Investigating the spread and toxicity of glycine-alanine dipeptides in C9orf72 ALS/FTD using *Drosophila melanogaster*

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Javier Morón Oset
aus Madrid, Spanien

Köln, 2020

Gutachter: **Prof. Dr. Linda Partridge**
Prof. Dr. Elena Rugarli

Tag der mündlichen Prüfung: 02. Dezember 2020

Table of Contents

Acknowledgements V

Abbreviations VII

Summary IX

1. INTRODUCTION 1

1.1. Neurodegenerative disorders 2

1.1.1 Commonalities among neurodegenerative disorders: role of ageing and protein deposits 2

1.1.2 Current challenges posed by neurodegenerative diseases in today's ageing society .. 6

1.2. Mechanisms for pathology dissemination during disease progression 7

1.2.1 Selective neuronal vulnerability vs cell-to-cell pathogenic spread 7

1.2.2 Seeding, release and uptake of pathological proteins 12

1.3. The *C9orf72* gene and neurodegenerative diseases 13

1.3.1 Amyotrophic Lateral Sclerosis (ALS): pathobiology, epidemiology and therapeutics 14

1.3.1.1 Clinical presentations and selective vulnerability of somatic motor neurons 14

1.3.1.2 Environmental and genetic risk factors for ALS 15

1.3.1.3 TDP43 pathology is the overarching hallmark lesion of ALS 16

1.3.1.4 Approved drugs and novel therapeutic avenues under investigation for ALS 17

1.3.2 Frontotemporal Dementia (FTD): pathobiology, epidemiology and therapeutics 18

1.3.2.1 Clinical variants and vulnerability of the frontotemporal lobes 18

1.3.2.2 Environmental and genetic risk factors for FTD 19

1.3.2.3 Neuropathological signatures of FTD: tau, TDP43 and FUS 20

1.3.2.4 Approved drugs and novel therapeutic avenues under investigation for FTD 20

1.3.3 The C9orf72 gene 21

1.3.3.1 The discovery of the C9orf72 mutation in fALS and fFTD 21

1.3.3.2 Neuropathological and clinical hallmarks in C9orf72 mutation carriers 22

1.3.3.3 Toxicity perpetrators of the C9orf72 mutation: experimental and human evidence 25

1.3.3.4 The role of repeat length in C9orf72-associated toxicity: clinical and preclinical evidence 28

1.3.3.5 Glycine-alanine (GA) DPRs: synthesis, molecular features and toxicity mechanisms 31

1.4 Aims of the PhD Thesis 34

2. MATERIALS AND METHODS 35

2.1. Fly work 36

2.1.1 Maintenance of flies 36
2.1.2 Transgenic flies used in this study 36
2.1.3 Fly genetics 37
2.1.4 Egg-to-adult viability assay and eye phenotypes 37
2.1.5 Lifespan assay 37
2.2. Molecular biology 38
2.2.1 Generation of transgenic fly lines 38
2.2.1.1 Commonalities 38
2.2.1.2 Generation of mCherry-tagged DPR36 and DPR100 constructs 38
2.2.1.3 Generation of GFP- and FLAG-tagged GA100 40
2.2.1.4 Generation of untagged or mCherry-tagged GA200 42
2.2.1.5 Generation of GA400 42
2.2.2 Fly genotyping 43
2.2.3 RNA extraction, cDNA synthesis and Quantitative Real-Time PCR (qRT-PCR) 43
2.3. Protein biochemistry 43
2.3.1 Soluble-insoluble protein fractionation and western blotting 43
2.3.2 Dot blotting 44
2.3.3 In-gel proteasome activity assay and subsequent protein detection 45
2.4. Histology 46
2.4.1 Immunostainings of adult Drosophila brains 46
2.4.2 Stainings of adult Drosophila antennae 46
2.4.3 Phalloidin stainings of adult Drosophila brains 47
2.4.4 TUNEL assay 47
2.4.5 Imaging of adult Drosophila brains 47
2.4.6 Quantification of confocal images 48
2.5. Statistical analysis 49
3. RESULTS 50
3.1 Spread of GA DPRs in the fly brain 52
3.1.1 GA DPRs spread rapidly in a repeat length- and age- dependent manner in the fly brain 52
3.1.1.1 Generation and validation of DPR-mCherry constructs 52
3.1.1.2 GA100 DPRs, but not GR100 or PR100 DPRs, spread rapidly in the fly brain 57
3.1.1.3 GA repeat length modulates the aggregation and spread of GA DPRs 62
3.1.1.4 GA DPRs exhibit an age-related increase in spreading 69
3.1.2 Consequences of GA transmission in the fly brain 71
3.1.3 Mechanisms of GA transmission in the fly brain 75

3.1.3.1 GA200 transmission from ORNs is reduced upon down-regulation of genes involved in exosomal release 75
3.1.3.2 GA200 transmission from ORNs is affected by modulating genes involved in exocytosis and neuronal activity 78
3.2 Effect of tags on GA toxicity, aggregation pattern and cellular responses in the fly brain 83
3.2.1 Tags reduce GA100-mediated lifespan shortening upon pan-neuronal expression .. 83
3.2.2 Large tags affect GA aggregation profile without interfering with puncta formation in the fly brain 86
3.2.3 Large tags interfere with cellular consequences of GA expression in the fly brain 89
3.3 Role of repeat length in GA toxicity, aggregation pattern and cellular responses in the fly brain 93
3.3.1 GA400 DPRs are more toxic than GA100 or GA200 DPRs 93
3.3.1.1 Generation and validation of GA400 fly lines 93
3.3.1.2 GA400 DPRs are more toxic than GA100 and GA200 DPRs upon expression in neuronal tissue 95
3.3.2 The subcellular location and aggregation propensity of GA DPRs is repeat-length-dependent 100
3.3.3 The repeat length of GA DPRs affects their cellular responses in a non-linear manner 103
3.3.3.1 GA expression modulates Ref(2)P levels in a repeat length-dependent manner 103
3.3.3.2 GA100 and GA200 exacerbate DNA damage more than GA400 108
4. Discussion 111
4.1. Spread of GA DPRs in the fly brain 112
4.1.1 GA DPRs spread rapidly in a repeat length- and age-dependent manner in the fly brain 112
4.1.2 GA transmission is associated with lower Ref(2)P levels in recipient tissue 116
4.1.3 GA transmission depends on exocytosis of exosomes and synaptic vesicles 118
4.2. Tags affect GA toxicity, aggregation pattern and cellular responses 120
4.2.1 Tags reduce GA100-mediated lifespan shortening upon pan-neuronal expression 121
4.2.2 Large tags modify GA aggregation 122
4.2.3 Large tags interfere with cellular responses triggered by GA100 123
4.3. Repeat length affects GA toxicity, aggregation pattern and cellular responses 124
4.3.1 GA400 is more toxic than GA100 and GA200 125
4.3.2 The subcellular location and aggregation propensity of GA DPRs is repeat length-dependent 126
4.3.3 GA DPRs may cause toxicity by affecting repeat length-dependent mechanisms ... 127
5. List of figures 130

6. List of tables 132
7. Contributions 133
8. Bibliography 134
9. Signed thesis declaration 149
10. Publications 150

Acknowledgements

First of all, I want to thank Prof. Dr. Linda Partridge for her mentorship, trust and constant feedback. Thank you for your kind words when the first projects that I started working on were not going anywhere, and reminding me of the relevance of staying focused to be efficient and to produce good quality of work in research. I am also very grateful for the freedom that you gave me from the beginning to design and pursue my projects, as well as to attend conferences to stay abreast of the field. A very heart-felt “danke schön” also goes to Dr. Sebastian Grönke for his supervision, guidance and inspiration. Thanks to our long scientific discussions, I have become a more self-critical, but also self-confident researcher.

I would like to thank Prof. Dr. Elena Rugarli, Prof. Dr. Jan Riemer and Dr. Joris Deelen for being part of my PhD thesis committee, and my TAC members Dr. David Vilchez and Dr. Natalia Kononenko for their constructive feedback.

I also have to acknowledge the great crew of people that I have been extremely lucky to share these last 4 years with. A very special thank you to my dearest friends (DJ) Víctor and Carinchen for helping me kick off in the lab. I cannot put into words how thankful I am for your patience and constant willingness to teach me various lab techniques and to help me design my experiments. I also have to acknowledge my good and inspiring friend Maarouf. Thank you so much for always taking the time to discuss results, neuroscience papers and ideas with me. You’re an amazing scientist and I have honestly learnt so much from you (so, Lebanon does NOT equal desert). A big shout-out to Paulinchen, for our awesome conversations in the fly lab, to Lisa, for always being there to listen to my problems and helping so much with German paperwork, to Oliver (I know you’re going to kill me for not writing Dr. Hahn, well there it is!), for being a supportive friend and helping me become a more ambitious scientist, to Dr. Luke Tain, for being there to discuss experiments and lecture me on weird Spanish customs that only he knows about, and to Ralf, for always being willing to answer my lab questions and enabling me to experience *Giardia intestinalis* infection first-hand ☺. Thanks to Dr. Joris Deelen, for your constructive feedback and support and for being my tennis know-it-all buddy, to Dr. Jerome Korzelius, for your enthusiastic attitude and making me aware of NEB stable bacteria (life saver!), to Sandra, for your kindness, fun spirit and help to improve my Deutsch, to Jacky, for always being there to discuss protocols and being an awesome lab bench buddy. I also have to deeply thank Dr. Christine Lesch, for her constant willingness to help me solve bureaucratic problems and making our lab meetings, retreats and a few other million things possible. Thank you to Oliver Hendrich for essentially keeping the lab afloat and always being there to discuss protocols. Gracias to everyone else in the LP Department for, some way or another, having helped me develop my projects.

I also have to thank the three talented students that I feel honoured to have supervised: Simone, Tessa and Lilly. I hope you (will) remember fondly your time in the lab. How fun it was (is) to work with you!

I would also like to acknowledge the FACS & Imaging Facility at the MPI AGE, especially Dr. Christian Kukat and Marcel, for being so helpful to answer my microscopy questions. I am also very thankful to the CGA coordinators and directors for putting together such a fantastic programme. A very special thanks to Dr. Daniela Morrick for literally ALWAYS being happy to

answer my questions and have a chat. I also want to acknowledge Prof. Dr. Adrian Isaacs for a lot of constructive feedback.

And last but not least, I have to thank the people who are there at home, in Cologne and in Spain, and without whom there is no way I would have not lost my sanity during the past years. I can simply not express how thankful I am to mi gordi, Almu, for being the most supportive, loving and whole-package girlfriend&scientist on the planet. I would need a separate thesis to do justice to your moral contribution to this work. Te amo. Mamá, papá, Marty, Nacho, yaya, prim@s, thank you for your unmeasurable moral support (but not so much for reminding me of how much better food and the weather are in Madrid). Os quiero. Thank you, mamá, for being my scientific role model. Thanks to my Copic and Domyos friends for making life a lot more fun.

Abbreviations

Aβ	Amyloid β
ACC	Anterior cingulate cortex
AD	Alzheimer's disease
ALS	Amyotrophic lateral sclerosis
ASO	Antisense oligonucleotide
BAC	Bacterial artificial chromosome
BMN	Bulbar motor neuron
BvFTD	Behavioural-variant frontotemporal dementia
C9orf72	Chromosome 9 open reading frame 72
CA	Constitutively active
CBD	Corticobasal degeneration
DMPK	Dystrophia myotonica protein kinase
DN	Dominant negative
DPR	Dipeptide repeat
DsDNA	Double-stranded deoxyribonucleic acid
DTI	Diffusion tensor imaging
Eag	Ether-a-gogo
FI	Fronto-insular cortex
FMRI	Funcional magnetic resonance imaging
FTD	Frontotemporal dementia
FTLD	Frontolobar degeneration
FUS	Fused to sarcoma
G4C2	GGGGCC
GA	Glycine-alanine
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
GOF	Gain-of-function
GP	Glycine-proline
GR	Glycine-arginine
GRN	Progranulin
HD	Huntington's disease
HnRNP3	Heterogeneous nuclear ribonucleoprotein 3
HRE	Hexanucleotide repeat expansion
Hsp	Heat shock protein
LBD	Lewy body dementia
LMN	Lower motor neuron
LOF	Loss-of-function
mHtt	Mutant Huntingtin
MNCs	Median neurosecretory cells
MRI	Magnetic resonance imaging
NCT	Nucleocytoplasmatic transport
NSF1	N-ethylmaleimide-sensitive fusion protein 1
OL	Optic lobe
OPTN	Optineurin
OrkΔ1	Open rectifier potassium channel 1

ORNs	Olfactory receptor neurons
P62	Protein 62
PA	Proline-alanine
pATM	Phosphorylated ataxia telangiectasia mutation
PD	Parkinson's disease
PET	Positron Emission Tomography
PNFA	Primary non-fluent aphasia
PR	Proline-arginine
PrP	Prion protein
PSP	Progressive supranuclear palsy
P-Tau	Hyperphosphorylated tau
PTM	Post-translational modification
Rab	Ras-associated binding
RAN	Repeat-associated non-AUG translation
Ref(2)P	Refractory to sigma P
Rpt6	Regulatory particle triple ATPase 6
RS	Restriction site
RT	Room temperature
SC	Spinal cord
SD	Semantic dementia
Shi	Shibire
SMN	Spinal motor neuron
SNP	Single nucleotide polymorphism
SNpc	Substantia nigra pars compacta
SOD1	Superoxide dismutase 1
Stip1	Stress-induced phosphoprotein 1
SV2	Synaptic vesicle 2
SYA	Sugar-yeast-agar
Syntx	Syntaxin
TARDBP	Transactive response DNA binding protein
TBPK1	TANK-binding kinase 1
TDP43	Transactive response DNA binding protein 43
TUNEL	Terminal deoxynucleotidyl transferase dUTP nick end labeling
UMN	Upper motor neuron
VEN	Von economo neurons
VTA	Ventral tegmental area
WT	Wild-type

Summary

Hexanucleotide repeat expansions of variable size in *C9orf72* are the most prevalent genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). The role of repeat size in disease onset and severity in humans remains controversial. Transcripts of the expansions are translated into five dipeptide repeat (DPR) proteins. Most preclinical studies have used relatively short and tagged poly-DPR constructs to investigate DPR-mediated toxicity, and shown that poly-GR, poly-PR and, to a lesser extent, poly-GA DPRs are neurotoxic. Consequently, a major emphasis has been placed on understanding poly-GR- and poly-PR-mediated toxicity. However, poly-GA is the most abundant DPR in patient tissue. Transmission of protein aggregates may be a major driver of toxicity in neurodegeneration. In this study, I show for the first time that only poly-GA DPRs can spread trans-neuronally *in vivo* using the adult fly brain. Repeat length and tissue age modulate this phenomenon, and exosomes and synaptic vesicles are relevant in the extracellular release of GA DPRs.

I also compared the toxicity, aggregation and cellular responses of GA100 DPRs carrying or not commonly used tags. Expression of tagged GA100 was markedly less toxic. GA100 tagged with GFP and mCherry exhibited aggregation differences and failed to cause DNA damage or proteostasis stress compared to untagged GA100 and GA100FLAG. These findings highlight the need to use untagged DPRs as controls when investigating their pathobiology.

Finally, I tested the role of repeat size in modulating GA toxicity, subcellular localization, aggregation and cellular responses by comparing these in flies expressing untagged GA100, GA200 and GA400 DPRs. While aggregation propensity and proteostasis stress hold a positive correlation with repeat length, and GA400 was markedly more toxic than GA100, the latter was in turn more toxic than GA200. This highlights a non-linear correlation between repeat length and toxicity. GA100 and GA200 formed numerous puncta-like aggregates both in the soma and axons of neurons and, especially GA200, exhibited spreading, whereas GA400 resided only in somata and did not spread. Surprisingly, GA200 caused more DNA damage than GA100, but this effect was not observed upon GA400 expression.

Collectively, I show that GA DPRs have a unique ability to spread *in vivo*, and their toxicity may have been previously underestimated by the use of short and tagged constructs. Therefore, my data support the further characterization of GA DPRs of a clinically relevant composition to develop strategies with therapeutic potential for *C9orf72* mutation carriers.

1. INTRODUCTION

1.1. Neurodegenerative disorders

Memory loss, speech difficulty, paralysis, mood changes. These are some examples of the symptoms that affected more than 200 million Europeans suffering from a neurological condition in 2017 (Raggi and Leonardi 2020). A large portion of the most common neurological disorders comprises neurodegenerative diseases. These are devastating pathological conditions that stem from the progressive impairment and eventual demise of neurons and synapses in specific regions of the central and peripheral nervous systems. These disorders typically progress in an irreversible manner and have large socioeconomic and personal costs. The most common neurodegenerative disorders are Alzheimer's disease (AD) and Parkinson's disease (PD). Some other classical examples relevant for this doctoral thesis include Amyotrophic Lateral Sclerosis (ALS), Frontotemporal Dementia (FTD) and Huntington's disease (HD) (Mure and Jellinger 2010).

1.1.1 Commonalities among neurodegenerative disorders: role of ageing and protein deposits

While the clinical presentation of these disorders can vary a great deal depending on the regions affected by neuronal dysfunction (Table 1), there are several traits that they all have in common. Firstly, these are multifactorial diseases whose onset and progression are influenced by diverse genetic and environmental factors (Table 1). A wide spectrum of genes has been associated with increased risk of developing these conditions (Table 1), with a typically small percentage of cases being accounted for by dominant mutations that define the familial cases of these disorders (Bertram and Tanzi 2005) (Table 1). The latter are mostly disease-specific and have been instrumental in developing the majority of the experimental models that have been, and are currently used, to investigate the complex aetiology of these disorders and to test therapeutic strategies. However, sporadic cases are by far the norm and numerous environmental factors have also been correlated with increased risk, including exposure to specific toxins and lifestyle habits, such as smoking, diet, sport practice and education level (Brown, Lockwood, and Sonawane 2005) (Table 1). In fact, disease presentation and severity can substantially vary within the same family of mutation carriers, which highlights the relevance of environmental factors in determining inter-individual variability (Xi et al. 2014). Furthermore, certain non-neurological disorders, such as mid-life

obesity, hypertension and heart disease, increase the probability of developing neurodegenerative diseases (Whitmer et al. 2005) (Table 1).

Ageing, defined as the time-dependent deterioration of physiological integrity that affects most organisms, is the main risk factor for developing the majority of these disorders (Hou et al. 2019). While some extreme familial cases have been reported to show symptoms in their 20s, this is uncommon and disease onset usually occurs at the age of 40-50 at the earliest, with greater risk and exacerbated clinical presentation as individuals age. For instance, AD prevalence between 65-69 years of age in the US is approximately 20 per 1,000 individuals, and this increases to roughly 450 per 1,000 individuals aged over 95 (Hou et al. 2019). This correlation between increasing age and exacerbated risk for neurodegenerative conditions has led to the hypothesis that neurodegenerative diseases may simply reflect accelerated brain ageing in afflicted individuals (Wyss-Coray 2016). Of note, in some diseases, such as PD, neuronal demise commences decades earlier than symptoms (Cheng, Ulane, and Burke 2010), suggesting that human brains can maintain functionality despite severe neuronal connectivity loss and ageing. This is also supported by magnetic resonance imaging (MRI) studies showing that specific brain regions of infants that carry susceptibility alleles for sporadic neurodegenerative diseases are significantly different from non-carriers (Dean et al. 2014). Therefore, the loss of neuroplasticity during ageing may be the key factor for the onset of symptoms at advanced ages despite disease commencing much earlier.

Nine critical “hallmarks of ageing” have been defined based on their occurrence during normal ageing and the possibility of influencing healthspan and lifespan upon their experimental manipulation. These can be categorized into three subgroups. (López-Otín et al. 2013). First, genomic instability, telomere attrition, epigenetic alterations and loss of proteostasis are regarded as the main driving forces of the ageing process and are therefore termed as the “primary hallmarks”. Second, mitochondrial dysfunction, cellular senescence and deregulated nutrient sensing constitute pathways that cells exploit to mitigate initial damage following disruption of the “primary hallmarks” of the ageing process, thus representing the “antagonistic hallmarks”. However, their excessive activity upon chronic activation results in them becoming deleterious themselves. Third, stem cell exhaustion and altered intercellular communication, known as the “integrative hallmarks”, arise when tissue homeostatic mechanisms cannot withstand the cumulative damage inflicted by the “primary and antagonistic hallmarks” (López-Otín et al. 2013).

Table 1. Summary of epidemiological, clinical and neuropathological traits of AD, PD, ALS and bvFTD.

These data are variable across different geographical populations and disease stages. Sources: *1 = (Prusiner 2001), *2 = (Fu, Hardy, and Duff 2018), *3 = (Takatori et al. 2019), *4 = (Klein and Westenberger 2012).

