interconnectedness and the generalized topological overlap measure. *BMC Bioinformatics* **8**, 1–14 (2007).
211. Palubinsky, A. M. *et al.* CHIP Is an Essential Determinant of Neuronal Mitochondrial Stress Signaling. *Antioxid. Redox Signal.* **23**, 535–549 (2015).
212. Ullah, K. *et al.* The E3 ubiquitin ligase STUB1 attenuates cell senescence by promoting the ubiquitination and degradation of the core circadian regulator BMAL1. *J. Biol. Chem.* **295**, jbc.RA119.011280 (2020).
213. Rao, L., Sha, Y. & Eissa, N. T. The E3 ubiquitin ligase STUB1 regulates autophagy and mitochondrial biogenesis by modulating TFEB activity. *Mol. Cell. Oncol.* **4**, (2017).
214. Sha, Y., Rao, L., Settembre, C., Ballabio, A. & Eissa, N. T. STUB1 regulates TFEB-induced autophagy-lysosome pathway. *EMBO J.* **36**, 2544–2552 (2017).
215. Moriwaki, Y. *et al.* L347P PINK1 mutant that fails to bind to Hsp90/Cdc37

- chaperones is rapidly degraded in a proteasome-dependent manner. *Neurosci. Res.* **61**, 43–48 (2008).
216. A, W. *et al.* Pink1 Parkinson mutations, the Cdc37/Hsp90 chaperones and Parkin all influence the maturation or subcellular distribution of Pink1. **17**, 602–616 (2008).
217. Joo, J. H. & Kundu, M. ULK1-Hsp90-Cdc37 and AMPK: Novel Insight Into the Regulation of Mitophagy. *Blood* **120**, 987–987 (2012).
218. Jinwal, U. K. *et al.* Cdc37/Hsp90 protein complex disruption triggers an autophagic clearance cascade for TDP-43 protein. *J. Biol. Chem.* **287**, 24814–24820 (2012).
219. Heir, P. *et al.* DCNL1 Functions as a Substrate Sensor and Activator of Cullin 2-RING Ligase. *Mol. Cell. Biol.* **33**, 1621–1631 (2013).
220. Cai, W. & Yang, H. The structure and regulation of Cullin 2 based E3 ubiquitin ligases and their biological functions. *Cell Division* vol. 11 (2016).
221. Uchida, T. *et al.* CUL2-mediated clearance of misfolded TDP-43 is paradoxically affected by VHL in oligodendrocytes in ALS. *Sci. Rep.* **6**, (2016).
222. Cook, C. N. *et al.* C9orf72 poly(GR) aggregation induces TDP-43 proteinopathy. *Sci. Transl. Med.* **12**, (2020).
223. De Boer, E. M. J. *et al.* TDP-43 proteinopathies: a new wave of neurodegenerative diseases. *J. Neurol. Neurosurg. Psychiatry* **92**, 86–95 (2020).
224. Gracia, L., Lora, G., Blair, L. J. & Jinwal, U. K. *Therapeutic Potential of the Hsp90/Cdc37 Interaction in Neurodegenerative Diseases. Frontiers in Neuroscience* vol. 13 1263 (Frontiers Media S.A., 2019).
225. Wagner, M. *et al.* Clinico-genetic findings in 509 frontotemporal dementia patients. *Mol. Psychiatry* 2021 2610 **26**, 5824–5832 (2021).
226. Guennec, K. Le *et al.* ABCA7 rare variants and Alzheimer disease risk. *Neurology* **86**, 2134–2137 (2016).
227. Zhao, Q. F. *et al.* ABCA7 Genotypes Confer Alzheimer’s Disease Risk by Modulating Amyloid- β Pathology. *J. Alzheimer’s Dis.* **52**, 693–703 (2016).
228. Cuyvers, E. *et al.* Mutations in ABCA7 in a Belgian cohort of Alzheimer’s disease patients: A targeted resequencing study. *Lancet Neurol.* **14**, 814–822 (2015).
229. Steinberg, S. *et al.* Loss-of-function variants in ABCA7 confer risk of Alzheimer’s disease. *Nat. Genet.* **47**, 445–447 (2015).