Disease	Mean age of onset	Prevalence (per 100,000)*1	Major initial symptoms	Genes of familial forms	Genes associated with increased risk	Environmental risk factors	Nature of protein deposits	Most vulnerable neurons*2
AD	≈ 80	1,450	memory loss, learning difficulties	*APP, PSEN1, PSEN2*	*APOE4, CLU, PICALM, TREM2, BIN1*3*	Smoking, hypertension, low educational background, brain injuries	Amyloid-β, tau	large pyramidal neurons in entorhinal cortex
PD	≈ 65	360	tremors, muscle rigidity, bradykinesia, postural impairment	*SNCA, LRRK2, GBA PINK1, PARKIN, DJ-1* *4	*SNCA, UCHL1, LRRK2, PARK 16, GAK*4*	Exposure to pesticides, use of antidepressants and psychostimulants, including metamphetamines and cocaine	α-synuclein	dopaminergic neurons in substantia nigra pars compacta
ALS	≈ 55	1-7	Lack of control of voluntary muscles	*C9orf72, SOD1, FUS, TARDBP*	*C9orf72, SOD1, FUS, TARDBP, UNC13A, ATXN1, ATXN2, EPHA4*	smoking, male gender and strenuous physical activity	TDP43	fast-fatigable motor neurons in spinal cord, brainstem and motor cortex
bvFTD	≈ 70	14	Personality and behavioural changes, language deficits	*MAPT, GRN, C9orf72, VCP, CHMP2B, OPTN, FUS*	*MAPT, GRN, C9orf72*	Traumatic brain injuries, high educational background	TDP43, tau, FUS	von Economo neurons in anterior cingular cortex

While all of the hallmarks of the ageing process influence brain ageing and the risk of developing a neurodegenerative disease (Hou et al. 2019), loss of proteostasis is likely to play a pivotal role. Aged brains typically exhibit accumulation of misfolded proteins and abnormal organelles involved in proteostasis, such as lysosomes and autophagosomes. In addition, abundant RNA-protein structures known as stress granules, as well as subcellular structures, such as lipofuscin, consisting of oxidized proteins, lipids and carbohydrates accumulate in the aged brain, probably as a result of lysosomal dysfunction (Wyss-Coray 2016). The nature of the aggregated proteins is variable (Table 1) and many of them can also co-exist in the same disorder (Robinson et al. 2018), potentially exerting an additive or synergistic effect on pathology and toxicity (Z. He et al. 2018). Aggregates primarily formed by the same protein are found in groups of neurodegenerative diseases, thus constituting the hallmark lesion of that family of disorders. For instance, this is the case for the microtubule-associated protein tau, whose insoluble deposits characterize the so-called tauopathies, including AD, corticobasal degeneration (CBD) and progressive supranuclear palsy (PSP). While all of these are characterized by the accumulation of aggregated tau, tauopathies significantly differ in their clinical presentation, as well as in the morphology and location of their protein aggregates and affected cell types (Höglinger, Respondek, and Kovacs 2018). This has led to the well-established notion that protein deposits can form different aggregate strains depending on the co-aggregated proteins and/or structure, thus rendering specific cell types more vulnerable and accounting for the clinical heterogeneity of diseases sharing a common protein as their hallmark lesion (Kaufman et al. 2016; Narasimhan et al. 2017; Sanders et al. 2014).

Protein aggregates have hogged the spotlight in this field for some decades now because of the close correlation between the brain areas where these protein aggregates progressively accumulate, the brain regions most affected by neurodegeneration and the clinical symptoms shown by patients (Nelson et al. 2012). In fact, the accurate diagnosis of some of these disorders has long been possible exclusively at postmortem when examining the location and nature of brain protein deposits (Bang, Spina, and Miller 2015; Heiko Braak et al. 2003). Great ongoing efforts to diagnose these diseases at an early stage are focused on developing non-invasive strategies to detect the levels and distribution of these specific proteins in a disease-specific manner (Palmqvist et al. 2020; Vogels et al. 2020). Furthermore, many of the genes where mutations are associated with dominant penetrance directly impinge on the

aggregation propensity of these proteins (Johnson et al. 2009). However, their contribution to disease has also been somewhat disputed given that many of these proteins, such as amyloid-β (Aβ), hyperphosphorylated tau (p-tau), α-synuclein and transactive response DNA binding protein 43 (TDP43), also accumulate in the brains of aged individuals that do not exhibit clear signs of neurodegenerative diseases (Elobeid et al. 2016). In addition, several therapeutic strategies that reduce the brain load of these proteins have failed to halt disease progression (Panza et al. 2019). Nevertheless, there is a consensus that brain protein aggregates play a key role in initiating disease, but other ageing-related factors, such as increased inflammation (Glass et al. 2010), as well as the presence of compensatory, protective mechanisms, are also pivotal in mediating disease risk and progression.

1.1.2 Current challenges posed by neurodegenerative diseases in today's ageing society

Thanks to major dietary and healthcare improvements, mean human lifespan expectancy has essentially doubled in the last 200 years (Partridge, Deelen, and Slagboom 2018) and is currently 72 years of age on a global scale according to the World Health Organization (see https://apps.who.int/gho/data/view.main.SDG2016LEXREGv?lang=en). According to estimates by the United Nations, 1 in 11 individuals worldwide were aged at least 65 in 2019, and by 2050 this proportion is projected to rise to 1 in 6 individuals (United Nations, World Population Ageing 2019), which clearly shows that the elderly will be an increasingly prevalent demographic group. For example, it is estimated that the US population aged above 65 will increase from 56 million in 2020 to 88 million by 2050 (Alzheimer'sAssociation 2020). While these figures clearly emphasize the key role that scientific advances can play in promoting human lifespan, they pose the challenge that the population percentage afflicted with age-related diseases, including neurodegenerative disorders, will continue to increase in the coming years. Apart from the terrible personal consequences, this also entails a substantial financial burden, with brain disorders being particularly costly. Only in 2010 brain disorders required a cost of 800 billion euros in Europe, with cardiovascular disorders and cancer accounting for roughly a fourth of this cost each (DiLuca and Olesen 2014).

To meet the need to develop strategies that can efficiently prevent, slow down or cure neurodegenerative diseases, scientists have generated a plethora of disease-specific *in vitro* and *in vivo* experimental models. The majority of these models rely on the endogenous

expression or over-expression of mutation-carrying human alleles associated with familial neurodegenerative diseases. This is per se a limitation of the currently available tools to investigate the pathobiology of many of these diseases because, albeit not always, familial cases often exhibit more aggressive phenotypes than their sporadic counterparts (Naumann et al. 2019). However, as many of these systems are genetically malleable and given the great advances in genetic manipulations, more and more accurate models are being developed to investigate specific aspects of each disease. In addition, great efforts are being devoted to understanding how central biological mechanisms of ageing could be targeted to promote healthy brain ageing that would result in lower risk for neurodegenerative diseases. Equally relevant is the advent of non-invasive biomarkers to track disease progression and therefore to faithfully test the efficacy of therapies. Altogether, ongoing endeavors aim to reverse the current situation of no or very few effective treatments for neurodegenerative diseases.

1.2. Mechanisms for pathology dissemination during disease progression

One key aspect of neurodegenerative diseases that has traditionally struck neurologists is the progressive and changing nature of symptoms shown by afflicted patients over time. A good example of this is found in AD, where the first symptoms are typically memory loss and learning difficulties, followed by mood and personality changes and finally by impairment in language, visuospatial and executive functions (Thies and Bleiler 2012). But what are the underlying drivers of this progressive clinical deterioration at the molecular and cellular levels?

1.2.1 Selective neuronal vulnerability vs cell-to-cell pathogenic spread

Nearly 30 years ago a seminal histopathological study using numerous post-mortem human brains reported that there is a conserved spatiotemporal pattern in the location of p-tau deposits in the brains of AD patients, which strongly correlates with their clinical deterioration (H. Braak and Braak 1991). Similar findings have been presented for other neurodegenerative diseases ever since, including a strong association between α-synuclein aggregates and PD (Heiko Braak et al. 2003) and TDP43 inclusions and ALS and FTD (Brettschneider et al. 2013,

2014). More recent cross-sectional studies using positron emission tomography (PET) tracers have largely corroborated some of these findings in living patients (Maass et al. 2017; Schwarz et al. 2016; L. Wang et al. 2016), thus supporting the notion that the progressive accumulation of disease-specific protein aggregates in select brain regions plays a key role in the time-dependent clinical impairment of neurodegenerative disorders.

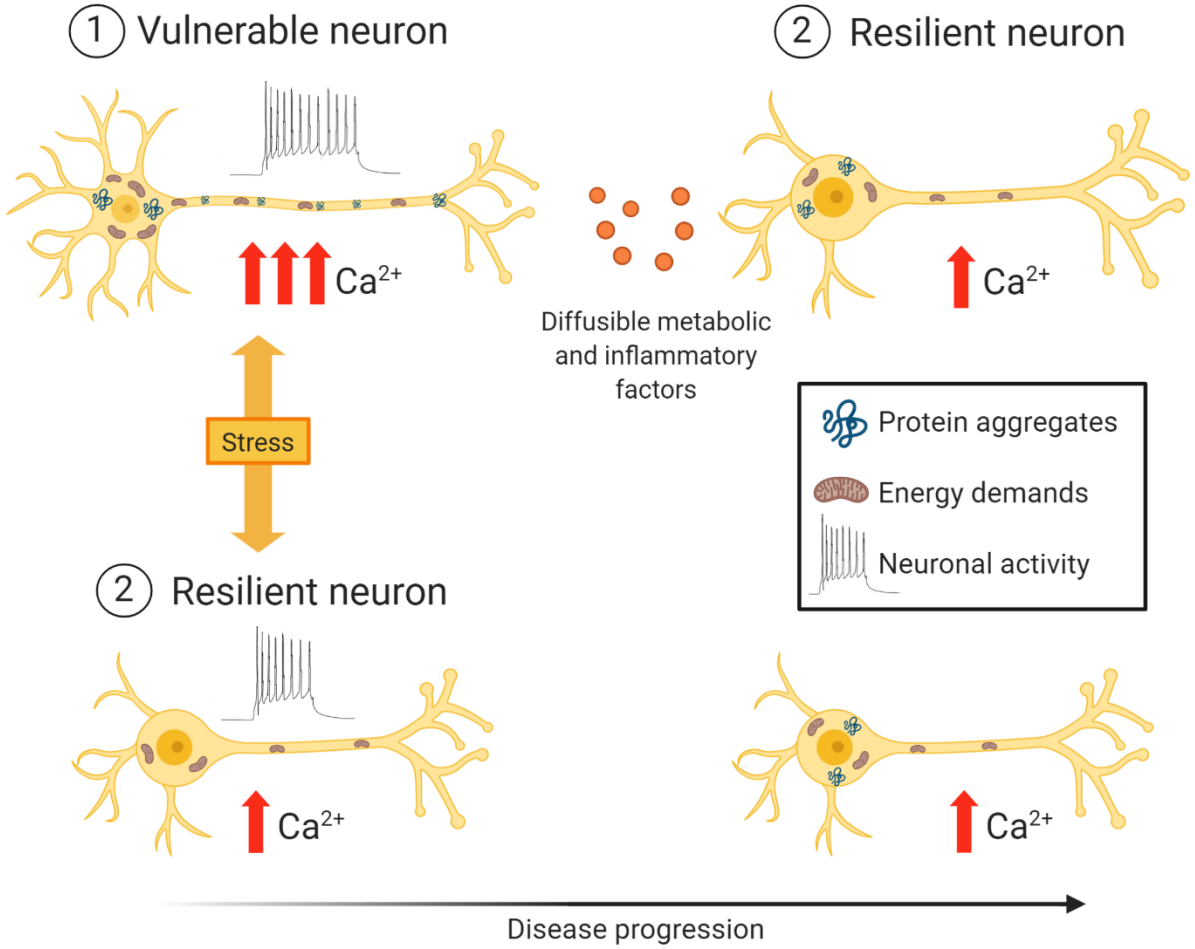

Figure 1. Selective neuronal vulnerability hypothesis.
Certain properties intrinsic to some neurons may make them more vulnerable to developing pathological lesions, such as protein aggregates, and degeneration. These include morphological features, such as having long axons and highly branched dendrites, as well as functional aspects, including low levels of calcium-buffering proteins and subsequent disturbed calcium homeostasis, higher energy demands and increased neuronal activity. Vulnerable neurons may release diffusible metabolic and inflammatory factors, thus making nearby or synaptically connected neurons less resistant to various kinds of stress. This might result in initially resilient neurons progressively developing similar lesions. Alternatively, resilient neurons may become vulnerable when their intrinsic properties become less advantageous at some point during disease progression (Fu, Hardy, and Duff 2018). This cartoon was created with Biorender.

Two mutually non-exclusive theories have been put forward to explain the stereotypical dissemination of pathology in neurodegenerative diseases. On the one hand, the selective neuronal vulnerability hypothesis proposes that, upon stress, select neurons are intrinsically more susceptible to developing specific lesions, such as the aggregation of a particular protein, and to subsequent degeneration through cell-autonomous mechanisms (Surmeier, Obeso,

and Halliday 2017) (Figure 1). In addition, this theory proposes that the transduction of diffuse metabolic and inflammatory factors from the most vulnerable cells to the initially less vulnerable ones could also contribute to the increasing number of cells affected by pathological lesions (Walsh and Selkoe 2016) (Figure 1). This hypothesis is strongly supported by the fact that all of the disease-associated proteins are expressed as soluble physiological entities by all neurons throughout life, yet they aggregate in selective regions during the course of particular diseases (Walsh and Selkoe 2016).

Despite our currently limited understanding of the factors that underlie the vulnerability to pathology of a particular neuronal subset, some intrinsic biochemical, morphological and electrophysiological properties are starting to be recognized (Figure 1). First, different neuronal subsets display distinct efficiency of clearance mechanisms that keep aggregation-prone proteins in check. For instance, a study reported a transcriptional, protein homeostasis signature in healthy brains that was specific to areas known to exhibit the hallmark protein aggregates of AD. This consisted of high levels of a specific subset of proteins that co-aggregate with Aβ and tau, as well as low levels of protein homeostasis components (Freer et al. 2016). In addition, mutant Huntingtin (mHtt) is cleared more readily by cortical neurons than by striatal neurons, the latter being particularly vulnerable in HD (Tsvetkov et al. 2013). Intrinsic subpopulation differences have also been extensively studied in the context of PD pathogenesis, where dopaminergic neurons within the substrantia nigra pars compacta (SNpc) strongly degenerate, whereas the very similar dopaminergic neurons in the ventral tegmental area (VTA) remain largely spared (Fu, Hardy, and Duff 2018). The former express lower levels of calcium-buffering proteins, such as calbindin, and display clear electrophysiological differences, e.g., SNpc dopaminergic neurons rely on calcium channels to maintain their spontaneous, pacemaker-like firing, whereas VTA dopaminergic neurons rely on sodium channels, which results in increased cytosolic calcium levels in the former (Brichta and Greengard 2014). Increased cytosolic calcium can cause oxidative stress and mitochondrial dysfunction and therefore, calcium dysregulation is thought to play a relevant role in neurodegeneration (Marambaud, Dreses-Werringloer, and Vingtdeux 2009). Furthermore, SNpc dopaminergic neurons exhibit longer and more branched axons, with a remarkable large number of synaptic spines, compared to VTA dopaminergic neurons. This results in the former having a larger energy demand and being more susceptible to deficits in energy supply, e.g., upon mitochondrial dysfunction (Brichta and Greengard 2014). Finally, differential expression

of neurotransmitter receptors may also contribute to vulnerability, as specific subunits of glutamatergic, GABAergic and dopaminergic receptors have been found in particularly vulnerable neurons. Since some of these features are individually shared by a large number of neurons within the human brain, it is the collective possession of these and other cell-autonomous factors that is most likely to contribute to the pattern of primary and secondary selective regional vulnerabilities in neurodegenerative diseases.

Cell non-autonomous processes may also contribute to disease progression. The pathogenic spread hypothesis postulates that there is a physical transmission of pathological proteins from neuron to neuron, and from brain region to region, to induce disease progression (Fu et al. 2019) (Figure 2). The ability of pathological proteins to physically travel from expressing cells to non-expressing ones has been widely tested. Two main types of *in vivo* models are worth highlighting: the injection of recombinant proteins or brain extracts from patients or animal models into specific locations of the rodent brain (Clavaguera et al. 2013; Kaufman et al. 2016; Smolek et al. 2019; Ulusoy et al. 2013) and the genetic expression of the pathological protein restricted to a precise brain region either through site-specific drivers (De Calignon et al. 2012; Liu et al. 2012) or injection of viral vectors (Asai et al. 2015). The appearance of pathological protein entities outside of the injection or expression areas, and frequently within synaptically connected brain regions, supports the active spread of pathological proteins across neuronal networks. Along with numerous *in vitro* studies, these *in vivo* models have corroborated experimentally the pathogenic spread hypothesis for, among others, tau, α-synuclein and TDP43 (Peng, Trojanowski, and Lee 2020). In support of this hypothesis, treatment of animal models with antibodies specific for some of these proteins reduce, among others, tau and α-synuclein pathology (Jadhav et al. 2019). Given that these are intracellular proteins, the effectiveness of immunotherapies is thought to at least partially stem from their ability to block the trans-cellular transmission of pathological seeds.

Despite the difficulties in testing the pathogenic spread hypothesis in living humans, the development of PET tracers for each spread-prone pathological protein, which is a challenge due to their varied ultrastructural conformations, and their longitudinal application to patients as their symptoms advance, are likely to be insightful. To date, two studies have addressed the trans-neuronal spread of tau in living AD patients. Combining longitudinal tau PET measurements and functional MRI (fMRI), a close association between functionally connected brain areas and similar tau accumulation over time was found (Franzmeier et al. 2020). Given

that brain regions that work together functionally are often, but not always, connected by axons, this study suggested that tau accumulation is dictated by synaptic connections. Even more recently, researchers used an *in silico* approach to model tau spread starting from the entorhinal cortex using either diffusion tensor imaging (DTI) data, which is based on axons, or fMRI (Vogel et al. 2020). The predictions were then compared to cross-sectional tau PET scans. The DTI prediction provided the best fit, thus reinforcing the notion that tau pathology spreads through anatomical connections in living AD patients.

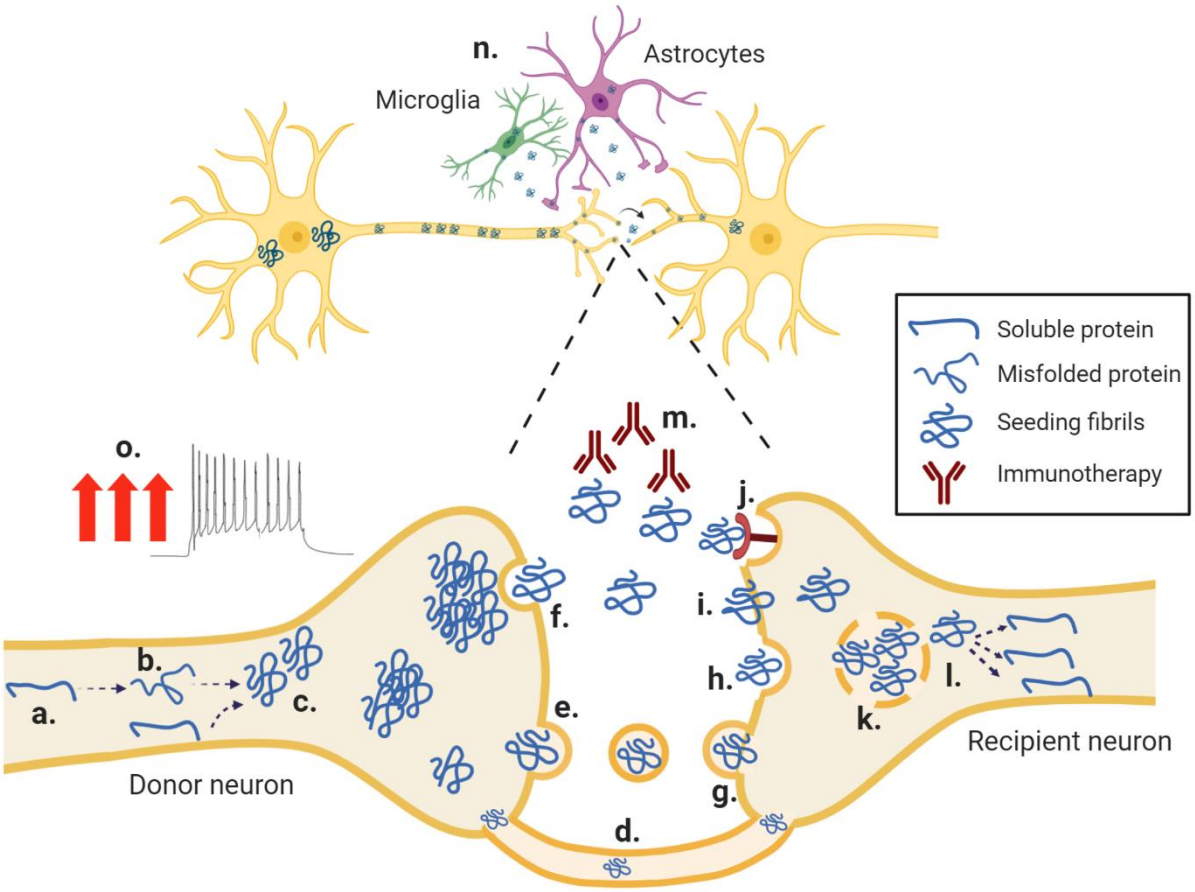

Figure 2. Hypothesis of pathological spread of toxic proteins and associated mechanisms.
According to this theory, some proteins (a) are prone to misfold upon certain stress cues, thus forming pathological misfolded proteins (b) that can template the misfolding of soluble monomers into highly reactive fibrils with seeding properties (c). Pathological proteins may be released through several membrane-dependent and -independent mechanisms, including tunneling nanotubules (d), synaptic vesicles and exosomes (e), as well as direct membrane translocation of naked proteins (f). Recipient neurons may actively incorporate pathological proteins through endocytosis (g), micropinocytosis (h), direct membrane translocation (i) or antibody-mediated endocytosis (j). After uptake, pathological proteins can disrupt endocytotic vesicles (k), thus gaining access to the cytoplasm of recipient neurons, where they can initiate new rounds of templated misfolding (l). Immunotherapies based on antibodies against pathological proteins (m) may be useful in preventing this phenomenon, thus slowing down disease progression. Glial cells (n) may also incorporate travelling pathological proteins and either contribute to their further dissemination or promote their elimination. Increased neuronal activity can boost the trans-cellular spread of toxic proteins (o). This cartoon was created with Biorender.

Overall, the current consensus is that the specific location of pathology in neurodegenerative diseases is influenced by both intrinsic vulnerability of specific brain areas, which dictates the spatial onset of pathology, and the ability of pathological proteins to be transmitted across connections. However, the great clinical heterogeneity of neurodegenerative disorders suggests that the spatiotemporal routes of pathology dissemination may be affected by various genetic and environmental factors, thus reducing the accuracy of prediction models in some cases (Peng, Trojanowski, and Lee 2020).

1.2.2 Seeding, release and uptake of pathological proteins

The spread of pathological proteins can be divided into various stages (Figure 2). First, the proteins need to become propagation-competent. Then, they have to be released from donor cells and subsequently be taken up by recipient cells. Finally, they enter the cytoplasm of recipient cells and cause damage.

While there is no common consensus as to the specific factors that trigger "propagation competence", it is clear that protein monomers first misfold and aggregate on their way to acquiring the ability to spread. These initial aggregates then initiate a seeding process in a prion-like manner by inducing the misfolding of monomers of their physiologically folded counterparts, which can subsequently template the misfolding of additional naïve molecules. This self-perpetuating process is known as template-directed misfolding or seeded nucleation and causes a rapid elongation of protein aggregates first into soluble and highly reactive oligomers, and subsequently into insoluble fibrils (Brettschneider et al. 2015) (Figure 2). However, this process may not always be required for spreading, as pathological tau was found to be transmitted even in the absence of endogenous tau (Wegmann et al. 2015). Overall, the jury is still out regarding the specific species that physically travel from cell to cell, with some *in vitro* reports indicating that misfolded trimers (Mirbaha et al. 2015) or even monomers (Mirbaha et al. 2018) could be the pathological moving seed. Furthermore, many of the neurodegeneration-associated proteins are subjected to various post-translational modifications (PTMs), including phosphorylation, acetylation and truncation. However, their role in the initiation of pathology initiation and dissemination *in vivo* remains largely unresolved (Katsinelos et al. 2018).

Various mechanisms have been associated with the release of neurodegeneration-associated proteins mostly in *in vitro* systems (Figure 2), including the secretion of these through synaptic vesicles (Pooler et al. 2013) or within exosomes (Asai et al. 2015; Emmanouilidou et al. 2010; Ngolab et al. 2017; Saman et al. 2012; Y. Wang et al. 2017), direct translocation across the plasma membrane mediated by sulfated proteoglycans and release as free protein (Katsinelos et al. 2018; Kfoury et al. 2012), and transport through tunneling nanotubes (Tardivel et al. 2016). In addition, several uptake mechanisms by recipient cells have been experimentally validated (Figure 2), such as bulk endocytosis (J. W. Wu et al. 2013), micropinocytosis by heparan sulfate proteoglycans (Holmes et al. 2013), dynamin- and clathrin-mediated endocytosis (Calafate et al. 2016; Evans et al. 2018) and even receptor-mediated endocytosis (Rauch et al. 2020). After the pathological seeds have entered the cell, they may damage and leak through internalized vesicles into the cytosol and initiate new rounds of templated misfolding (Calafate et al. 2016) (Figure 2).

Various additional factors may contribute to the spread of pathological proteins. For instance, microglia (Asai et al. 2015) and astroglia (Martini-Stoica et al. 2018) can take up pathological proteins and modulate their trans-neuronal spreading (Figure 2). Of note, some of the neurodegeneration-associated proteins are known to be physiologically secreted into the extracellular space (Avila 2010), which is exacerbated by synaptic activity (Pooler et al. 2013; Y. Wang et al. 2017; Yamada et al. 2014; Yamada and Iwatsubo 2018) (Figure 2). Optogenetically increasing the activity of neurons in the entorhinal cortex of mice expressing tau also accelerated its trans-synaptic spread (J. W. Wu et al. 2016).

Overall, while various pathways for release and uptake of pathological proteins have been identified in *in vitro* systems, further studies are needed to assess their contributions *in vivo*. In the current thesis, I have characterized this phenomenon in the context of familial ALS and FTD.

1.3. The *C9orf72* gene and neurodegenerative diseases

While genes carrying dominant mutations are mostly disease-specific (Table 1), some mutations in the same gene can cause neurodegenerative diseases with different clinical symptoms depending on additional factors that are in some cases unresolved. For instance, the same mutation in the gene encoding the prion protein (PrP) can cause either fatal familial

insomnia, characterized by compromised sleep, the demise of thalamic neurons and no PrP aggregates, or a familial form of Creutzfeld-Jakob disease, a cognitive disorder where cortical neurons typically degenerate and exhibit abundant PrP deposits. The nature of a polymorphic codon at position 129 determines which disease develops (Petersen et al. 1996). Similarly, a mutation in the chromosome 9 open reading frame 72 (*C9orf72*) gene can primarily cause both/either familial ALS (fALS) and/or familial FTD (fFTD).

1.3.1 Amyotrophic Lateral Sclerosis (ALS): pathobiology, epidemiology and therapeutics

1.3.1.1 Clinical presentations and selective vulnerability of somatic motor neurons

ALS, also known as motor neuron disease (MND) and Lou Gehrig's disease, was first described by the French neurologist Jean-Martin Charcot back in 1874 (Kumar et al. 2011). This is a prototypical neurodegenerative disorder characterized by the progressive dysfunction and demise of selective groups of motor neurons. Briefly, voluntary movements in human beings are controlled by the somatic motor system, which comprises two types of motor neurons: upper motor neurons (UMNs) and lower motor neurons (LMNs). The former, also known as corticospinal motor neurons, reside in the primary motor cortex and project their axons to LMNs, which can be located in the brainstem or in the spinal cord (SC). Brainstem LMNs are known as bulbar motor neurons (bMNs) and control the movements of face muscles, whereas SC motor neurons are known as somatic or spinal motor neurons (sMNs) and control the movements of limb muscles (Javed and Daly 2019). From a clinical perspective, ALS patients can initially manifest different symptoms based on the first motor neurons affected, and be stratified into four main subgroups: i) progressive muscular atrophy, characterized by the initial dysfunction of sMNs and weakness and atrophy of limbs; ii) primary lateral sclerosis, where UMNs are initially affected, thus causing hyperreflexia and spasticity, but little muscle atrophy; iii) bulbar ALS, where bMNs that innervate the tongue degenerate, thereby impairing tongue movements, swallowing and chewing, and causing speech difficulties; and iv) pseudobulbar palsy, where UMNs innervating bMNs degenerate, thus causing spasticity and reflexes of the tongue, and slurred speech (Taylor, Brown, and Cleveland 2016). Disease rarely starts simultaneously in multiple areas. However, while disease onset is focal, both UMNs and LMNs are eventually affected in an anatomically contiguous fashion, thus ultimately leading

to insidious disease spreading throughout the motor system and paralysis of almost all skeletal muscles. In fact, unlike primarily cognitive neurodegenerative disorders, ALS is strongly fatal, with death occurring 3-5 years after diagnosis due to respiratory failure (Taylor, Brown, and Cleveland 2016).

Of note, some sMNs are preserved throughout disease. On the one hand, these include neurons regulating eye movements (oculomotor neurons, trochlear neurons and neurons in the abducens nuclei), as well as those in the Onuf's nuclei in the lumbosacral spinal cord, which regulate sphincter and sexual functions. Consequently, eye movement, as well as sexual and bladder functions remain relatively preserved, even in the advanced stages of this disease (Kanning, Kaplan, and Henderson 2010). Moreover, within a given motor pool, certain functional subtypes of motor neuron are more vulnerable than others, with fast fatigable motor neurons being the most vulnerable (Kanning, Kaplan, and Henderson 2010).

While ALS has traditionally been considered a movement disorder, it is increasingly clear that it also boasts extramotor features, with up to 50% of patients showing cognitive and behavioural decline due to dysfunction of the frontotemporal and frontostriatal pathways. Furthermore, approximately 15% of patients initially diagnosed ALS develop FTD symptoms, thus constituting an additional phenotypic variant of ALS known as ALS-FTD (Al-Chalabi et al. 2016).

1.3.1.2 Environmental and genetic risk factors for ALS

ALS is the most predominant adult motor neuron disease with an estimated incidence of 1-7 per 100,000. Age of disease onset is variable but classically occurs in late midlife, around 55 years of age (Taylor, Brown, and Cleveland 2016). Ageing is considered as a major risk factor for ALS, e.g., its prevalence in the USA has been shown to increase steadily up to 80 years, from 5 in 100,000 in the 40-49 years period to 20 in 100,000 in the 70-79 window (Hou et al. 2019). Albeit difficult to study due to the relatively low prevalence of this disease, several environmental factors have also been associated with increased risk for ALS, such as smoking and strenuous physical activity (Oskarsson, Horton, and Mitsumoto 2015).

Up to 10% of ALS patients have at least one affected relative and are therefore diagnosed as familial ALS (fALS). The remaining predominant cases have sporadic ALS (sALS). Superoxide dismutase 1 (*SOD1*) was the first gene to ever be associated with fALS back in 1993 and is currently estimated to account for approximately 15 to 30% of all fALS cases in European and

Asian populations, respectively. Since its discovery, dominant mutations have been identified in more than 50 genes, including TAR DNA binding protein (*TARDBP*), fused to sarcoma (*FUS*) and optineurin (*OPTN*), the majority of which represent a very small proportion of fALS cases (Mejzini et al. 2019). However, in 2011 a mutation in the *C9orf72* gene was identified, which is estimated to account for over 30% of European fALS cases (Mejzini et al. 2019). Interestingly, a recent study showed that the prevalence of each disease subtype, including the presentation of cognitive deficits, is strongly affected by sex, and age of onset, as well as genetic factors (Chiò et al. 2020).

1.3.1.3 TDP43 pathology is the overarching hallmark lesion of ALS

Neuropathologically, ALS is characterized by the accumulation of cytoplasmic, ubiquitinated aggregates positive for the otherwise primarily nuclear protein TDP43 in the most vulnerable brain regions, including SC and primary motor cortex (I. R. A. Mackenzie et al. 2015; Neumann et al. 2006). These are observed in over 90% of all ALS cases, with patients carrying *SOD1* (I. R. A. Mackenzie et al. 2007) or *FUS* (Vance et al. 2009) mutations being the only cases lacking this neuropathological signature. TDP43 aggregates comprise C-terminally cleaved and hyperphosphorylated TDP43 molecules at various residues (Hasegawa et al. 2008; Neumann et al. 2006). Interestingly, the clinical progression of ALS may be explained by the sequential spreading of hyperphosphorylated TDP43 pathology across conserved neuronal circuits (Brettschneider et al. 2013) and there is, therefore, a great interest in understanding how ALS-associated genetic and environmental factors impinge on TDP43 biology.

TDP43 is ubiquitously expressed and is mostly located in the nucleus, where it binds mRNA molecules and regulates their splicing, translation, transport and degradation. TDP43 targets include genes involved in diverse cellular processes, such as neuronal plasticity and stress granule formation (Gao et al. 2018). Since pathological TDP43 aggregates in the cytoplasm, this leads to an overall reduction of nuclear TDP43, which results in a plethora of transcriptional changes that are thought to contribute to toxicity and constitute loss-of-function (LOF) consequences of TDP43 pathology (Humphrey et al. 2017). For instance, TDP43 deficiency is associated with retention of cryptic exons during splicing that result in the formation of non-functional truncated proteins, such as stathmin-2, whose function is essential for axonal regeneration of motor neurons (Melamed et al. 2019). Furthermore, TDP43 insoluble aggregates are per se toxic, as they can sequester numerous proteins, such

as components of the nuclear pore complex and the nucleocytoplasmic transport (NCT) machinery (Chou et al. 2018), these being considered as gain-of-function (GOF) properties of TDP43 proteopathy.

1.3.1.4 Approved drugs and novel therapeutic avenues under investigation for ALS

Currently, only two drugs have been approved for ALS treatment. Riluzole was approved in 1995 and has been reported to mildly extend survival, especially in bulbar-onset cases (Bensimon, Lacomblez, and Meininger 1994). Some studies have also suggested that riluzole can mildly delay disease progression, especially if administered at early disease stages and in late-onset cases (Bellingham 2011). While it was initially thought to act primarily via the inhibition of post-synaptic glutamate receptors, riluzole is currently considered to act by reducing overall neuronal firing and dampening down the release of excitatory neurotransmitters, like glutamate, from presynapses, thus reducing excitotoxicity (Bellingham 2011). Due to its relatively insignificant side effects and the failure of numerous clinical trials testing other compounds, riluzole remained the only approved drug for ALS for 22 years. In 2017, edaravone was approved for use in some countries and it is believed to function by activating the nuclear factor erythroid-derived 2-like 2 (Nrf2) antioxidant pathway (M. Zhang et al. 2019). However, similarly to riluzole, edaravone has relatively mild effects on reducing motor functional decline and its beneficial effects might be restricted to the first 6 months of treatment (Sawada 2017).

While efforts to develop efficient drugs have been scarcely fruitful, there is hope that alternative disease-modifying approaches can be developed. For instance, great preclinical efforts are being devoted to the development of antisense oligonucleotides (ASOs) that target mutated forms of ALS-relevant genes that act through gain-of-function mechanisms, such as *SOD1* (Miller et al. 2013) and *C9orf72* (Jiang et al. 2016). This strategy has already proved successful in treating other motor diseases, such as spinal muscular atrophy (Singh et al. 2017). Others are focusing on understanding the molecular differences that underpin the differential vulnerability to disease among motor neurons, which could be therapeutically exploited. For instance, large, vulnerable motor neurons express higher levels of matrix metalloproteinase-9 (Kaplan et al. 2014) and the ephrin receptor Epha4 (Van Hoecke et al. 2012). In fact, LOF mutations in the *EPHA4* gene are associated with longer survival in ALS patients (Van Hoecke et al. 2012). Similarly, synaptotagmin 13 is preferentially expressed in

oculomotor neurons, which do not degenerate in ALS, and increasing its expression levels in sMNs affords protection (Nizzardo et al. 2020).

1.3.2 Frontotemporal Dementia (FTD): pathobiology, epidemiology and therapeutics

1.3.2.1 Clinical variants and vulnerability of the frontotemporal lobes

FTD was first diagnosed by the Czech psychiatrist Arnold Pick back in 1892 in a patient that presented with aphasia, lobar atrophy and presenile dementia (Pick 1892). After several terms were coined, the clinical term "FTD" currently refers to a spectrum of syndromes that arise as a consequence of neurological damage in the frontal and temporal lobes. Patients can be stratified into three main clinical subgroups: i) behavioural variant FTD (bvFTD), ii) non-fluent variant primary progressive aphasia (PNFA) and iii) semantic dementia (SD), aka semantic variant primary progressive aphasia. First, bvFTD patients typically exhibit personality and behaviour changes manifested by impulsive behaviour, sudden opinion changes (e.g., alterations in food preferences and reversals in political or religious beliefs), apathy, loss of sympathy, lack of appropriate emotional responses and mental inflexibility. All of these symptoms are closely related to the frontal lobes and constitutive the most common form of FTD disease onset, which typically evolves during a very slow disease course. Second, PNFA is characterized by progressive deficits in speech production, grammar usage and comprehension of complex sentences, which primarily arise due to dysfunction of the left frontal lobe. Third, SD is rather a progressive disorder of semantic knowledge and naming, without speech fluency being affected, and it occurs following damage in the left temporal lobe (Graff-Radford and Woodruff 2007).

Over time, symptoms of the three possible clinical syndromes may converge as disease progresses from the temporal lobes to the frontal lobes. In addition, about 40% of FTD patients develop a variety of motor deficits, including parkinsonism and ALS. The former occurs in up to 20% of FTD patients, especially in bvFTD, and presents as either CBD and PSP. On the one hand, CBD is a complex syndrome characterized by asymmetric rigidity, sensory-motor cortical dysfunction, alien limb phenomena and dystonia. On the other hand, PSP is characterized by progressive parkinsonism-like movement abnormalities (e.g., early postural imbalance with falls), supranuclear gaze palsy (i.e., inability to move eyes voluntarily in a particular direction) and axial rigidity. Furthermore, up to 12.5% of FTD patients, mostly

exhibiting bvFTD-like symptoms, develop ALS during the course of disease (Bang, Spina, and Miller 2015).

Initial neuropathological changes in bvFTD take place in discrete cortical regions, including the anterior cingulate cortex (ACC) and the fronto-insular (FI) cortex. More specifically, von Economo neurons (VENs) in the ACC, as well as VENs and fork cells in the FI, specifically degenerate in bvFTD, unlike in other dementias, such as AD (Kim et al. 2012). These brain areas of the limbic system are considered as key nodes to integrate the neural representation of human feelings in consciousness and interoception, which might account for the lack of contextual sensitivity typical of bvFTD (Kim et al. 2012).

1.3.2.2 Environmental and genetic risk factors for FTD

While the diagnostic challenges posed by the very variable clinical presentations of FTD render the real prevalence of this group of disorders unclear, FTD is thought to be the third most common cause of dementia after AD and Lewy body dementia (LBD), representing 5-10% of all dementia cases. This entails an overall prevalence estimated at 15-22/100,000 individuals (Onyike and Diehl-Schmid 2013). In addition, FTD accounts for 3-26% of early-onset dementia cases, which makes FTD the second most common type of dementia in patients younger than 65 after AD (Graff-Radford and Woodruff 2007). Similarly to most neurodegenerative diseases, FTD risk increases with advancing age (Niccoli, Partridge, and Isaacs 2017) and the mean age at diagnosis is about 70 years (Nilsson et al. 2014). Death usually occurs about 6-11 years after symptom onset, which is often caused by pneumonia or other secondary infections (Bang, Spina, and Miller 2015). Environmental risk factors raising risk for FTD are poorly understood, but FTD patients tend to have a higher educational background than AD patients (Borroni et al. 2008). This is counterintuitive, as a higher educational background has been associated with the cognitive reserve hypothesis, which postulates that higher education allows people to better deal with brain damage (Roe et al. 2007). Traumatic brain injuries also increase the risk of FTD (Deutsch, Mendez, and Teng 2015).

FTD has a strong genetic component, with about 40% of cases showing a family history of dementia, yet only 10% of cases show a clear autosomal dominant inheritance. This discrepancy might be due to a contribution of unknown genes, environment or both (Panza et al. 2020). The tau-encoding gene *MAPT* was the first gene to be ever associated with familial FTD back in 1998 (Hutton et al. 1998). Later on, rarer mutations in other genes have been

associated with familial FTD, including progranulin (*GRN*), TANK-binding kinase 1 (*TBPK1*), *FUS* and *TARDBP*. However, similarly to ALS, mutations in the *C9orf72* were recently found to be the most common genetic cause of FTD (Panza et al. 2020).

1.3.2.3 Neuropathological signatures of FTD: tau, TDP43 and FUS

The complex clinical spectrum of FTD is known to reflect a number of molecular signatures that overall trigger frontotemporal lobar degeneration (FTLD). Consequently, patients can be classified into three main neuropathological subgroups, which are mutually exclusive, based on whether the major component of the protein deposits observed at postmortem is formed by tau (FTLD-tau), TDP43 (FTLD-TDP) or FUS (FTLD-FET) (Panza et al. 2020). FTLD-tau is present in 40% of all FTD cases and its clinical presentation is very heterogeneous, only being absent in ALS-FTD patients. FTLD-TDP represents 50% of all FTD patients and it can, in turn, be subclassified into four additional pathological groups based on the type and cortical laminar distribution of neuronal inclusions, which are often associated with mutations in specific genes (I. R. Mackenzie and Neumann 2017). The most common subtype is FTLD-TDP type B, which represents 35% of cases and is characterized by robust pathology across all cortical layers and predominantly diffuse, granular, neuronal, cytoplasmic inclusions. This subtype is often associated with *C9orf72* mutations and is absent in PSP (I. R. Mackenzie and Neumann 2017). FTLD-FET is much rarer, occurring in only 5-10% of FTD cases, and is typically associated with an unusual early-onset bvFTD clinical presentation with severe and rapidly progressive neuropsychiatric symptoms. Last, a few FTD cases exhibit brain aggregates positive for ubiquitin-only or p62-only, or no inclusions at all (Panza et al. 2020).

1.3.2.4 Approved drugs and novel therapeutic avenues under investigation for FTD

The only approved disease-modifying drug for the treatment of FTD is riluzole, but treatment is in general focused on the management of behavioral abnormalities, some of which can improve with drugs that inhibit serotonin reuptake, and antipsychotics (Panza et al. 2020). Drugs approved for treatment of AD have been tested for FTD, such as memantine, but these either provide no benefits or even cause severe side effects on cognition when given to FTD patients (Boxer et al. 2013).