230. Hollingworth, P. *et al.* Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat. Genet.* **43**, 429–436 (2011).
231. De Michele, G. *et al.* Spinocerebellar ataxia type 48: last but not least. *Neurol. Sci.* **41**, 2423–2432 (2020).
232. Zhang, S., Hu, Z. wei, Mao, C. yuan, Shi, C. he & Xu, Y. ming. CHIP as a therapeutic target for neurological diseases. *Cell Death and Disease* vol. 11 1–12 (2020).
233. Kanack, A. J., Newsom, O. J. & Scaglione, K. M. Most mutations that cause spinocerebellar ataxia autosomal recessive type 16 (SCAR16) destabilize the protein quality-control E3 ligase CHIP. *J. Biol. Chem.* **293**, 2735–2743 (2018).
234. Oddo, S. *et al.* Blocking A β 42 Accumulation Delays the Onset and Progression of Tau Pathology via the C Terminus of Heat Shock Protein70-Interacting Protein: A Mechanistic Link between A β and Tau Pathology. *J. Neurosci.* **28**, 12163 (2008).
235. Imai, Y. *et al.* CHIP is associated with Parkin, a gene responsible for familial Parkinson's disease, and enhances its ubiquitin ligase activity. *Mol. Cell* **10**, 55–67 (2002).
236. Mahrouf, N. *et al.* Characterization of Cullin-box Sequences That Direct Recruitment of Cul2-Rbx1 and Cul5-Rbx2 Modules to Elongin BC-based Ubiquitin Ligases. *J. Biol. Chem.* **283**, 8005–8013 (2008).
237. Yasukawa, T. *et al.* NRBP1-Containing CRL2/CRL4A Regulates Amyloid β Production by Targeting BRI2 and BRI3 for Degradation. *Cell Rep.* **30**, 3478–3491.e6 (2020).
238. Jinwal, U. K. *et al.* The Hsp90 kinase co-chaperone Cdc37 regulates tau stability and phosphorylation dynamics. *J. Biol. Chem.* **286**, 16976–16983 (2011).
239. Joo, J. H. *et al.* Hsp90-Cdc37 Chaperone Complex Regulates Ulk1- and Atg13-Mediated Mitophagy. *Mol. Cell* **43**, 572–585 (2011).
240. Kimura, Y. *et al.* Cdc37 is a molecular chaperone with specific functions in signal transduction. *Genes Dev.* **11**, 1775–1785 (1997).
241. Ando, M. *et al.* The PINK1 p.I368N mutation affects protein stability and ubiquitin kinase activity. **12**, (2017).
242. Seibler, P. *et al.* Mitochondrial Parkin Recruitment Is Impaired in Neurons Derived from Mutant PINK1 Induced Pluripotent Stem Cells. *J. Neurosci.* **31**, 5970–5976 (2011).

243. Cai, Q., Zakaria, H. M., Simone, A. & Sheng, Z. H. Spatial parkin translocation and degradation of damaged mitochondria via mitophagy in live cortical neurons. *Curr. Biol.* **22**, 545–552 (2012).
244. Rakovic, A. *et al.* Phosphatase and Tensin Homolog (PTEN)-induced Putative Kinase 1 (PINK1)-dependent Ubiquitination of Endogenous Parkin Attenuates Mitophagy: STUDY IN HUMAN PRIMARY FIBROBLASTS AND INDUCED PLURIPOTENT STEM CELL-DERIVED NEURONS. *J. Biol. Chem.* **288**, 2223–2237 (2013).
245. van Laar, V. S. *et al.* Bioenergetics of neurons inhibit the translocation response of Parkin following rapid mitochondrial depolarization. *Hum. Mol. Genet.* **20**, 927–940 (2011).
246. Joselin, A. P. *et al.* ROS-dependent regulation of Parkin and DJ-1 localization during oxidative stress in neurons. *Hum. Mol. Genet.* **21**, 4888–4903 (2012).