Current efforts are mostly focused on interfering with tau or TDP43 aggregation, as FTLD-tau and FTLD-TDP43 are the most common neuropathological causes of this disorder. This includes the use of ASOs against mutation-carrying transcripts in familial FTD cases, modulators of the pathways involved in the degradation of protein aggregates, drugs that inhibit PTMs known to promote tau aggregation, as well as immunotherapies that aim to prevent the trans-cellular spread of TDP43 and tau (Panza et al. 2020).

Of note, while different FTLD subtypes are well-characterized, these cannot be accurately distinguished during clinical examination, which hurdles the correct stratification of patients into clinical trials testing disease-modifying drugs (Panza et al. 2020). The only exception is bvFTD developing to ALS, which is invariably associated with FTLD-TDP (Rohrer et al. 2011). Therefore, great ongoing endeavors also focus on identifying biomarkers that can differentiate between FTD subtypes, as well as FTD and other unrelated dementias, such as AD (Gossye, Van Broeckhoven, and Engelborghs 2019).

1.3.3 The C9orf72 gene

1.3.3.1 The discovery of the C9orf72 mutation in fALS and fFTD

As discussed above, ALS and FTD are strongly clinically connected, which had long suggested the existence of a prevalent common genetic cause. In 2006, a conserved haplotype of 11 Mb located at 9p13.2-9p21.3 was identified in a large Dutch family with autosomal dominant ALS-FTD by using linkage analysis as a key location of a major genetic risk factor for both ALS and FTD, which harboured 103 known genes (Vance et al. 2006). Then, the group of Bryan Traynor further investigated single nucleotide polymorphisms (SNPs) by using genome-wide genotyping arrays of Finnish patients with sALS, fALS and controls, which were chosen on the basis of the relatively high ALS incidence and genetic homogeneity of the Finnish population. This deep analysis revealed a robust, unannotated SNP cluster over the center of chromosome 9p, including three genes, *MOBKL2B*, *IFNK* and *C9orf72*, specific to sALS and fALS patients. After sequencing the coding regions, as well as exon-intron boundaries, for all these three genes, no major genetic suspects were identified (Laaksovirta et al. 2010). Finally, two independent laboratories headed by Bryan Traynor and Rosa Rademakers investigated the non-coding areas of these genes and observed that healthy individuals exhibited a GGGGCC (G4C2) hexanucleotide repeated from three to 23 times between non-coding exons 1a and 1b

of the *C9orf72* gene, whereas ALS, FTD and ALS-FTD patients had a G4C2 hexanucleotide repeat expansion (HRE) of about 700-1,600 repeats (DeJesus-Hernandez et al. 2011; Renton et al. 2011) (Figure 3). The prevalence of this expansion in their cohorts was estimated at 46%, 21% or 29% (Renton et al. 2011) and at 24%, 4% and 12% (DeJesus-Hernandez et al. 2011) in fALS, sALS and fFTD cases, respectively. It is currently known that patients typically carry several 100s-1000s repeats, and the epidemiological aspect of this mutation varies greatly worldwide, but it is accepted as the most common genetic cause of ALS, FTD and ALS-FTD (Cooper-Knock, Shaw, and Kirby 2014).

Given that many consider ALS and FTD as opposite sides of the same disease spectrum, the discovery of this mutation revolutionized the field by providing the most solid causal link as of yet between these diseases. Significant headway has been made in unravelling how this mutation causes toxicity in the hope to find therapeutic approaches that can be applied to both disorders.

1.3.3.2 Neuropathological and clinical hallmarks in C9orf72 mutation carriers

The *C9orf72* gene was not the first gene to ever be found to comprise an expansion of nucleotide repeats in association with a neurodegenerative disease. Among others, CAG repeats had been previously identified in coding DNA sequence in HD (Snell et al. 1993) and CGG repeats in non-coding DNA cause fragile X syndrome (Verkerk et al. 1991). Collectively, this group of diseases is known as microsatellite expansion or codon reiteration disorders, and currently includes more than 40 different disorders (Pattamatta, Cleary, and Ranum 2018).

The *C9orf72* gene is made up of 11 coding and non-coding exons and the transcription of its sense strand gives rise to three transcript variants (V1-V3) (Figure 3). For transcript variants 1 and 3, the G4C2 HRE lies within intron 1 and is therefore incorporated into their pre-mRNAs. However, in transcript V2 the G4C2 HRE is located in the promoter region and is therefore not included into this transcript variant. V1 and V3 pre-mRNA transcripts are subjected to alternative splicing, with V1 comprising only 4 coding exons and producing a short C9orf72 isoform (C9orf72-S), while V3 has 10 coding exons and forms a long C9orf72 protein isoform (C9orf72-L). V2 transcripts also generate the C9orf72-L isoform upon translation (Balendra and Isaacs 2018) (Figure 3). Albeit controversial, several studies have consistently found decreased levels of V2 mRNA in lymphoblast and frontal cortex (DeJesus-Hernandez et al. 2011; Peters et al. 2015) and lower C9orf72-L levels in specific disease-related (e.g., frontal cortex) and

disease-unrelated (e.g., occipital cortex) regions (Saberi et al. 2018; Xiao et al. 2015) in mutation carriers. Many have also reported epigenetic differences in the *C9orf72* gene in the presence of the mutation, including increased histone H3 and 4 trimethylation (Belzil et al. 2013) and promoter hypermethylation (Russ et al. 2015), which may collectively account for the *C9orf72* mRNA and protein expression differences between controls and *C9orf72* mutation carriers.

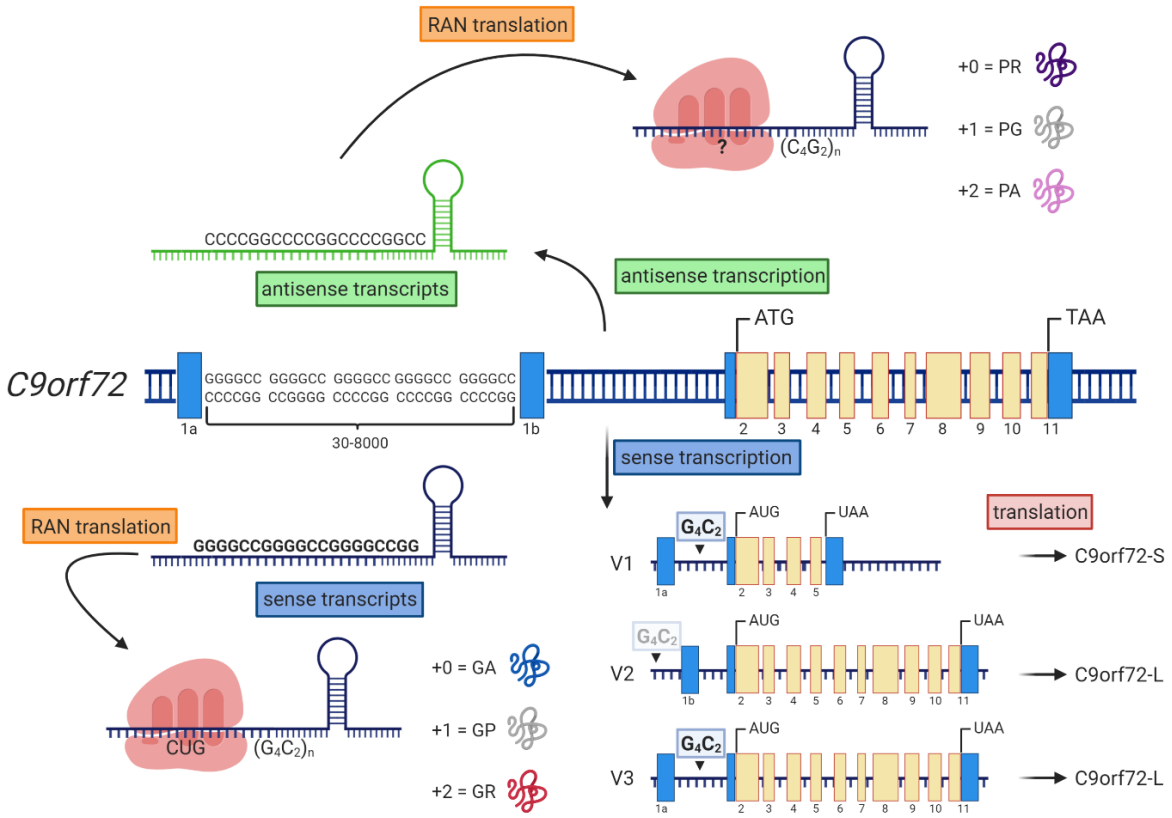

Figure 3. The C9orf72 mutation and its molecular products.
The G4C2 HRE is located in a non-coding part of the *C9orf72* gene and may contain from 30 to several thousand repeats. Upon transcription of the sense strand, 3 transcript variants are formed (V1-V3), of which V1 encodes C9orf72-S, while V2 and V3 encode C9orf72-L. The intronic HRE is present in the pre-mRNAs of transcripts V1 and V3 and is not degraded during splicing. Instead, it forms transcripts with G-quadruplex and hairpin structures, which can interact with various nuclear proteins. In the presence of the G4C2 HRE, the antisense strand is also transcribed, thus forming antisense transcripts that acquire similar structures as the sense ones. Both can undergo RAN translation, which allows for the production of 3 different DPRs from each reading frame of each strand. RAN translation of the sense strand starts at a CUG codon upstream of the HRE sequence. Whether RAN translation of the antisense strand also depends on non-canonical codons remains undetermined. This cartoon was created with Biorender.

Intronic sequences are typically spliced out and subsequently degraded. However, the G4C2 repeats in V1 and V3 transcripts are either retained within the mature mRNA transcripts or spliced out but not degraded (Figure 3). Consequently, the repeat-rich mRNAs can form RNA aggregates known as RNA foci in various brain regions and cell types, but these are primarily found in the nucleus of neurons (DeJesus-Hernandez et al. 2011). In-depth neuropathological studies revealed that, similarly to other repeat-expansion disorders, antisense transcription

also occurs from HRE-containing *C9orf72* loci (Figure 3), thus forming both sense and antisense RNA foci, which may co-exist in the same cells (Lagier-Tourenne et al. 2013; Mizielinska et al. 2013).

Finally, neuropathologists found that various tissues of *C9orf72* mutation carriers contained neuronal cytoplasmic inclusions that were negative for TDP43 but positive for p62 and ubiquitin. Recent additional studies showed that, in fact, the core of these inclusions are dipeptide repeat (DPR) proteins that are generated by repeat-associated non-AUG-initiated (RAN) translation of the sense and antisense HRE-containing transcripts (Figure 3). Many aspects of this non-canonical translation remain largely unknown, but it is known to affect transcripts forming hairpin structures, which is favoured by highly repetitive sequences and allows for translation to occur at various non-AUG codons, as well as at multiple sites within the repetitive sequence (Zu et al. 2011). RAN translation of the G4C2 HREs occurs in the three possible reading frames of the sense and antisense transcripts, thereby forming poly(glycine-alanine) (GA), poly(glycine-proline) (GP) and poly(glycine-arginine) (GR) DPRs from the sense strand, and poly(proline-alanine) (PA), poly (proline-arginine) (PR) and poly(proline-glycine) (PG) DPRs from the antisense strand (Arzberger et al. 2013; Gendron et al. 2013; Mori, Weng, et al. 2013; Zu et al. 2013) (Figure 3). Some of these DPRs co-exist in the same cells (Mori, Weng, et al. 2013), which might be due to the formation of chimeric DPRs from the same strand as a result of RAN errors (McEachin et al. 2020).

Apart from these three pathognomonic features, *C9orf72* mutation carriers show robust TDP43 pathology and neurodegeneration in ALS and FTD-relevant areas, respectively (MacKenzie et al. 2013; I. R. A. Mackenzie et al. 2015).

From a clinical perspective, FTD patients carrying this mutation are typically diagnosed familial bvFTD and exhibit a particular prominence of psychotic symptoms, especially complex paranoid delusions, and a low frequency of apathy. Strikingly, MRI studies have found atypical imaging features in bvFTD carriers, with apparent preservation of the frontotemporal lobes and greater atrophy in the precuneus in the parietal lobule (Devenney et al. 2014). In contrast, this mutation increases the likelihood of developing bulbar-onset ALS (Chiò et al. 2020) and one study found that *C9orf72* fALS cases tend to have an older age of onset and shorter disease duration than fALS cases associated with other mutations (Millecamps et al. 2012).

Finally, it is worth mentioning that the *C9orf72* mutation is not fully penetrant until the age of 80 and expansion carriers can exhibit clinical phenotypes other than FTD and ALS, including

PD, AD and LBD, yet this might be confounded by co-morbodities that appear in otherwise healthy carriers during ageing. It is not clear what disease modifiers determine the clinical presentation of mutation carriers (Cooper-Knock, Shaw, and Kirby 2014). Both environmental and genetic modifiers are under investigation, e.g., a silencing SNP in TMEM106B increases risk of FTD (Nicholson et al. 2013). As discussed below, one potentially critical aspect may be the repeat length number in specific brain areas.

1.3.3.3 Toxicity perpetrators of the C9orf72 mutation: experimental and human evidence

As discussed above, three main neuropathological features have so far been found to be specific to *C9orf72* mutation carriers: i) reduction in one or more *C9orf72* transcript and protein variants, ii) the production of repeat-containing sense and antisense transcripts that accumulate in the form RNA foci and iii) the accumulation of five different DPRs as cytoplasmic deposits derived from the non-canonical translation of the G4C2 HRE sequence. These findings suggest that both LOF and GOF mechanisms could play a role in disease and great efforts have been devoted to teasing out the relative contribution of each of them.

The contribution of the C9orf72 LOF has until recently gathered the smallest scientific support. Based on sequence homology comparisons, the C9orf72-L protein was first predicted to be structurally related to the DENN class of proteins, which regulate membrane fusion and budding events by modulating Rab GTPases (Levine et al. 2013). Most subsequent studies have focused on the long isoform of C9orf72 given its greater abundance and reliable detection (Viodé et al. 2018). C9orf72-L is a component of the autophagic machinery, but the specific steps where it is involved are still under investigation (Nassif, Woehlbier, and Manque 2017). Autophagy is a cellular process involved in the degradation of dysfunctional organelles and protein aggregates through their uptake into membranous structures called autophagosomes, which subsequently fuse with lysosomes for substrate elimination (Menzies et al. 2017). Cultured cells subjected to C9orf72 knock-down show lower autophagic flux, which indicates that C9orf72 is involved in promoting autophagy (Sellier et al. 2016; Webster et al. 2016). Mechanistically, the C9orf72 proteins interact with Rab1a and the ULK1 complex, and regulate the translocation of the ULK1 complex to the developing phagophore, which is essential for autophagy initiation (Webster et al. 2016). In addition, C9orf72-L forms a complex that acts as a GDP/GTP exchange factor for Rab8a and Rab39b, which are essential for autophagy initiation (Sellier et al. 2016). Others have shown that C9orf72-L and p62 interact

to eliminate stress granules via autophagy (Chitiprolu et al. 2018). Given that neurons are particularly vulnerable to autophagy defects (Hara et al. 2006), these findings suggest that C9orf72 LOF could strongly contribute to neurodegeneration in mutation carriers. However, none of the various *C9orf72* knock-out mouse models exhibit clear signs of neurodegeneration, motor phenotypes or cognitive impairment (Atanasio et al. 2016; Burberry et al. 2016; Jiang et al. 2016; O'Rourke et al. 2016). Nevertheless, these different mouse models did show increased systemic and brain inflammation markers, which may contribute to disease.

To decipher whether the HRE-containing transcripts and/or the DPRs cause toxicity, numerous studies have used *in vitro* and *in vivo* models with restricted expression of the G4C2 RNAs or the DPRs and measured toxicity. The fruit fly *Drosophila melanogaster* has been a central model in these investigations due to its particular suitability for genetic manipulations. A key pioneering study discriminated between RNA and protein toxicity by comparing the effects of RNA-only and DPR-only constructs in flies (Mizielinska et al. 2014). The former consisted of stretches of 12 G4C2 repeats interrupted by stop codons up to a maximal length of 288 repeats. These transcripts formed foci in neuroblastoma cells and in the fly salivary glands, and exhibited the expected G-quadruplex structure characteristic of GC-rich molecules. However, upon constitutive eye-specific or adult-onset pan-neuronal expression, these constructs failed to damage eye morphology, impair development or shorten lifespan, respectively, indicating that G4C2 transcripts alone cannot cause toxicity, even at very high expression levels. To test whether DPRs alone can cause toxicity, independent constructs encoding 36 or 100 GA, GR, PA and PR DPRs were generated by using alternate codons different from G4C2 that did not form foci. PA DPRs exerted no toxicity, while the arginine-rich DPRs strongly induced eye degeneration, impaired development and shortened lifespan. GA DPRs did not disrupt development or caused eye morphology defects, but they robustly reduced fly lifespan too. Many other studies using mammalian cells (May et al. 2014; Wen et al. 2014) and flies (Solomon et al. 2018; Tran et al. 2015; W. Xu and Xu 2018) have also shown that expression of single DPRs alone, particularly GR and PR, and to a lesser extent GA, is toxic. In contrast, expression of G4C2 RNA alone in mammalian cells (May et al. 2014) or in flies (Moens et al. 2018; Tran et al. 2015) is not. Regarding higher organisms, expression of sense and antisense repeat RNA in zebrafish models did cause toxicity independent of RAN DPR products (Swinnen et al. 2018), and HRE-containing RNAs have been found to bind to

numerous RNA-binding proteins (Cooper-Knock et al. 2014; Donnelly et al. 2013; Z. Xu et al. 2013), which could disrupt splicing and transport of numerous RNAs. Finally, several independent mouse models expressing one of the three toxicity-associated DPRs, namely GA, GR and PR, have been generated, all of which show cognitive and motor symptoms to varying degrees (LaClair et al. 2020; Schludi et al. 2017; Y. J. Zhang et al. 2016, 2018, 2019). Only one very recent study simultaneously compared GA and PR toxicity in congenic mice and found that only GA-expressing mice showed ALS-specific features, such as MN loss, muscle wasting, TDP43 pathology and a characteristic pro-inflammatory microglial signature (LaClair et al. 2020). Furthermore, behavioural and motor symptoms, as well as neuropathology, in mice expressing a bacterial artificial chromosome (BAC) that harboured the human *C9orf72* allele with 450 G4C2 repeats (BAC-450) were strikingly ameliorated upon treatment with GA-binding antibodies, which was not achieved by a GP-specific antibody (Nguyen et al. 2019).

Human studies have vaguely supported a more central role to GOF mechanisms so far. Reduction in the *C9orf72* transcripts and proteins has been reported in extracts from both disease-relevant and disease-unrelated brain areas (Saberi et al. 2018; Xiao et al. 2015). In addition, no LOF mutations in coding regions of *C9orf72* have been associated with ALS (Harms et al. 2013), and neither disease severity nor onset seem to change between homozygous and heterozygous carriers (Higginbottom et al. 2013). Regarding the GOF mechanisms, abundant RNA foci and DPR accumulation have been identified in the temporal lobe of a healthy carrier in the absence of TDP43 pathology or neurodegeneration, decades prior to the onset of her FTD symptoms (Vatsavayai et al. 2016), suggesting that these could be responsible for the prodromal phase (Edbauer and Haass 2016). In addition, quantitative neuropathological analysis revealed that GA DPRs, which current models correlate with mild toxicity, are the most abundantly detected DPRs and the arginine-rich DPRs, demonstrated to be the most toxic protein species experimentally, are relatively scarce (I. R. A. Mackenzie et al. 2015). Moreover, a poor correlation between DPR deposits and neurodegeneration at postmortem has been established, with DPRs being particularly abundant in the cerebellum and hippocampus and scarce in the SC and the frontal cortex (MacKenzie et al. 2013; I. R. A. Mackenzie et al. 2015). However, one recent study found that GR is the only DPR that predominantly accumulates in disease-relevant areas and that significantly co-localizes with phosphorylated TDP43 in ALS (Saberi et al. 2018). Also of note, two independent studies using a small cohort of C9-ALS patients found that, while RNA foci are not particularly abundant in

disease-related brain areas, antisense RNA foci were statistically more predominant in sMNs with cytoplasmic TDP43 accumulation than in those without. In contrast, sense foci were equally abundant in sMNs with or without TDP43 pathology, despite sense foci being more frequently detected than their antisense counterparts (Aladesuyi Arogundade et al. 2019; Cooper-Knock et al. 2015). These results do not rule out or strongly support the contribution of GOF or LOF mechanisms, as the great majority only used a very small number of postmortem tissues, which only allows for acute correlations and precludes cause-or-consequence extrapolations.

Overall, current experimental models support a major role for DPR toxicity, but no RNA-only studies have been conducted in mammals *in vivo* and RNA toxicity can therefore not be excluded. In addition, since the three neuropathological hallmarks of the *C9orf72* mutation co-exist, they may affect each other in complex ways that may even be tissue- and context-specific. In support of the latter, a recent study showed that reducing or ablating *C9orf72* in mice exacerbated their cognitive and motor deficits upon expression of BAC-450, which suggests synergism between GOF and LOF mechanisms (Zhu et al. 2020). Furthermore, *C9orf72* knock-out mice exhibit shorter lifespan in an environment-dependent manner (Burberry et al. 2020), thus highlighting that environmental factors may also determine the contribution of each of the *C9orf72* mutation hallmarks to disease.

1.3.3.4 The role of repeat length in C9orf72-associated toxicity: clinical and preclinical evidence

Since the discovery of the *C9orf72* mutation, it was of interest to test whether the HRE length could predict aspects of clinical relevance, such as age of onset, clinical phenotype (i.e., ALS vs FTD, and within these what clinical syndrome) and disease severity. One common difficulty faced by laboratories is the great variation in repeat sizes among ALS and FTD carriers, and even among tissues within the same individual, with overall mean repeat lengths being 1,667 in cerebellum, 2,717 in blood and roughly 5,000 repeats in various cortical areas, including the frontal cortex (van Blitterswijk et al. 2013). Somatic heterogeneity has been reported for other microsatellite disorders, such as myotonic dystrophy, with the degree of size heterogeneity increasing with advancing age, probably as a consequence of ageing-related DNA replication errors (Jones, Houlden, and Tabrizi 2017; Wong et al. 1995).

Regarding whether *C9orf72* HRE length correlates with age of onset, findings from cross-sectional studies are quite contradictory. In FTD mutation carriers longer repeat sizes in the frontal cortex correlated with older age of onset, suggesting that longer repeats are protective (van Blitterswijk et al. 2013). However, longer repeat sizes also correlated with older age at collection (van Blitterswijk et al. 2013), raising the possibility that patients with an older disease onset may have experienced more replication errors, thus showing longer repeats upon disease onset even if they initially had shorter repeats. In other words, repeat size in the frontal cortex may simply reflect the patient's age. Conversely, a different study using blood and brain-derived DNA found that carriers with fewer than 80 repeats had a significantly later age of onset (Gijselinck et al. 2016). However, the largest study so far found that longer repeats in blood resulted in later age of onset, but this association completely depended on the age at collection in patients and repeat numbers markedly changed throughout lifespan, making any conclusions unreliable (Fournier et al. 2019).

In other microsatellite diseases, such as myotonic dystrophy and HD, the repeat size expands across generations, which results in earlier disease onset in carriers' offspring (Duyao et al. 1993; Jaspert et al. 1995). This phenomenon is known as disease anticipation and this has also been explored in the *C9orf72* field. One study used 13 individual parent-child pairs and methylation as a proxy for repeat length. Intergenerational repeat amplification and earlier age of onset were reported, supporting disease anticipation (Gijselinck et al. 2016). An additional study using 36 multigenerational families found that, over four generations, the average age of onset transitioned from 62 to 49, with no effect on age at death or whether the mutation came from the fathers or the mothers. Repeats were not sized in this study (Van Mossevelde, Van Der Zee, et al. 2017). The largest study so far found signs of both expansions and contractions across generations, and even within the same individuals using blood DNA, thus concluding that blood HREs cannot be used to determine if anticipation occurs from these samples (Fournier et al. 2019). One major confounding factor in these studies are the age at collection because younger generations, especially those in a family with ALS or FTD history, tend to be more aware of their disease risk and therefore seek for medical diagnosis at an earlier age. This phenomenon is called ascertainment bias and accounts for any effect of blood HRE on predicting age of onset (Fournier et al. 2019). Overall, no clear consensus exists here either.

Whether repeat length affects the clinical presentation or severity also remains largely controversial. One study using several tissues found no association (van Blitterswijk et al. 2013), while a more recent one using blood samples suggested that longer repeats are associated with ALS rather than with FTD (Dols-Icardo et al. 2014). Of note, while the cerebellum is not considered a disease-relevant area for either ALS or FTD, longer repeats in the cerebellum correlated with shorter survival after symptom onset, with older age at collection not associated with longer repeats in this tissue (van Blitterswijk et al. 2013). These findings may be particularly interesting because the cerebellum is poorer in dividing glial cells and richer in non-dividing terminally differentiated neurons than other brain regions (Chong et al. 1995). Therefore, cerebellar repeat lengths may be less affected by replication errors, thus constituting an ideal organ to identify the "original" repeat length of an individual for correlation post-mortem research analyses (van Blitterswijk et al. 2013).

Regarding the evidence from experimental models, the majority of studies have used relatively short sequences of about 36 to 100 repeats because of technical difficulties while cloning long repetitive sequences. Nonetheless, expression of approximately 1,000 G4C2 repeats interrupted by stop codons did not cause toxicity in flies, suggesting that repeat length may not dictate RNA-mediated toxicity (Moens et al. 2018). In contrast, expression of each of the five DPRs with 36 to 400 impure repeats in cortical neurons showed that arginine-rich DPRs cause neuronal death in a partially repeat length-dependent manner (Wen et al. 2014). This has also been supported by *Drosophila* studies, where GR36 and PR36 caused less eye and developmental toxicity than GR100 and PR100, respectively (Mizielinska et al. 2014). Moreover, DPRs of different lengths are localized to different compartments, e.g., GR20 localizes to the nucleolus (Kwon et al. 2014), while GR80 is cytoplasmic (Yang et al. 2015). Therefore, DPR proteins of different lengths may preferentially interact with different sets of proteins, thus causing differential levels of toxicity. One report found dramatic differences when comparing the effects of varying lengths of impure GA/GR/PR/PA DPRs on proteostasis, nucleolar stress and electrophysiology using a human neuroblastoma cell line (Callister et al. 2016).

In conclusion, technical hurdles to accurately size large expansions, small sample sizes and the use of variable tissues of different ages have led to considerable conflicting results among *C9orf72* clinical studies, but current evidence suggests that repeat size does not correlate with age of onset or disease presentation in a straightforward manner. Large-scale and longitudinal

studies are ongoing to solve current discrepancies and to clarify the prognostic value of repeat size and other disease modifiers, as well as whether and how repeat sizes in peripheral tissues, especially blood and fibroblasts, vary over time and across generations. Given that RAN translation efficiency increases with longer repeats (Zu et al. 2011), there is a great interest in understanding how pure DPRs of longer lengths behave in experimental models, which could be more clinically relevant. This question has been addressed in this thesis.

1.3.3.5 Glycine-alanine (GA) DPRs: synthesis, molecular features and toxicity mechanisms

As previously discussed, neuropathological studies have confirmed that GA is the most abundantly detected DPR in postmortem tissue (Mori, Arzberger, et al. 2013; Mori, Weng, et al. 2013). Three main possibilities may account for this: i) GA RAN translation may be more efficient than that of the other DPRs, ii) GA DPRs may be more stable than the other DPRs and/or iii) GA DPR-expressing cells may undergo less toxicity because GA expression affords protection or because it is simply not toxic.

Whether the GA reading frame can be more efficiently RAN translated has been experimentally tested in cell culture and rabbit reticulocyte lysates. Authors used a construct that harbored G4C2 repeats followed by three possible tags downstream of the repeats in each reading frame and preceded by several nucleotides from the human *C9orf72* intron 1 with no AUG codons. The GA reading frame was substantially more efficient than the GP and GR frames, and this was dependent on a CUG codon positioned 24 nucleotides upstream of the repeats, which happens to be in frame with GA (Figure 3). The production of the other DPRs may be due to occasional skipping over the CUG codon or ribosomal frameshifting during translation (Green et al. 2017; Tabet et al. 2018).

Several studies have also examined the aggregation propensity of the DPRs using synthetic fibers and concluded that GA is the only DPR that rapidly aggregates, assembling into fibrils of amyloid nature (Brasseur et al. 2020; Chang et al. 2016; Flores et al. 2016). This is in line with several cell culture, fly and mouse studies, showing that GA forms puncta-shaped aggregates (May et al. 2014; Schludi et al. 2017; Yang et al. 2015). While all DPRs appear as cytoplasmic aggregates in patient tissue, many preclinical studies expressing GR, PR, PA and GP alone in experimental systems have found a rather diffuse cytoplasmic or nuclear pattern (Wen et al. 2014; Yang et al. 2015). This might be due to the specific constructs used in these studies, which typically consist of a short number of repeats attached to a large soluble tag, like GFP.

Alternatively, GR, PR, PA and GP may simply aggregate more slowly than GA, but the short-term nature of preclinical studies may prevent reaching that state. Moreover, GR, PR, PA and GP could also exist as soluble molecules in patient tissue, but these may not exist at postmortem or not be detected by currently available antibodies. Importantly, a recent study found that chimeric GA:GP DPRs can be detected in patient tissue, where GP does have a punctated appearance (McEachin et al. 2020). This is supported by neuropathological studies showing that almost all DPR inclusions contain GA (Mori, Weng, et al. 2013). Therefore, the intrinsic aggregation-prone nature of GA may facilitate its long-term stability and subsequent detection.

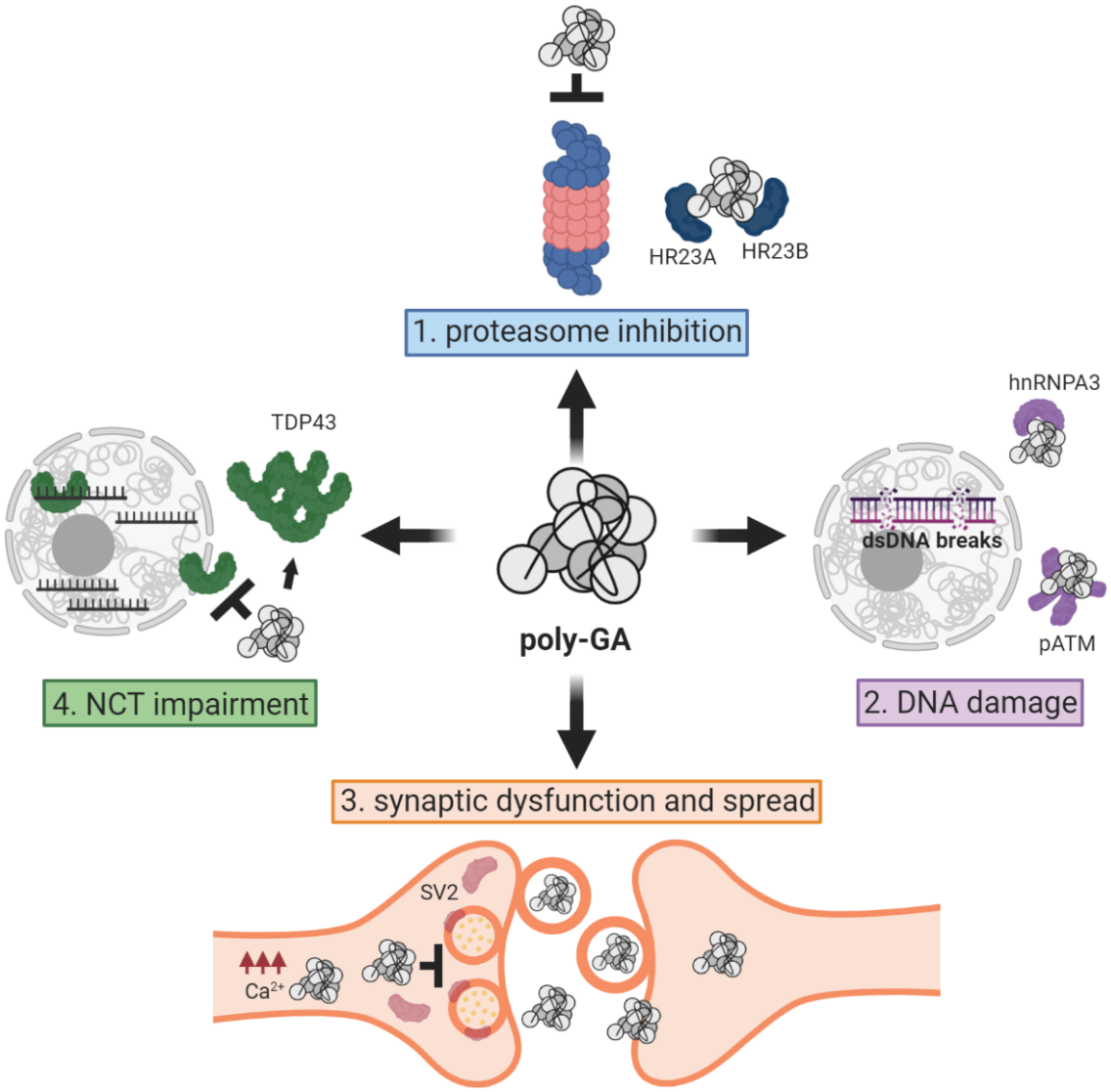

Figure 4. Proposed toxicity mechanisms mediated by GA DPRs.

(Below)

Figure 4. Proposed toxicity mechanisms mediated by GA DPRs.
GA DPRs rapidly aggregate into insoluble deposits and exert pleiotropic effects: 1) GA DPRs can directly interact and inhibit the proteasome, as well as bind to and sequester several proteasome-interacting proteins. These include HR23A and HR23B, which bind to ubiquitinated substrates and transfer them to nuclear proteasomes (Khosravi et al. 2020; Y. J. Zhang et al. 2016). 2) GA DPRs can directly bind to several proteins involved in DNA damage repair, such as hnRNPA3 and pATM dimers. This results in elevated double-stranded DNA (dsDNA) breaks upon GA expression (Nihei et al. 2020). 3) GA DPRs can travel along axons and inhibit synaptic vesicle release by reducing the protein levels of synaptic vesicle 2 (SV2). However, GA DPRs increase synaptic calcium influx (Jensen et al. 2020). Additionally, GA DPRs can be secreted and spread from cell to cell (Westergard et al. 2016). 4) Expression of GA DPRs leads to disruption of NCT by inhibiting the nuclear transport of importin-dependent substrates, such as TDP43, thus leading to TDP43 cytoplasmic retention and reduced nuclear TDP43 functions (Khosravi et al. 2017, 2020).

GA toxicity has been somewhat controversial due to conflicting data across studies using different experimental system expressing untagged or tagged GA with variable repeat lengths, and assessing different toxicity read-outs. The current consensus is that GA DPRs are mildly toxic and cause damage by disrupting various pathways, such as synaptic vesicle release (Jensen et al. 2020), DNA repair (Nihei et al. 2020; Y. J. Zhang et al. 2016), ER stress (Y. J. Zhang et al. 2014), NCT (Khosravi et al. 2017; Y. J. Zhang et al. 2016) and the proteasome (Khosravi et al. 2020; May et al. 2014; Y. J. Zhang et al. 2014) (Figure 4). Importantly, GA is the only DPR that has consistently been shown to cause cytoplasmatic accumulation of TDP43 (Khosravi et al. 2017) due to its ability to inhibit the proteasome, which causes TDP43 ubiquitination and this, in turn, prevents TDP43 from being transported into the nucleus (Khosravi et al. 2020) (Figure 4). Interestingly, GA is the only DPR that has consistently been found to spread from cell to cell in *in vitro* systems (Khosravi et al. 2020; Westergard et al. 2016; Zhou et al. 2017) (Figure 4), thus having the unique potential to cause damage through non cell autonomous mechanisms.

Overall, current evidence strongly indicates that GA DPRs are toxic and have the unique ability to recapitulate key aspects of ALS and FTD. However, most studies have been performed *in vitro,* and used tagged and short lengths of GA. Therefore, our current knowledge on GA-mediated neurotoxicity would greatly profit from *in vivo* studies using long and untagged GA constructs, which have a better potential to be clinically relevant. In this thesis I have addressed whether GA DPRs have a greater propensity to be trans-neuronally transmitted than the arginine-rich DPRs *in vivo* using the fly brain as a model*,* and explored the underlying mechanisms and non-cell-autonomous responses to this phenomenon. Furthermore, I have systematically tested whether GA toxicity is influenced by commonly used artificial tags and by repeat length.

1.4 Aims of the PhD Thesis

1. Characterizing the spread of the toxic *C9orf72*-derived DPRs in the adult fly brain
    - Unbiased comparison of the transmission potential of GA, GR and PR DPRs
    - Generation of novel fly lines expressing tagged and untagged GA200
    - Characterization of the contribution of repeat length and age to GA spread
    - Analysis of non-cell-autonomous effects of GA DPRs
    - Identification of modulators of GA spread

2. Deciphering the effect of commonly used tags on the pathobiology of GA DPRs
    - Generation of novel fly lines expressing tagged GA
    - Analysis of the toxicity inflicted by expression of tagged and untagged GA
    - Comparison of the aggregation propensity and cellular responses elicited by tagged and untagged GA

3. Test the role of repeat length in the pathobiology of GA DPRs
    - Generation of novel fly lines expressing GA400
    - Analysis of the toxicity inflicted by expression of GA100, GA200 and GA400
    - Comparison of the aggregation propensity, subcellular distribution and cellular responses elicited by GA100, GA200 and GA400

2. MATERIALS AND METHODS

2.1. Fly work

2.1.1 Maintenance of flies

Fly stocks were kept at 65% humidity on a 12:12 hour light:dark cycle and fed a standard sugar/yeast/agar (SYA) diet (Bass et al. 2007). For experiments using the constitutive eye-specific GMR-Gal4 and the inducible pan-neuronal elavGS drivers, experimental flies developed and were allowed to mate for two days at 25°C, after which female flies were sorted to SYA food at a density of 20 flies/vial and maintained at 25°C for the indicated number of days. In the case of elavGS flies, their food contained either 200 µM RU486 (Mifepristone) dissolved in EtOH or the same amount of EtOH-only. Flies used for propagation experiments expressed the temperature-sensitive Gal4 inhibitor Gal80^{ts} to minimize the expression of the UAS transgenes during development. This inhibitor is active at 18°C and can be inhibited to activate Gal4 activity by shifting flies to 29°C (McGuire et al. 2003). Therefore, flies used for propagation experiments developed and were allowed to mate for two days at 18°C, after which female flies were sorted into SYA food at a fly density of 20 flies/vial and maintained at 18°C or 29°C as indicated for each experiment.

2.1.2 Transgenic flies used in this study

The following transgenic fly lines were obtained from the Bloomington *Drosophila* Stock Center: GMR-Gal4, tubulin-Gal80^{ts} (BDSC_7019), orco-Gal4 (BDSC_23292), R9D03-Gal4 (BDSC_40726; hereafter referred to as OL-Gal4), UAS-eGFP.NLS (BDSC_4776), UAS-syt.eGFP (BDSC_6926), UAS-Rab11DN (BDSC_9762), UAS-Rab3CA (BDSC_9765), UAS-Rab3DN (BDSC_9766), UAS-ShiDN (BDSC_44222), UAS-*stip1* RNAi (BDSC_32979), UAS-Ork1Δc (BDSC_8928), UAS-eagDN (BDSC_8187) and UAS-NaChBac (BDSC_9469). The following stocks were obtained from the Vienna *Drosophila* Resource Center: UAS-*hsp90* RNAi (VDRC_108568), UAS-*syntx1* RNAi (VDRC_33112) and UAS-*comt* RNAi (VDRC_105552). The elavGS driver line was obtained as a generous gift from Dr. Hervé Tricoire (CNRS, France) (Osterwalder et al. 2001). Dilp3-Gal4 was obtained from Dr. Cathy Slack (Slack et al. 2011). Flies carrying the following transgenes inserted in the attP40 locus were obtained from Dr. Sebastian Grönke: UAS-GA36, UAS-GA100, UAS-GR36, UAS-GR100, UAS-PR36, UAS-PR100 (Mizielinska et al. 2014). The rest of the fly lines used were generated for this study.

2.1.3 Fly genetics

For all experiments, female Gal4 driver flies were crossed with UAS or wild-type (WT) male flies. To generate the final genotypes of the driver flies used for the propagation experiments, the orco-Gal4 and R9D03-Gal4 genes were recombined with UAS-syt.eGFP and UAS-eGFP.NLS on chromosome II, respectively. These flies were then crossed with tub-Gal80^{ts} flies and stable stocks were generated carrying the following genotypes: w^{-}; w, tub-Gal80^{ts}; w, orco-Gal4, w, UAS-syt.eGFP and w^{-}; w, tub-Gal80^{ts}; w, R9D03-Gal4, w, UAS-eGFP.NLS. For the genetic miniscreen, the UAS-GA200 (attP40) gene was recombined with tub-Gal80^{ts} on chromosome II. These were then crossed with orco-Gal4 flies and stable stocks were generated carrying the genotype: w-, w, UAS-GA200, w, tub-Gal80^{ts}; w, orco-Gal4.

2.1.4 Egg-to-adult viability assay and eye phenotypes

Five virgin GMR-Gal4 females were mated with five UAS males for two days. Then, flies were transferred to experimental vials and allowed to lay eggs for 5 hours at 25°C on SYA food. Eggs were counted, and then vials were incubated at 18°C or 25°C to achieve low or high transgene expression levels, as the Gal4-UAS system is temperature sensitive (Duffy, 2002). Adult flies were counted, and egg-to-adult viability was calculated by dividing the number of adult flies by the number of eggs. 10 replicates per genotype and temperature were used.
Eye images of 2-days-old flies expressing the indicated constructs under the constitutive eye-specific GMR-Gal4 driver were taken using a Leica M165 FC stereomicroscope equipped with a motorized stage and a multifocus tool (Leica application suite software).

2.1.5 Lifespan assay

Flies were reared at a standard density on SYA medium at 25°C and subsequently allowed to mate for two days after eclosion. Then, flies were sorted into experimental vials at a density of 15 flies per vial containing SYA medium with or without 200 µM RU486 to induce transgene expression. 10 independent biological replicates per condition were tested (i.e., n = 150 female flies per genotype and treatment). Flies were tipped to fresh vials 3-4 times per week and, at the same time, deaths in each vial were scored. Data are shown as survival curves and comparisons across genotypes and treatments were performed using a log-rank test.