247. Mokranjac, D. & Neupert, W. The many faces of the mitochondrial TIM23 complex. *Biochim. Biophys. Acta - Bioenerg.* **1797**, 1045–1054 (2010).
248. McLelland, G. L. *et al.* Mfn2 ubiquitination by PINK1/parkin gates the p97-dependent release of ER from mitochondria to drive mitophagy. *Elife* **7**, (2018).
249. Sun, N. *et al.* Measuring in vivo mitophagy. *Mol. Cell* **60**, 685 (2015).
250. Winsor, N. J., Killackey, S. A., Philpott, D. J. & Girardin, S. E. An optimized procedure for quantitative analysis of mitophagy with the mtKeima system using flow cytometry. *Biotechniques* **69**, 249–256 (2020).
251. Frew, J. & Nygaard, H. B. Neuropathological and behavioral characterization of aged Grn R493X progranulin-deficient frontotemporal dementia knockin mice. *Acta Neuropathol. Commun.* **9**, 1–16 (2021).
252. Gerst, J. L. *et al.* Role of oxidative stress in frontotemporal dementia. *Dement. Geriatr. Cogn. Disord.* **10 Suppl 1**, 85–87 (1999).
253. Jenner, P. *et al.* Oxidative stress in Parkinson's disease. *Ann. Neurol.* **53**, S26–S38 (2003).
254. Huang, W. J., Zhang, X. & Chen, W. W. Role of oxidative stress in Alzheimer's disease. *Biomed. Reports* **4**, 519 (2016).
255. Gómez-Sánchez, R. *et al.* PINK1 deficiency enhances autophagy and mitophagy induction. *Mol. Cell. Oncol.* **3**, (2016).
256. Klein, C. *et al.* PINK1-interacting proteins: Proteomic analysis of overexpressed

- PINK1. *Parkinsons. Dis.* (2011) doi:10.4061/2011/153979.
257. Watzlawik, J. O. *et al.* Sensitive ELISA-based detection method for the mitophagy marker p-S65-Ub in human cells, autopsy brain, and blood samples. *Autophagy* **17**, 2613–2628 (2021).
258. Evans, C. S. & Holzbaur, E. L. F. Degradation of engulfed mitochondria is rate-limiting in optineurin-mediated mitophagy in neurons. *Elife* **9**, (2020).
259. Mendsaikhan, A., Tooyama, I. & Walker, D. G. Microglial Progranulin: Involvement in Alzheimer's Disease and Neurodegenerative Diseases. *Cells* **8**, 230 (2019).
260. Evers, B. M. *et al.* Lipidomic and Transcriptomic Basis of Lysosomal Dysfunction in Progranulin Deficiency. *Cell Rep.* **20**, 2565–2574 (2017).
261. Sun, A. Lysosomal storage disease overview. *Ann. Transl. Med.* **6**, 476.-476. (2018).
262. Tanaka, Y. *et al.* Progranulin regulates lysosomal function and biogenesis through acidification of lysosomes. *Hum. Mol. Genet.* **26**, 969–988 (2017).
263. Beel, S. *et al.* Progranulin functions as a cathepsin D chaperone to stimulate axonal outgrowth in vivo. *Hum. Mol. Genet.* **26**, 2850–2863 (2017).
264. Zhou, D. *et al.* PGRN acts as a novel regulator of mitochondrial homeostasis by facilitating mitophagy and mitochondrial biogenesis to prevent podocyte injury in diabetic nephropathy. *Cell Death Dis.* 2019 107 **10**, 1–16 (2019).
265. Chang, M. C. *et al.* Progranulin deficiency causes impairment of autophagy and TDP-43 accumulation. *J. Exp. Med.* **214**, 2611 (2017).
266. Oakes, J. A., Davies, M. C. & Collins, M. O. TBK1: a new player in ALS linking autophagy and neuroinflammation. *Mol. Brain* **10**, 1–10 (2017).
267. Weil, R., Laplantine, E., Curic, S. & Génin, P. Role of Optineurin in the Mitochondrial Dysfunction: Potential Implications in Neurodegenerative Diseases and Cancer. *Front. Immunol.* **9**, (2018).