2.2. Molecular biology

2.2.1 Generation of transgenic fly lines

2.2.1.1 Commonalities

The sequences of all primers used for this study are included in Table 2. List of primers used for this thesis.Table 2. In addition, a summary with all generated plasmids and fly lines can be found in Table 3 and Table 4, respectively. To achieve high expression levels, all constructs contained the CACC Kozak sequence before their ATG initiation site. Except for GA400 cloning, all other plasmids below were amplified using chemically competent OneShot TOP10 *E. coli* (Life Technologies) according to manufacturer's instructions for plasmid amplification. The sequence of all plasmids was verified by Sanger sequencing (Eurofins Genomics). Constructs were inserted into the fly genome using the phiC31 and attP/attB site-directed integration system (Bischof et al. 2007). All constructs were injected in house by Jacqueline Eβer with the exception of the UAS-GA400 plasmid, which was injected by BestGene Drosophila Embryo Injection Service. For comparisons across the different DPRs and for the genetic miniscreen (i.e., results section 1.3), the landing site attP40 was used, whereas for comparisons across the different repeat lengths or across tagged constructs of GA the landing site attP2 was used.

2.2.1.2 Generation of mCherry-tagged DPR36 and DPR100 constructs

To generate the mCherry-tagged DPR constructs, I first PCR amplified mCherry from a plasmid template provided by Christian Kukat (MPI AGE) and using the Phusion polymerase (NEB) and the primers JOL13 and JOL14, which allowed for the addition of an N-terminal NotI restriction site (RS) followed by the linker GGTAGTGGAAGTGGTAGT, as well as a C-terminal KpnI RS after the stop codon. This amplicon was then ligated into the pUAST attB *Drosophila* transgenesis vector to generate the pUAST-mCherry-C plasmid. In parallel, I PCR amplified the sequences for GA36, GR36, PR36, GA100, GR100 and PR100 (Mizielinska et al. 2014) using the TaKaRa LA Taq polymerase (Takara Bio Inc.) and the following primers: GA36fwd: JOL26; GA36rev: JOL33; GR36fwd: JOL26; GR36rev: JOL34; PR36fwd: JOL26; PR36rev: JOL35; GA100fwd: JOL26; GA100rev: JOL28; GR100fwd: JOL26; GR100rev: JOL30; PR100fwd: JOL26 and PR100rev: JOL32.

Table 2. List of primers used for this thesis.

Primer name	Primer sequence	Purpose
JOL9	AAAAGAATTCCAACATGGTGAGCAAG GGCGAG	Generation of the mCherry-only pUAST plasmids
JOL13	ATATGCGGCCGCCGGTAGTGGAAGTG GTAGTGTGAGCAAGGGCGAGGAG	Generation of pUAS T-mCherry-C
JOL14	CCCCGGTACCTCACTTGTACAGCTCGT CCATG	Generation of pUAS T-mCherry-C and mCherry-only pUAST plasmids
JOL26	ATATGAATTCGGATCCCACCATG	Generation of GA36mCherry, GR36mCherry, PR36mCherry, GA100mCherry, GR100mCherry, PR100mCherry, GA100GFP, GA200 and GA200-mCherry plasmids
JOL28	AAAAGCGGCCGCTGATGCTC	Generation of GA100-mCherry and GA200-mCherry plasmid
JOL30	AAAAGCGGCCGCTGAACGTC	Generation of GR100-mCherry plasmid
JOL32	AAAAGCGGCCGCTGATCGAG	Generation of PR100-mCherry plasmid
JOL33	AAGCGGCCGCTGAAGCG	Generation of GA36-mCherry plasmid
JOL34	AAGCGGCCGCTGATCTGC	Generation of GR36-mCherry plasmid
JOL35	AAGCGGCCGCTGATCTGG	Generation of PR36-mCherry plasmid
JOL36	AGCAACCAAGTAAATCAACTGCA	Amplification of UAS-transgenes from gDNA
JOL37	TGTCCAATTATGTCACACCACAG	Amplification of UAS-transgenes from gDNA
JOL43	GAATTCGGATCCCACCATGTCTAGAG GAGCT	Generation of the GA200 and GA200-mCherry plasmids
JOL44	CTTGCGGCCGCTTATGCTCC	Generation of the GA200 and GA200-mCherry plasmids
JOL69	CCGCGGCCGCTCTAGACCCGGGTGAT GCTCCTGCTCC	Generation of the GA200 and GA200-mCherry plasmids
JOL112	ACATCCCCGACTACTTGAAGC	qRT-PCR for mCherry
JOL113	ACCTTGTAGATGAACTCGCCG	qRT-PCR for mCherry
JOL117	AAAAGCGGCCGCTTACTTATCGTCGTC GTCCTTGTAATCTGCTCCTGCT	Generation of GA100FLAG plasmid
JOL124	ATATGCGGCCGCCGGTAGTGGAAGTG GTAGTATGGTGAGCAAGGGCGAGGA GCTGTTCAC	Generation of pUAS T-GFP-C
JOL125	AAAAGGTACCTCACTTGTACAGCTCGT CCATGCGGAGAGTGAT	Generation of pUAS T-GFP-C and GFP-only pUAST plasmids
JOL126	ATATGAATTCCAACATGGTGAGCAAG GGCGAGGAG	Generation of the GFP-only pUAST plasmids
SOL268	ATATGCTAAGCTGTCGCACAAATGG	qRT-PCR for Rpl32
SOL269	GATCCGTAACCGATGTTGGGCA	qRT-PCR for Rpl32

This resulted in the addition of an N-terminal EcoRI RS followed by the ATG initiation site, as well as a C-terminal NotI RS. These amplicons were first ligated into the pBlueScript SK(+) plasmid for amplification and subsequently subcloned into the pUAST-mCherry-C plasmid. As a control, I also PCR amplified mCherry using the primers JOL9 and JOL14, which resulted in the addition of an N-terminal EcoRI RS followed by the ATG initiation site, as well as a C-terminal NotI RS after the stop codon. This amplicon was then directly ligated into the pUAST attB plasmid.

Table 3. List of plasmids generated for this thesis.

Plasmid name	Vector backbone	Detailed description of restriction sites surrounding transgenes
pBS-GA100-NotI	pBlueScript SK(+)	EcoRI-Kozak-ATG-GA100-NotI
pBS-GA100-SmaI-XbaI	pBlueScript SK(+)	EcoRI-Kozak-ATG-GA100-SmaI-XbaI
pBS-GA200-NotI	pBlueScript SK(+)	EcoRI-Kozak-ATG-GA100-SmaI-XbaI-GA100-NotI
pBS-GA200-Stop	pBlueScript SK(+)	EcoRI-Kozak-ATG-GA100-SmaI-XbaI-GA100-Stop-NotI
pBS-GA36-NotI	pBlueScript SK(+)	EcoRI-Kozak-ATG-GA36-NotI
pBS-GA400-Stop	pBlueScript SK(+)	EcoRI-Kozak-ATG-GA100-SmaI-GA200-XbaI-GA100-Stop-NotI
pBS-GR100-NotI	pBlueScript SK(+)	EcoRI-Kozak-ATG-GR100-NotI
pBS-GR36-NotI	pBlueScript SK(+)	EcoRI-Kozak-ATG-GR36-NotI
pBS-PR100-NotI	pBlueScript SK(+)	EcoRI-Kozak-ATG-PR100-NotI
pBS-PR36-NotI	pBlueScript SK(+)	EcoRI-Kozak-ATG-PR36-NotI
pMK-RQ-SmaI-GA200-XbaI	pMK-RQ	SmaI-GA200-XbaI
pUAST-GA100-FLAG	pUAST attB	EcoRI-Kozak-ATG-GA100-FLAG-Stop-NotI
pUAST-GA100-GFP	pUAST attB	EcoRI-Kozak-ATG-GA100-NotI-linker-GFP-Stop-KpnI
pUAST-GA100-mCherry	pUAST attB	EcoRI-Kozak-ATG-GA100-NotI-linker-mCherry-Stop-KpnI
pUAST-GA200-mCherry	pUAST attB	EcoRI-Kozak-ATG-GA100-SmaI-XbaI-GA100-NotI-linker-mCherry-Stop-KpnI
pUAST-GA200-Stop	pUAST attB	EcoRI-Kozak-ATG-GA100-SmaI-XbaI-GA100-Stop-NotI
pUAST-GA36-mCherry	pUAST attB	EcoRI-Kozak-ATG-GA36-NotI-linker-mCherry-Stop-KpnI
pUAST-GA400-Stop	pUAST attB	EcoRI-Kozak-ATG-GA100-SmaI-GA200-XbaI-GA100-Stop-NotI
pUAST-GFP	pUAST attB	EcoRI-Kozak-ATG-GFP-Stop-NotI
pUAST-GFP-C	pUAST attB	NotI-linker-GFP-Stop-KpnI
pUAST-GR100-mCherry	pUAST attB	EcoRI-Kozak-ATG-GR100-NotI-linker-mCherry-Stop-KpnI
pUAST-GR36-mCherry	pUAST attB	EcoRI-Kozak-ATG-GR36-NotI-linker-mCherry-Stop-KpnI
pUAST-mCherry	pUAST attB	EcoRI-Kozak-ATG-mCherry-Stop-NotI
pUAST-mCherry-C	pUAST attB	NotI-linker-mCherry-Stop-KpnI
pUAST-PR100-mCherry	pUAST attB	EcoRI-Kozak-ATG-PR100-NotI-linker-mCherry-Stop-KpnI
pUAST-PR36-mCherry	pUAST attB	EcoRI-Kozak-ATG-PR36-NotI-linker-mCherry-Stop-KpnI

2.2.1.3 Generation of GFP- and FLAG-tagged GA100

To generate GFP-tagged GA100, I first PCR amplified GFP from a pre-existing GFP-containing plasmid (Carina Weigelt, unpublished) using primers JOL124 and JOL125 and Phusion polymerase. This resulted in the addition of a NotI RS and the same linker as above at the N-terminus of GFP, as well as a C-terminal KpnI RS after the stop codon. This was ligated into the pUAST attB vector to form the pUAST-GFP-C plasmid. I then digested the GA100 sequence out of the pBlueScript SK(+)-EcoRI-ATG-GA100-NotI plasmid, and ligated it into the pUAST-GFP-C

plasmid. As a control, I also PCR amplified GFP using primers JOL125 and JOL126, which allowed for the addition of an N-terminal EcoRI RS followed by an ATG initiation site, as well as a C-terminal NotI RS after the stop codon. To create GA100FLAG, I PCR amplified the sequence for GA100 using JOL26 and JOL117, the latter containing the FLAG-coding sequence followed by a stop codon and a NotI RS. No linker was included between GA100 and FLAG for this plasmid. This amplicon was then directly ligated into the pUAST attB plasmid.

Table 4. List of fly lines generated for this thesis.

Fly line	Plasmid used	Integration site	Figures where it was used
UAS-GA100 (III)	pUAST-GA100 (Mizielinska et al., 2014)	attP2	12, 15, 16, 17, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37
UAS-GA100FLAG (III)	pUAST-GA100-FLAG	attP2	24, 25, 26, 27, 28
UAS-GA100GFP (III)	pUAST-GA100-GFP	attP2	24, 25, 26, 27
UAS-GA100mCherry (II)	pUAST-GA100-mCherry	attP40	5, 8, 10, 11
UAS-GA100mCherry (III)	pUAST-GA100-mCherry	attP2	12, 13, 14, 24, 25, 26, 27, 28
UAS-GA200 (II)	pUAST-GA200-Stop	attP40	21, 23, 30, 34
UAS-GA200 (III)	pUAST-GA200-Stop	attP2	12, 15, 16, 17, 18, 19, 20, 22, 29, 30, 31, 32, 33, 34, 35, 36, 37
UAS-GA200mCherry (III)	pUAST-GA200-mCherry	attP2	12, 13, 14
UAS-GA36mCherry (II)	pUAST-GA36-mCherry	attP40	5, 8, 9
UAS-GA36mCherry (III)	pUAST-GA36-mCherry	attP2	13
UAS-GA400 (III)	pUAST-GA400-Stop	attP2	29, 30, 31, 32, 33, 34, 35, 36, 37
UAS-GFP (III)	pUAST-GFP	attP2	24, 25, 26
UAS-GR100mCherry (II)	pUAST-GR100-mCherry	attP40	6, 8, 10, 11
UAS-GR36mCherry (II)	pUAST-GR36-mCherry	attP40	6, 8, 9
UAS-mCherry (II)	pUAST-mCherry	attP40	5, 6, 7, 9, 10, 11
UAS-mCherry (III)	pUAST-mCherry	attP2	13, 14, 17, 21, 23, 24, 25, 26
UAS-PR100mCherry (II)	pUAST-PR100-mCherry	attP40	7, 8, 10
UAS-PR36mCherry (II)	pUAST-PR36-mCherry	attP40	7, 8, 9

2.2.1.4 Generation of untagged or mCherry-tagged GA200

To clone the GA200 and GA200-mCherry constructs, I PCR amplified the GA100 sequence (Mizielinska et al. 2014) in two independent reactions using the TaKaRa LA Taq polymerase and then ligated them together. First, I used primers JOL26 and JOL69 to add an N-terminal EcoRI RS followed by the ATG initiation site, as well as a SmaI RS and an XbaI RS at the C terminus. This amplicon was ligated into the pBlueScript SK(+) plasmid to obtain an EcoRI-ATG-GA100-SmaI-XbaI pBlueScript SK(+) plasmid. Second, I used JOL43 and JOL44 to add an N-terminal XbaI RS and a C-terminal stop codon followed by a NotI RS to GA100. Alternatively, I used JOL43 and JOL28 to add an N-terminal XbaI RS and a C-terminal NotI RS without a stop codon to GA100. The former was ligated into the EcoRI-ATG-GA100-SmaI-XbaI pBlueScript SK(+) plasmid to generate an EcoRI-ATG-GA100-SmaI-XbaI-GA100-Stop-NotI pBlueScript SK(+) plasmid (hereafter referred to as pBS-GA200-Stop), which was then subcloned into the pUAST attB plasmid (hereafter referred to as GA200). The latter was ligated into the EcoRI-ATG-GA100-SmaI-XbaI pBlueScript SK(+) plasmid to generate an EcoRI-ATG-GA100-SmaI-XbaI-GA100-NotI pBlueScript SK(+) plasmid, which was subcloned into the pUAST-mCherry-C plasmid generating the GA200-mCherry plasmid.

2.2.1.5 Generation of GA400

To clone the GA400 vector, a pMK-RQ plasmid containing a codon-optimized DNA sequence coding for 200 GA repeats, flanked 5’ by a SmaI RS and 3’ by a XbaI RS was synthesized by Eurofins Genomics (Germany). After amplification, this plasmid and the pBS-GA200-Stop vector were digested with SmaI and XbaI. The SmaI-GA200-XbaI sequence, as well as the linearized pBS-GA200-Stop vector were purified from 1% agarose gel with the QIAquick Gel Extraction Kit (Qiagen), and ligated. Ligation reactions were transformed to NEB Stable Competent *E. coli* (New England Biolabs) according to the manufacturer’s instruction. The sequence coding for GA400, consisting of both GA100 sequences, separated by the SmaI-GA200-XbaI sequence, was subcloned into the pUAST attB expression vector, transformed to and amplified in NEB Stable Competent *E. coli* and sequenced. The cloning of this vector was performed with the help of Tessa Supèr as part of her Master’s thesis in the lab.

2.2.2 Fly genotyping

To perform fly genotyping, gDNA was first extracted from 3 whole adult female flies using the DNeasy Blood & Tissue Kit (Qiagen). 10 ng of DNA per sample were then used to PCR amplify the UAS transgenes using primers JOL36 and JOL37, as well as TaKaRa LA Taq polymerase. PCR products were loaded in a 2% agarose gel. To further verify the identity of the amplified transgenes, I performed digestion check-ups. To this end, half of the amplified reactions was subsequently digested with SmaI or XbaI in Cut Smart 10X digestion buffer (NEB) for 30 min at 37°C, while the other half was only mixed with the digestion buffer in the same conditions without restriction enzymes. Digestion products were separated in a 2% agarose gel.

2.2.3 RNA extraction, cDNA synthesis and Quantitative Real-Time PCR (qRT-PCR)

Total RNA was extracted using Trizol (Invitrogen) according to the manufacturer's instructions. To further minimize DNA contamination, RNA pellets were treated with DNase I (ThermoFischer) as indicated by the manufacturer. RNA concentration was measured by the Qubit BR RNA assay (ThermoFisher). cDNA of mRNA was generated using 600 ng total RNA, the SuperScript III first-strand synthesis kit (Invitrogen) and random hexamers, according to the manufacturer's instructions.

For qRT-PCR of cDNA, PowerUp SYBR Green Master Mix (ThermoFisher) was used according to the manufacturer's manual. Primers JOL112 and JOL113 were used for mCherry cDNA. Rpl32 was used as a loading control, for which primers SOL268 and SOL269 were used. qRT-PCR was performed with a QuantStudio7 (ThermoFisher). Relative expression (fold induction) was calculated using the ΔΔCT method and Rpl32 as a normalization control.

2.3. Protein biochemistry

2.3.1 Soluble-insoluble protein fractionation and western blotting

To separate the soluble and insoluble fractions of our protein samples, I employed a slightly modified version of the Cragnaz et al., 2014 protocol (Cragnaz et al. 2014). Briefly, 20 adult fly heads were homogenized in 200 µl of ice-cold RIPA supplemented with Complete mini without EDTA protease inhibitor (Roche) and PhosStop phosphatase inhibitors (Roche), and incubated

under agitation for 1 h at 4 °C. Samples were then centrifuged at 1,000 g for 10 min at 4 °C, after which 25% of the sample was taken and used as input. Then, samples were centrifuged at 100,000 g for 45 min at 4 °C in an Optima XPN-100 ultracentrifuge (Beckman Coulter) and the supernatant was collected as the soluble fraction. The remaining pellet was re-extracted by incubation in 60 µl of urea buffer (8 M urea, 50 mM Tris–HCl, pH 8, 1% CHAPS, and protease and phosphatase inhibitors) for 20 minutes at room temperature (RT). Samples were then spun down to remove any precipitate and the 8 M urea soluble material was collected as the insoluble fraction. Protein samples were separated on any-kD stain-free Criterion gels (Biorad) after loading 5% of the input, soluble and insoluble fractions, and subsequently transferred to 0.45 µm nitrocellulose membranes (GE Healthcare).

For western blotting experiments, 20 adult fly heads were homogenized in 100 µl of ice-cold RIPA supplemented with protease and phosphatase inhibitors, and incubated on ice for 30 min with occasional vortexing. Samples were then centrifuged at 13,000 g for 10 min at 4 °C, after which the supernatant was retrieved. 15 µl per sample were separated on any-kD stain-free Criterion gels, and subsequently transferred to 0.45 µm nitrocellulose membranes.

After transfer, protein loading was imaged by exposing membranes to UV light. Membranes were subsequently blocked with 5% non-fat dry milk for 1 hour at RT and incubated over night at 4°C with the following antibodies: mouse anti-GA (clone 5E9) (1:1,000; Merck Millipore, AB_2728663), mouse anti-tubulin (1:10,000; Sigma-Aldrich, AB_477593), rabbit anti-K48-poly-ubiquitination (1:1,000; Cell Signalling, AB_10859893), rabbit anti-mCherry (1:1,000, Abcam, AB_2571870) and rabbit anti-Ref(2)P (1:1,000; Abcam, catalog #178440). Following three washes in TBST, membranes were probed with HRP-conjugated anti-mouse (1:10,000, ThermoFischer, AB_2536527) or anti-rabbit (1:10,000, ThermoFischer, AB_2536530) secondary antibodies for 1 h at RT and detection was performed using an ECL chemiluminescence kit (GE Healthcare). ImageJ was subsequently used for band intensity quantifications.

2.3.2 Dot blotting

Head protein extracts were isolated as described above for western blotting. 10 ul of each biological protein sample were blotted in duplicates on a 0.45 µm nitrocellulose membrane using the Bio-Dot Microfiltration Apparatus (Biorad) according to the manufacturer’s

instructions. Replicates from the same sample were loaded onto different positions of the membrane to control for efficiency differences across locations of the vacuum pump. Membranes were subsequently blocked, immunoblotted and developed as described for western blotting. The average of the mean GA signals in GFP-, mCherry-, or driver-only flies was calculated and used as background, which was subtracted from the mean of the two technical replicates of each GA-expressing biological sample.

2.3.3 In-gel proteasome activity assay and subsequent protein detection

Proteasome activity of the 20S and 26S proteasome complexes was measured as previously described (Vernace et al. 2007). Briefly, 5 heads of adult female flies were manually dissected in four dissection rounds using alternating genotypes, i.e., in each round the dissection order of each genotype was changed to apply the same waiting time to all genotypes until all samples were dissected. Then, heads were homogenized in 25 µl of proteasome buffer (50 mM Tris–HCl, pH 7.4, 5 mM MgCl2, 1 mM ATP, 1 mM DTT and 10% glycerol) on ice, centrifuged at 16,000 g for 10 min at 4°C. 15 µl of each sample were loaded and separated on a Bio-Rad TGX 7.5% precast native gel. After 3 hours of electrophoresis at 125V, gels were incubated in proteasome buffer containing 0.4 mM Z-Leu-LeuGlu-AMC (Enzo Life Sciences) for 15 min at 37°C. Proteasome bands were visualized with UV light on a ChemiDoc station. For quantification of 20S and 26S intensity values, these were normalized to the WT control of the same dissection round. To quantify assembly efficiency, the 26S band intensity of each sample was divided by its own 20S band intensity.

Protein content in native gels was subsequently transferred to 0.45 µm PVDF membranes (GE Healthcare), after which gels were incubated in pre-made Coomassie solution (Biorad) and membranes were blocked in 5% BSA, both for 1h at RT. Coomassie staining of non-transferred proteins was used as loading control. Blocked membranes were incubated over night at 4°C with mouse anti-Rpt6 (1:1,000; Santa Cruz, AB_1118465) and mouse anti-GA (clone 5E9) (1:1,000; Merck Millipore, AB_2728663) antibodies. Membranes were washed and developed as described for western blotting.

2.4. Histology

2.4.1 Immunostainings of adult Drosophila brains

Brains of adult female flies were dissected in PBS and immediately fixed in 4% paraformaldehyde at 4°C for 2 hours. Tissues were then washed 4-6 x 30 min in PBT (PBS with 0.5% Triton X-100) at RT. For experiments in which the mCherry and eGFP signals were imaged, brains were subsequently incubated in 50% glycerol in PBS for 1 hour at RT after washing and mounted in VectaShield Antifade Mounting Medium with DAPI (Vectorlabs). For experiments where GA, GR, PR or Ref(2)P were immunostained, brains were blocked in PBT with 5% fetal bovine serum and 0.01% sodium azide for 1 hour at RT after initial washing and incubated with mouse monoclonal anti-GA (1:3,000; Merck Millipore, AB_2728663), 5H9 rat anti-polyGR (1:50; (Mori, Arzberger, et al. 2013)), rabbit polyclonal anti-PR (1:1,000; Proteintech, catalog #23979-1-AP) and rabbit polyclonal anti-Ref(2)P (1:1,000; Abcam, catalog #178440) antibodies overnight at 4°C. Following 4-6 x 30 min washes in PBT at RT, brains were incubated overnight at 4°C at 1:1,000 dilution with one of the following secondary antibodies: Alexa488 goat anti mouse (ThermoFischer, catalog #A11001), Alexa568 goat anti mouse (ThermoFischer, catalog #A11031), Alexa633 goat anti mouse (ThermoFischer, catalog #21050), Alexa488 goat anti rabbit (ThermoFischer, catalog #A11008), Alexa568 goat anti rabbit (ThermoFischer, catalog #A11011), Alexa633 goat anti rabbit (ThermoFischer, catalog #A21071), Alexa488 goat anti rat (ThermoFischer, catalog #A11006) or Alexa647 goat anti rat (ThermoFischer, catalog #A21247). Finally, brains were washed 4-6 x 30 min in PBT, incubated in glycerol-PBS and mounted in VectaShield antifade mounting medium with DAPI (Vectorlabs, catalog #H-1200).

2.4.2 Stainings of adult Drosophila antennae

Antennae were manually dissected from EtOH-washed heads of adult flies and subsequently fixed in 4% PFA in a Petri dish for 40 min at RT. Then, they were washed in PBST 3 x 10 min and incubated in 40% glycerol 1 x 5 min and 80% glycerol 1 x 5 min. Finally, they were carefully mounted in VectaShield mounting medium.

2.4.3 Phalloidin stainings of adult Drosophila brains

To label cell membranes, whole brains were dissected, fixed, washed and incubated or not with primary and secondary antibodies as explained above. Following washing after incubation with the secondary antibody, tissues were incubated in a rhodamine-conjugated phalloidin solution (ThermoFischer, catalog #R415) diluted in PBT at 0.2 U/ml for 15 minutes at RT. Brains were subsequently washed 3 x 30 min in PBT, incubated in glycerol-PBS and mounted in VectaShield mounting medium. Degenerative vacuoles were manually identified in a blinded manner using z-stacks of a posterior side of the brain behind the calyx of the mushroom body (Coelho et al. 2018). Their area was quantified using ImageJ. The average number of vacuoles on each posterior side of the same brain was calculated and used for comparisons across biological replicates and genotypes.

2.4.4 TUNEL assay

To detect dsDNA breaks in adult fly brains, I used the In Situ Cell Death Detection Kit, Fluorescein (Merck, catalog #11684795910) following the manufacturer's instructions. Briefly, brains were dissected, fixed, washed and incubated with primary and secondary antibodies as explained above. Following washing after incubation with the secondary antibody, brains were incubated in a solution containing the enzyme and label for 1 hour at 37°C under agitation. Brains were then washed in PBT 2 x 30 min, incubated in glycerol-PBS and mounted in VectaShield mounting medium.

2.4.5 Imaging of adult Drosophila brains

Series of 2-µm z-stacks across the whole fly brain were taken for each image using a Leica SP8-DLS confocal microscope and the same settings were used across genotypes and ages, unless otherwise stated. In experiments where DPR propagation was investigated, brains were imaged with settings where propagated puncta were over-exposed, both in the case of the GA100 and the GA200 constructs, and where the signal in the negative control, devoid of any DPR construct, was minimal. This was done in an attempt to maximize the detectability of signal. To further maximize the detectability of specific signal, HyD detectors, gating and the excitation wavelength that maximized the fluorescence emission of all fluorophores were used in all cases during imaging.

2.4.6 Quantification of confocal images

All confocal images acquired for experiments where DPR spread was measured were first processed using ImageJ before subjecting them to quantification analysis. First, maximum z-stack projections were obtained to identify the lamina surrounding the optic lobes, as well as distinct artifacts, which were cropped from the stacks. In addition, areas of initial expression induction were also removed. For the latter, brain regions positive for eGFP were identified and cropped in experiments where eGFP was co-expressed along with the relevant DPR. This included the ORN axons and synaptic terminals, or the medulla of the optic lobes, as well as a distinct region in the antennal lobes, when the OL driver was used. Alternatively, a rectangle spanning the visually detectable antennal lobes and the rest of the lower part of the central brain was drawn in experiments where eGFP was not co-expressed and its content was also cropped to ensure that no puncta within the axons or the terminals of ORNs were included in the quantification of propagated puncta. Puncta in the remaining brain areas were quantified from the cropped z-stacks in 3D using the image analysis software Imaris 9.2.0 (Oxford Instruments). After background correction, the built-in spot detection algorithm was used to identify spots with a minimum size of 1500 nm. Detection settings were adjusted based on the maximum intensity of the spots, which proved the most accurate filter to distinguish between strongly labelled spots (considered as real GA puncta) and weak/low quality spots from trachea or background. The same parameters were used for all of the conditions compared in the same experiment. For the quantification of the mean punctum intensity and volume, the built-in surfaces detection algorithm of Imaris 9.2.0 was used. For the quantification of Ref(2)P puncta, the built-in spot detection algorithm was used to identify spots with a minimum size of 1000 nm.

For the quantification of TUNEL positive cells, the antennal lobes were or not cropped in flies expressing the transgenes of interest in ORNs or pan-neuronally, respectively. Then, TUNEL positive cells were quantified using the cropped z-stacks in 3D and the spot detection algorithm of Imaris with a minimum size of 3000 nm. Maximum intensity was also used to distinguish between specific and non-specific signal.

For the quantification of synaptic eGFP or GA levels in ORNS, I used their fluorescent signal in the ORN terminals in the antennal lobes as a proxy for their overall levels. Briefly, whole-brain stacks were taken with non-saturating settings for the ORN eGFP or GA signals, maximum intensity projections were generated from each z-stack, and the mean intensity of eGFP or GA

in the synaptic terminals of ORNs was measured using ImageJ. The same settings were used for all of the conditions compared in the same experiment. For the quantification of Ref(2)P intensity, the same approach was used, but the signal of the central part of the whole brain was quantified. The antennal lobes were or not cropped from the images in flies expressing the transgenes of interest only in ORNs or pan-neuronally, respectively.

2.5. Statistical analysis

Statistical analysis was performed using GraphPad Prism. Individual statistical tests are indicated in the figure legends. For multiple comparison testing, One-way and Two-way ANOVA were used. As recommended in (S. Lee and Lee 2018), when more than two pairs of groups were compared, the Tukey-Kramer test was used. Instead, the Bonferroni post-hoc test was applied when only two pairs of groups were compared. *P* values < 0.05 were considered significant: **P*<0.05, ***P*<0.01, ****P*<0.001 and *****P*<0.0001.

3. RESULTS

3.1 SPREAD OF GA DPRS IN THE FLY BRAIN

3.1 Spread of GA DPRs in the fly brain

An emerging theme in the field of neurodegenerative diseases is that specific toxic proteins can spread trans-cellularly, thus contributing to the clinical progression shown by patients. Several independent studies have reported transmission of the *C9orf72* DPRs in cell culture models (Khosravi et al. 2020; Westergard et al. 2016; Zhou et al. 2017). However, whether and how this phenomenon occurs *in vivo* remains unexplored.

3.1.1 GA DPRs spread rapidly in a repeat length- and age- dependent manner in the fly brain

Part of the results presented in this chapter were used to assemble the already published article "Morón-Oset et al., 2019". The published version of this paper, as well as a detailed explanation about the specific contributions made to it by Javier Morón Oset, can be found in the "publications" section at the back of this thesis.

3.1.1.1 Generation and validation of DPR-mCherry constructs

To address whether toxicity-associated DPRs can spread trans-neuronally *in vivo*, we generated novel fly lines that expressed C-terminally mCherry-tagged GA, GR or PR with 36 or 100 repeats (termed hereafter GA36, GA100, GR36, GR100, PR36 and PR100) from a UAS-transgene inserted at the same genomic locus (attP40 landing site), to ensure equal transcriptional levels. mCherry tagged DPRs were used in an effort to avoid problems with differences in sensitivity of the different DPR-specific antibodies. Given the large size of the mCherry tag, I placed a spacer sequence consisting of five glycine-serine (Gly-Ser) dipeptides between the GA, GR and PR sequences and the fluorophore (Figure 5A, Figure 6A, Figure 7A). Gly-Ser linkers provide flexibility and are therefore commonly used to allow the tagged proteins to maintain their natural folding (X. Chen, Zaro, and Shen 2013).

To validate our constructs, I generated flies that pan-neuronally expressed each of the mCherry-tagged DPRs, using the inducible elavGS system, and imaged the mCherry signal in adult fly brains after induction of DPR expression for 3 days.

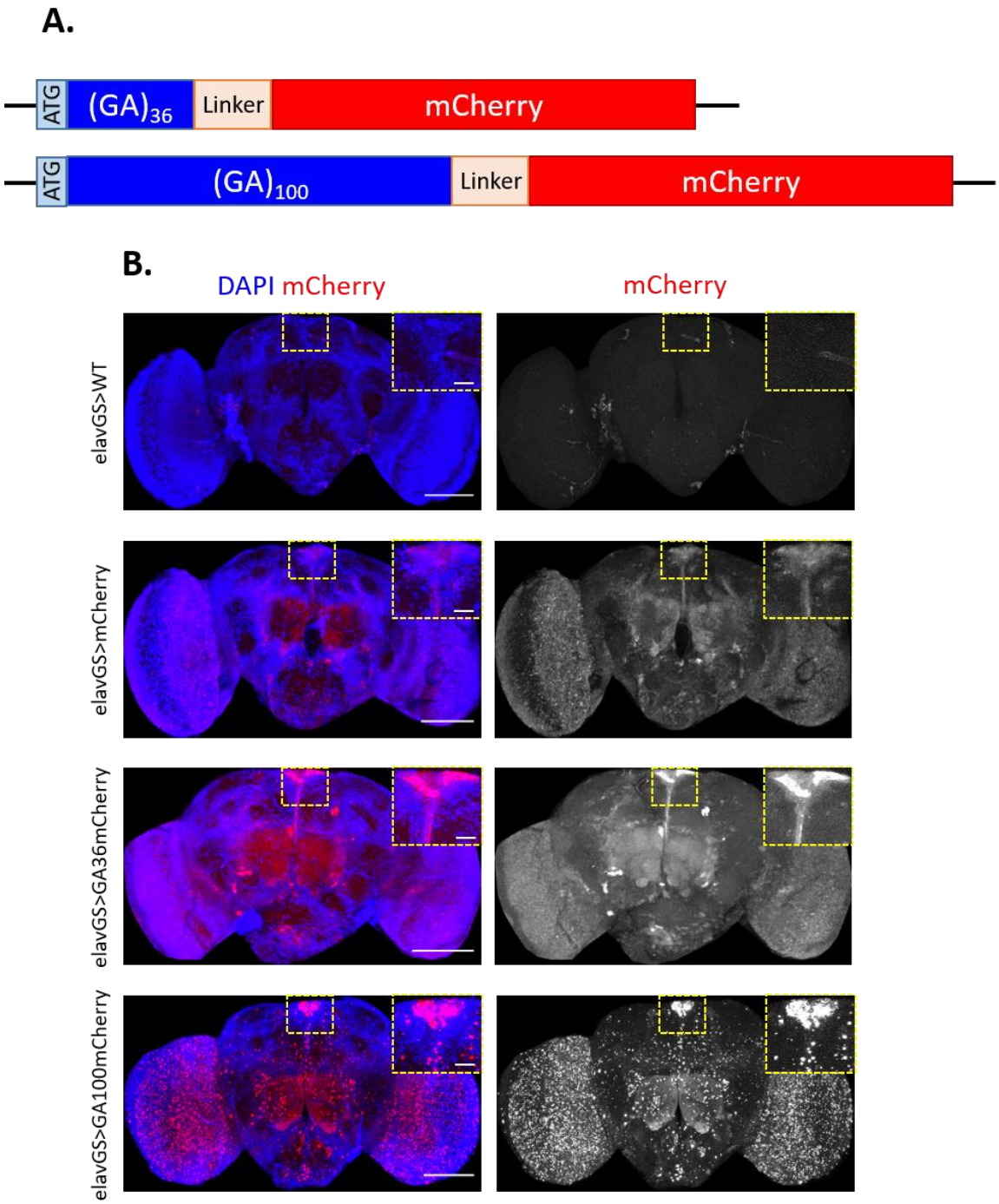

Figure 5. mCherry-tagged GA constructs can be detected by imaging their endogenous mCherry signal.
A Schematic of the GA36mCherry and GA100mCherry constructs generated and used in this study. **B** Representative images of 5-day-old adult fly brains from flies induced to express each of the indicated mCherry-tagged constructs for 3 days under the pan-neuronal driver elavGS. 10 times lower settings were used to image mCherry- and GA100mCherry-expressing brains, as the signal was much stronger in those genotypes. For WT and GA36mCherry, the settings were the same. No antibodies were used. Insets highlight the brain area where MNCs are located. Scale bars in images and insets are 100 µm and 10 µm, respectively.

Expression of both GA36mCherry and GA100mCherry fusion proteins could be detected by imaging their mCherry signals (Figure 5B). GA36mCherry exhibited a diffuse pattern, while GA100mCherry formed abundant puncta (Figure 5B), suggesting that the repeat length of GA

influences its propensity to aggregate. GR36mCherry expression was also detected through its mCherry signal (Figure 6B). However, the presence of GR100mCherry could only be verified by using a GR-specific antibody, which also stained GR36mCherry as well as untagged GR100 (Figure 6B).

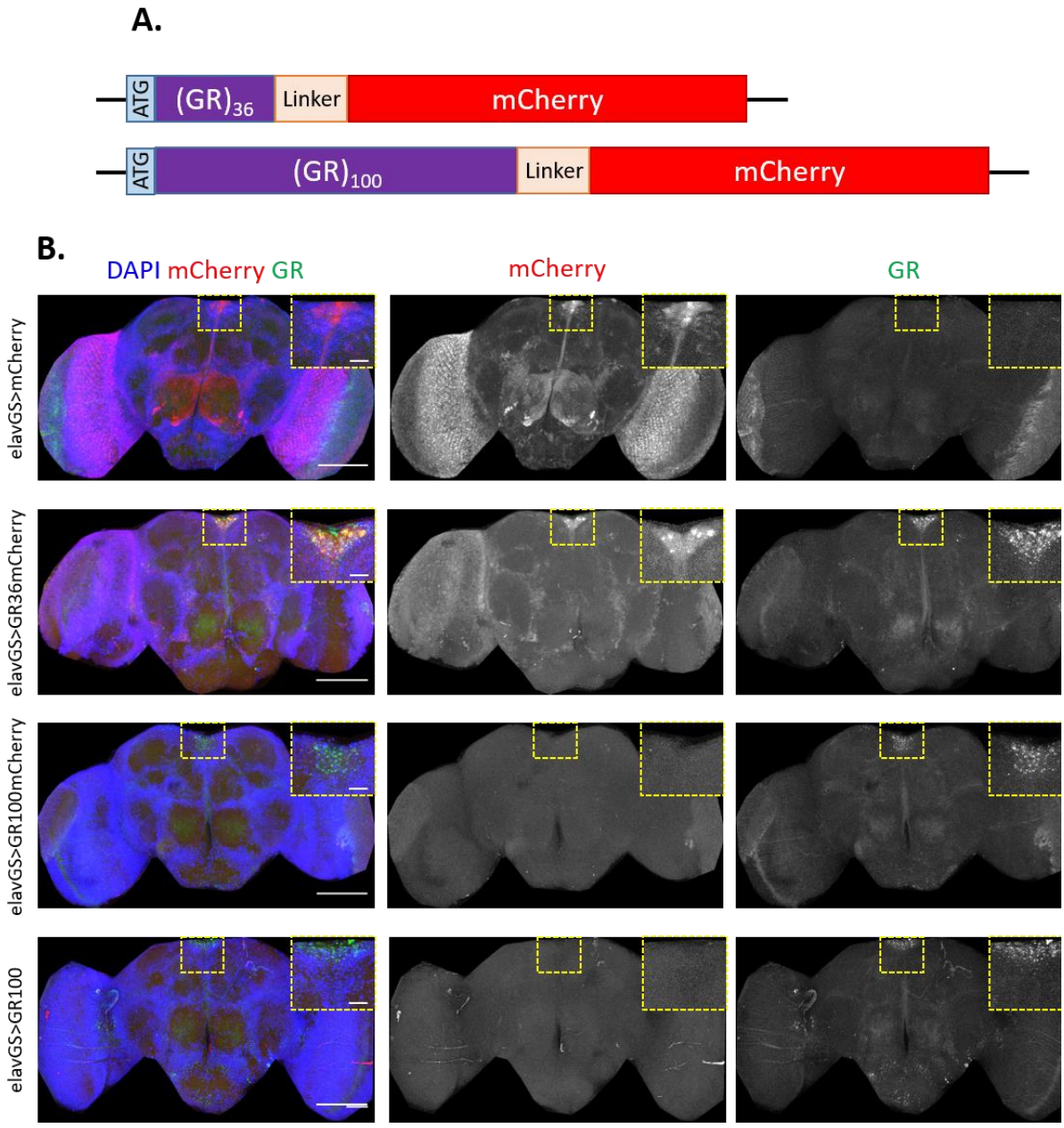

Figure 6. mCherry-tagged GR constructs can be detected upon pan-neuronal expression.
A Schematic of the GR36mCherry and GR100mCherry constructs generated and used in this study. **B** Representative images of 5-day-old adult fly brains from flies induced to express each of the indicated mCherry-tagged constructs for 3 days under the pan-neuronal driver elavGS. Brains were stained with an anti-GR antibody (green), and both the mCherry and the GR signals were imaged. 10 times lower settings were used to image the mCherry signal of mCherry-expressing brains than in the rest of the genotypes, as the mCherry signal was much stronger in mCherry-expressing brains. The same settings were used to image the GR signal in all of the genotypes. Insets highlight the brain area where MNCs are located. Scale bars in images and insets are 100 µm and 10 µm, respectively.

Expression of PR36mCherry and PR100mCherry was verified by imaging the mCherry signal, and by using a PR-specific antibody (Figure 7B). PR36mCherry was mostly nuclear, while both PR100mCherry and untagged PR100 were rather cytoplasmic (Figure 7B).