268. Wightman, D. P. *et al.* A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nat. Genet.* 2021 539 **53**, 1276–1282 (2021).
269. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat. Genet.* 2022 544 **54**, 412–436 (2022).
270. Lee, B., Yao, X. & Shen, L. Genome-Wide association study of quantitative biomarkers identifies a novel locus for alzheimer's disease at 12p12.1. *BMC Genomics* **23**, 1–13 (2022).

271. Kunkle, B. W. *et al.* Novel Alzheimer Disease Risk Loci and Pathways in African American Individuals Using the African Genome Resources Panel: A Meta-analysis. *JAMA Neurol.* **78**, 102–113 (2021).
272. Bertram, L. & Tanzi, R. E. Genome-wide association studies in Alzheimer's disease. *Hum. Mol. Genet.* **18**, R137 (2009).
273. Homann, J. *et al.* Genome-Wide Association Study of Alzheimer's Disease Brain Imaging Biomarkers and Neuropsychological Phenotypes in the European Medical Information Framework for Alzheimer's Disease Multimodal Biomarker Discovery Dataset. *Front. Aging Neurosci.* **14**, 195 (2022).
274. Schwartzenuber, J. *et al.* Genome-wide meta-analysis, fine-mapping, and integrative prioritization identify new Alzheimer's disease risk genes.
275. Nalls, M. A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* **46**, 989–993 (2014).
276. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet. Neurol.* **18**, 1091–1102 (2019).
277. Alfradique-Dunham, I. *et al.* Genome-Wide Association Study Meta-Analysis for Parkinson Disease Motor Subtypes. *Neurol. Genet.* **7**, e557 (2021).
278. Chang, D. *et al.* A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* **49**, 1511–1516 (2017).
279. Foo, J. N. *et al.* Identification of Risk Loci for Parkinson Disease in Asians and Comparison of Risk Between Asians and Europeans: A Genome-Wide Association Study. *JAMA Neurol.* **77**, 746–754 (2020).
280. Foo, J. N. *et al.* Genome-wide association study of Parkinson's disease in East Asians. *Hum. Mol. Genet.* **26**, 226–232 (2017).
281. Fan, L. *et al.* Analysis of 12 GWAS-Linked Loci With Parkinson's Disease in the Chinese Han Population. *Front. Neurol.* **12**, 404 (2021).
282. (IPDGC), I. P. D. G. C. & (WTCCC2), W. T. C. C. C. 2. A Two-Stage Meta-Analysis Identifies Several New Loci for Parkinson's Disease. *PLoS Genet.* **7**, e1002142 (2011).
283. Lill, C. M. & Bertram, L. Towards unveiling the genetics of neurodegenerative diseases. *Semin. Neurol.* **31**, 531–541 (2011).
284. Weydt, P. & La Spada, A. R. Targeting protein aggregation in neurodegeneration –

- lessons from polyglutamine disorders.
<http://dx.doi.org/10.1517/14728222.10.4.505> **10**, 505–513 (2006).
285. Ahmed, R. M. *et al.* Neuronal network disintegration: common pathways linking neurodegenerative diseases. *J. Neurol. Neurosurg. Psychiatry* **87**, 1234–1241 (2016).
 286. Kamat, P. K., Kalani, A., Kyles, P., Tyagi, S. C. & Tyagi, N. Autophagy of mitochondria: a promising therapeutic target for neurodegenerative disease. *Cell Biochem. Biophys.* **70**, 707–719 (2014).
 287. Goh, K. Il & Choi, I. G. Exploring the human diseasome: the human disease network. *Brief. Funct. Genomics* **11**, 533–542 (2012).
 288. Tian, R. *et al.* CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron* **104**, 239-255.e12 (2019).
 289. Liscic, R. M. Als and Ftd: Insights into the disease mechanisms and therapeutic targets. *Eur. J. Pharmacol.* **817**, 2–6 (2017).
 290. Liscic, R. M., Alberici, A., Cairns, N. J., Romano, M. & Buratti, E. From basic research to the clinic: innovative therapies for ALS and FTD in the pipeline. *Mol. Neurodegener.* 2020 151 **15**, 1–17 (2020).
 291. Kalmar, B., Lu, C. H. & Greensmith, L. The role of heat shock proteins in Amyotrophic Lateral Sclerosis: The therapeutic potential of Arimoclomol. *Pharmacol. Ther.* **141**, 40–54 (2014).
 292. Boxer, A. L. *et al.* Frontotemporal degeneration, the next therapeutic frontier: Molecules and animal models for frontotemporal degeneration drug development. *Alzheimer's Dement.* **9**, 176–188 (2013).
 293. Wang, P., Wander, C. M., Yuan, C. X., Bereman, M. S. & Cohen, T. J. Acetylation-induced TDP-43 pathology is suppressed by an HSF1-dependent chaperone program. *Nat. Commun.* **8**, (2017).
 294. Hochgräfe, K. *et al.* Preventive methylene blue treatment preserves cognition in mice expressing full-length pro-aggregant human Tau. *Acta Neuropathol. Commun.* **3**, 25 (2015).
 295. Liscic, R. M., Grinberg, L. T., Zidar, J., Gitcho, M. A. & Cairns, N. J. ALS and FTLD: two faces of TDP-43 proteinopathy. *Eur J Neurol* **15**, 772–780 (2008).
 296. Rutherford, N. J. *et al.* Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS Genet* **4**, e1000193 (2008).

297. Rhinn, H., Tatton, N., McCaughey, S., Kurnellas, M. & Rosenthal, A. Progranulin as a therapeutic target in neurodegenerative diseases. *Trends Pharmacol. Sci.* **0**, (2022).
298. Castillo, K. *et al.* Measurement of autophagy flux in the nervous system in vivo. *Cell Death Dis.* **2013 411 4**, e917–e917 (2013).
299. Mojsilovic-Petrovic, J. *et al.* FOXO3a Is Broadly Neuroprotective In Vitro and In Vivo against Insults Implicated in Motor Neuron Diseases. *J. Neurosci.* **29**, 8236 (2009).
300. Nagashima, T. *et al.* Discovery of Novel Forkhead Box O1 Inhibitors for Treating Type 2 Diabetes: Improvement of Fasting Glycemia in Diabetic db/db Mice. *Mol. Pharmacol.* **78**, 961–970 (2010).
301. Altieri, D. C., Stein, G. S., Lian, J. B. & Languino, L. R. TRAP-1, the mitochondrial Hsp90. *Biochim. Biophys. Acta* **1823**, 767–773 (2012).
302. Hong, D. S. *et al.* Targeting the molecular chaperone heat shock protein 90 (HSP90): lessons learned and future directions. *Cancer Treat. Rev.* **39**, 375–387 (2013).
303. Kong, A. *et al.* Phase 1B/2 study of the HSP90 inhibitor AUY922 plus trastuzumab in metastatic HER2-positive breast cancer patients who have progressed on trastuzumab-based regimen. *Oncotarget* **7**, 37680–37692 (2016).
304. Banerji, U. *et al.* Phase I pharmacokinetic and pharmacodynamic study of 17-allylamino, 17-demethoxygeldanamycin in patients with advanced malignancies. *J. Clin. Oncol.* **23**, 4152–4161 (2005).
305. Scrivo, A., Bourdenx, M., Pampliega, O. & Cuervo, A. M. Selective autophagy as a potential therapeutic target for neurodegenerative disorders. *Lancet Neurol.* **17**, 802–815 (2018).
306. Xu, W. *et al.* Selective autophagy as a therapeutic target for neurological diseases. *Cell. Mol. Life Sci.* **78**, 1369–1392 (2021).
307. del Campo, M. *et al.* New developments of biofluid-based biomarkers for routine diagnosis and disease trajectories in frontotemporal dementia. *Alzheimer's Dement.* (2022) doi:10.1002/ALZ.12643.
308. Ehrenberg, A. J. *et al.* Relevance of biomarkers across different neurodegenerative. *Alzheimer's Res. Ther.* **12**, 1–11 (2020).