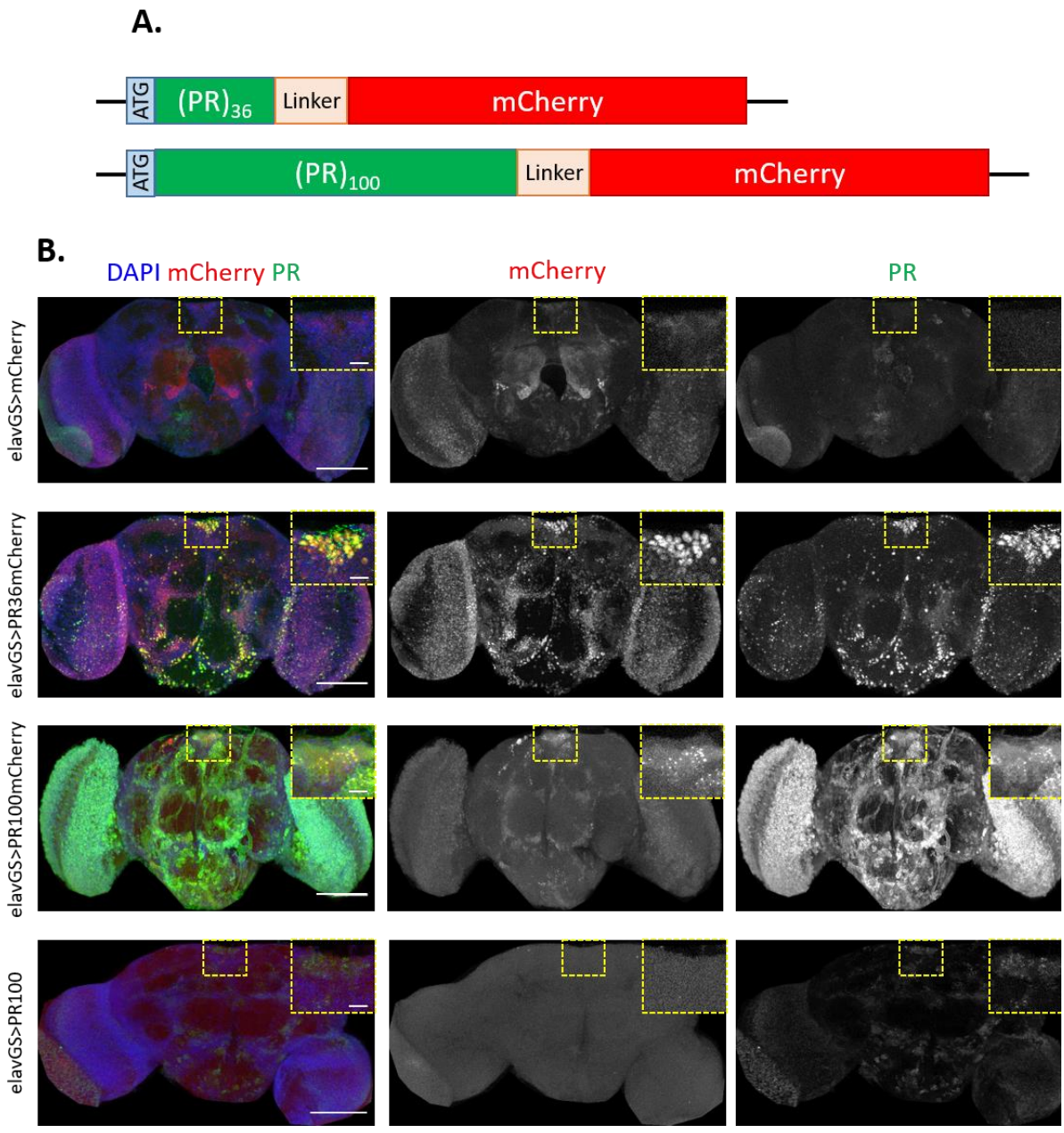

Figure 7. mCherry-tagged PR constructs can be detected upon pan-neuronal expression.
A Schematic of the PR36mCherry and PR100mCherry constructs generated and used in this study. **B** Representative images of 5-day-old adult fly brains from flies induced to express each of the indicated mCherry-tagged constructs for 3 days under the pan-neuronal driver elavGS. Brains were stained with an anti-PR antibody (green), and both the mCherry and the PR signals were imaged. 10 times lower settings were used to image the mCherry signal of mCherry- and PR36mCherry-expressing brains than in the rest of the genotypes, as the mCherry signal was much stronger in those genotypes. The same settings were used to image the PR signal in all of the genotypes. Insets highlight the brain area where MNCs are located. Scale bars in images and insets are 100 μm and 10 μm, respectively.

Interestingly, expression of GR100mCherry was almost exclusively detected in the median neurosecretory cells (MNCs) in the pars intercerebralis, where the expression levels of the

majority of the DPRs tested, but not mCherry only, was also particularly high (zoomed insets in Figure 5B, Figure 6B, Figure 7B). This suggests that MNCs may be particularly vulnerable to the accumulation of C9 DPRs. To further validate our constructs, I tested whether mCherry-tagged DPRs exerted similar toxicity to their untagged counterparts by generating flies with eye-specific expression of each construct using GMR-Gal4.

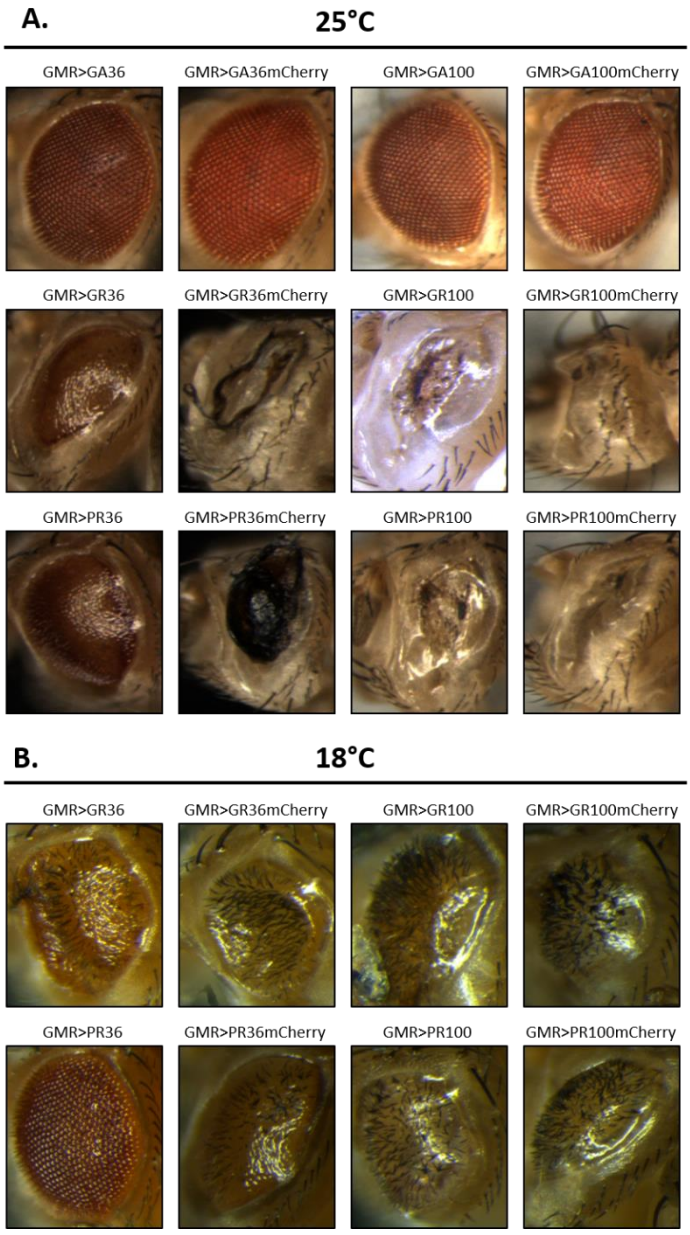

Figure 8. Eye-specific expression of mCherry-tagged, arginine-rich DPRs disrupts eye morphology.
A Representative eye images of 2-day-old flies expressing the indicated constructs under the eye-specific constitutive driver GMR-Gal4 after developing at 25°C. **B** Representative eye images of 2-day-old flies expressing the indicated constructs under GMR-Gal4 after developing at 18°C.

Expression of GR and PR in the fly eye at 25°C had previously been reported to cause severe eye degeneration, with 36 repeats being more toxic than 100. In contrast, neither GA36 nor

GA100 were reported to cause a rough eye phenotype (Mizielinska et al. 2014). In agreement, expression of GA36mCherry or GA100mCherry did not cause eye toxicity, whereas that of GR36mCherry, GR100mCherry, PR36mCherry or PR100mCherry triggered a rough eye phenotype (Figure 8A). GR36mCherry and PR36mCherry exerted greater toxicity than their untagged counterparts when expressed at 25°C (Figure 8A). To better discern toxicity differences between tagged and untagged DPRs, I also checked whether the arginine-rich DPRs would maintain their toxicity upon mCherry tagging when expressed at lower levels. To this end, I raised flies at 18°C. Given that the Gal4-UAS system is temperature-sensitive, lower transgene levels are expressed at lower temperatures (Duffy 2002). I found that GR36, GR100 and PR100 caused similar toxicity to their tagged counterparts at 18°C, whereas PR36 was more toxic upon tagging (Figure 8B).

Overall, these results show that our newly generated DPRmCherry constructs can be detected by immunostainings and they recapitulate the previously reported DPR-specific eye toxicity.

3.1.1.2 GA100 DPRs, but not GR100 or PR100 DPRs, spread rapidly in the fly brain

I next addressed whether the toxic DPRs have the ability to spread trans-neuronally *in vivo*. Given that a previous study reported the propagation of mutant huntingtin from Olfactory Receptor Neurons (ORNs) to other brain regions in *D. melanogaster* (Babcock and Ganetzky 2015), I also initiated expression in this brain area. I imaged the brains of flies where ORN-specific expression of GA36mCherry, GR36mCherry or PR36mCherry had been induced for 3 days in the adult fly using a temperature-inducible Gal80 and the ORN-specific or83b-Gal4 driver (Kreher, Kwon, and Carlson 2005; McGuire et al. 2003) (hereafter referred to as orco-Gal4). Since the cell bodies of ORNs are outside of the central brain and, therefore, only the axonal projections and synaptic terminals of ORNs can be detected in the adult central brain of *Drosophila* after dissection, I co-expressed eGFP-tagged synaptotagmin to label ORNs and control for driver specificity (Y. Q. Zhang, Rodesch, and Broadie 2002). No specific mCherry signal was found outside of ORNs (Figure 9A), suggesting that the short isoforms of the toxic DPRs cannot spread trans-neuronally, at least after short-term expression from this neuronal population.

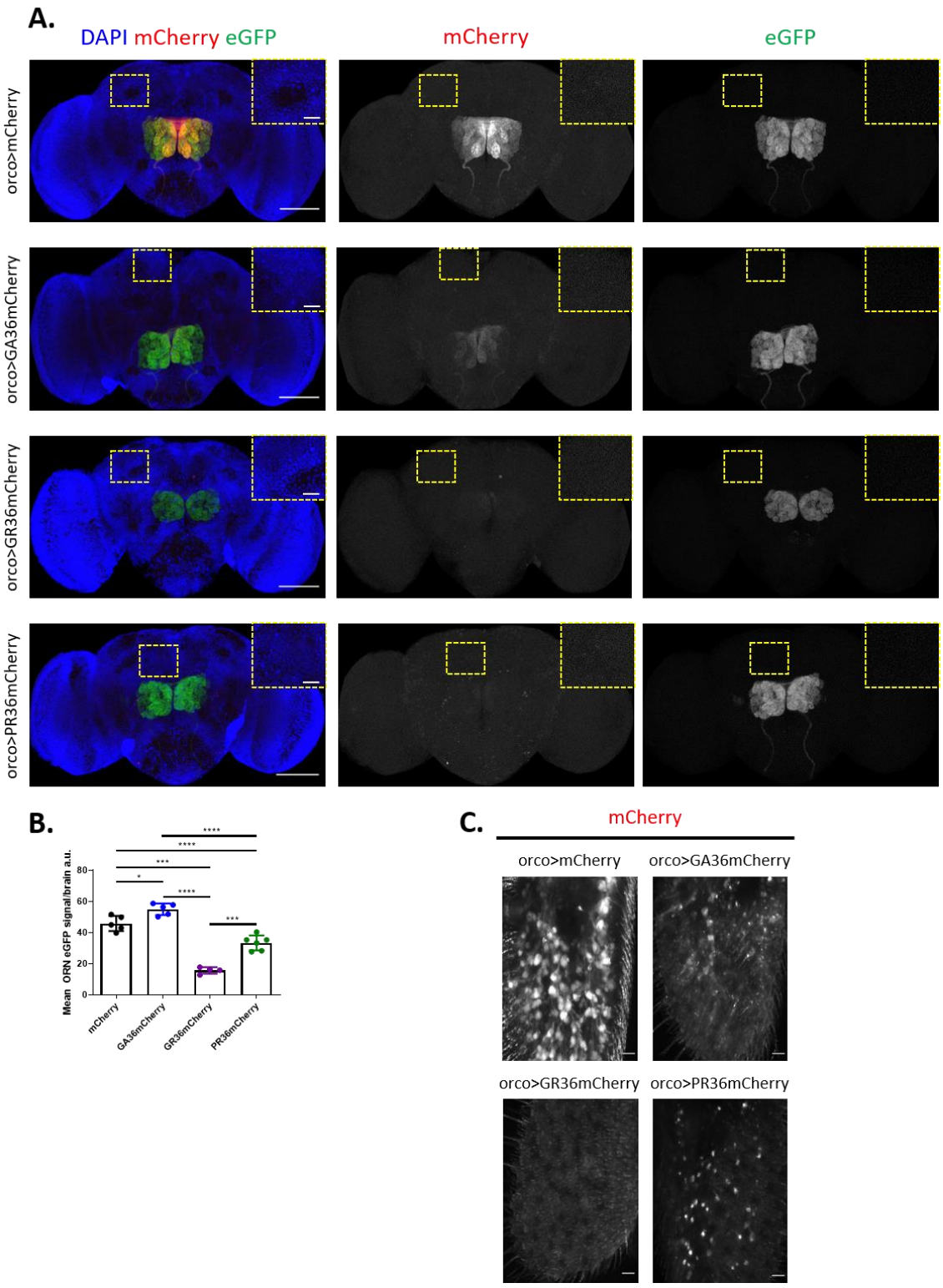

Figure 9. GA36mCherry, GR36mCherry and PR36mCherry do not spread from ORNs.
A Representative images of 5-day-old fly brains expressing GA36mCherry, GR36mCherry or PR36mCherry in ORNs for 3 days under the ORN-specific orco-Gal4 driver. Synaptotagmin-eGFP was co-expressed in all genotypes to identify ORNs. Flies expressing mCherry were used as a negative control to ensure that mCherry cannot spread by itself. No antibodies were used. Insets of the indicated areas are also shown to facilitate visualization. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B** Quantification of the eGFP signal detected within ORNs per brain (****P < 0.0001, ***P < 0.001 and *P < 0.05; One-way ANOVA with Tukey's multiple comparisons test, n = 4–6 brains). **C** Representative images of 5-days-old fly antennae induced to express the indicated constructs for 3 days. The same intensity settings were used to image the mCherry signal. Scale bars are 10 µm.

Interestingly, when expression of the longer DPRmCherry fusion proteins was induced in ORNs for 3 days, mCherry-positive puncta outside of the ORNs were exclusively detected in GA100-mCherry-expressing fly brains (Figure 10A-C), but not in brains expressing GR100mCherry or PR100mCherry, suggesting that longer GA DPRs may be more prone to spread trans-neuronally. Moreover, I detected a strong reduction in the eGFP fluorescence within the ORNs of flies expressing GR36mCherry, PR36mCherry, GR100-mCherry and PR100mCherry (Figure 9B, Figure 10D). This may result from the well-known inhibitory effect of the arginine-rich DPRs on protein translation (Kanekura et al. 2016; Moens et al. 2019). In contrast, both GA36mCherry and GA100mCherry mildly increased the eGFP signal in ORN terminals (Figure 9B, Figure 10D), which may be due to the ability of GA to inhibit the proteasome (Khosravi et al. 2020), thus interfering with synaptotagmin-eGFP degradation.

Since only GA36mCherry and GA100mCherry were detected in the synaptic terminals of ORNs, we speculated that GR36mCherry, PR36mCherry, GR100mCherry and PR100mCherry may not be transported along the axons and therefore may reside in the soma of these neurons, which are not visible upon dissection of the adult fly brain. Therefore, I dissected the antennae of flies of the same genotypes and age, and imaged their mCherry signal. I detected GA36mCherry, PR36mCherry, GA100mCherry and PR100mCherry in the cell bodies of antennal ORNs. However, neither GR36mCherry nor GR100mCherry were detected (Figure 9C, Figure 10E).

To confirm that GR did not spread and rule out a detection problem of this DPR in our ORN paradigm, I generated flies expressing GA100mCherry or GR100mCherry only in MNCs using dilp3-Gal4 (Buch et al. 2008). I targeted this neuronal subset because I had previously detected GR100mCherry almost exclusively in the MNCs upon pan-neuronal expression, which would therefore allow us to rule out a detectability problem. In agreement with our ORN paradigm, expression of GA100mCherry in MNCs led to the accumulation of mCherry-positive puncta both inside the axons of MNCs and outside of these neurons (Figure 11A), indicating that GA100mCherry can also spread from this neuronal subset. In contrast, GR100mCherry accumulated in the cell bodies of MNCs, but it was not detected in the axons of these neurons and it did not spread (Figure 11B).

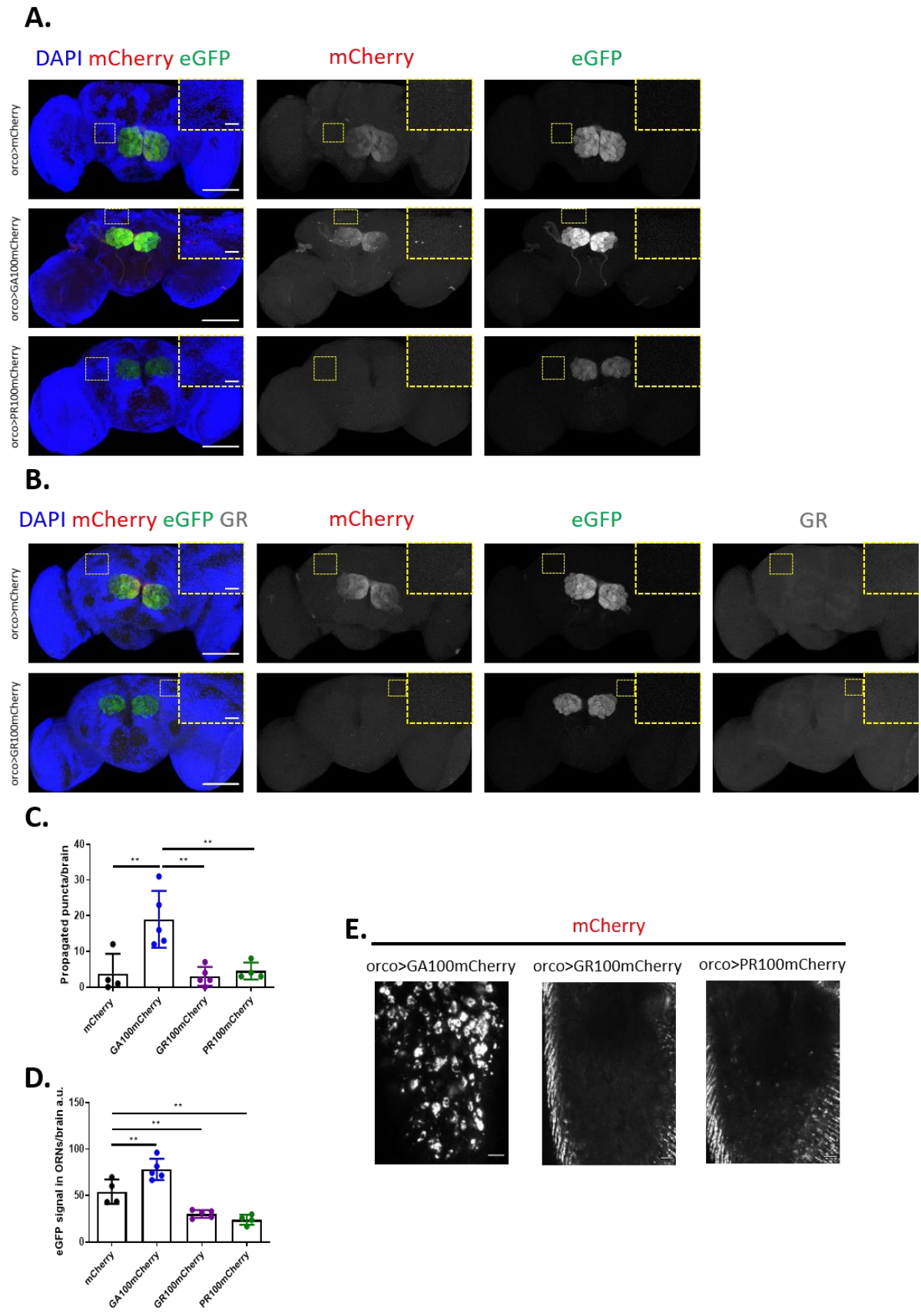

Figure 10. GA100mCherry, but not GR100mCherry or PR100mCherry spread outside of ORNs.

(Below)

Figure 10. GA100mCherry, but not GR100mCherry or PR100mCherry spread outside of ORNs.
A-B Representative images of 5-day-old fly brains expressing mCherry, GA100mCherry, PR100-mCherry or GR100-mCherry in ORNs for 3 days under the ORN-specific orco-Gal4 driver. The same settings were applied to all genotypes while imaging their eGFP and mCherry signals. Spreading was only observed in flies expressing GA100mCherry **(A)**. EGFP and mCherry were detected using fluorescence as read-out. A GR-specific antibody was used to detect GR100mCherry and mCherry **(B)**. Fly brains expressing only mCherry were used as controls to show that mCherry cannot spread by itself **(A)** and to verify antibody specificity **(B)**. Insets of the highlighted brain regions are shown. Scale bars in images and insets are 100 µm and 10 µm, respectively. **C** Quantification of the number of mCherry puncta detected outside of ORNs across genotypes per brain after induction for 3 days. **D** Quantification of the eGFP signal detected within ORNs per brain (**P < 0.01; One-way ANOVA with Tukey's multiple comparisons test, n = 4–6 brains). **E** Representative images of 5-day-old fly antennae induced to express the indicated constructs for 3 days. The same intensity settings were used to image the mCherry signal. Scale bars are 10 µm.

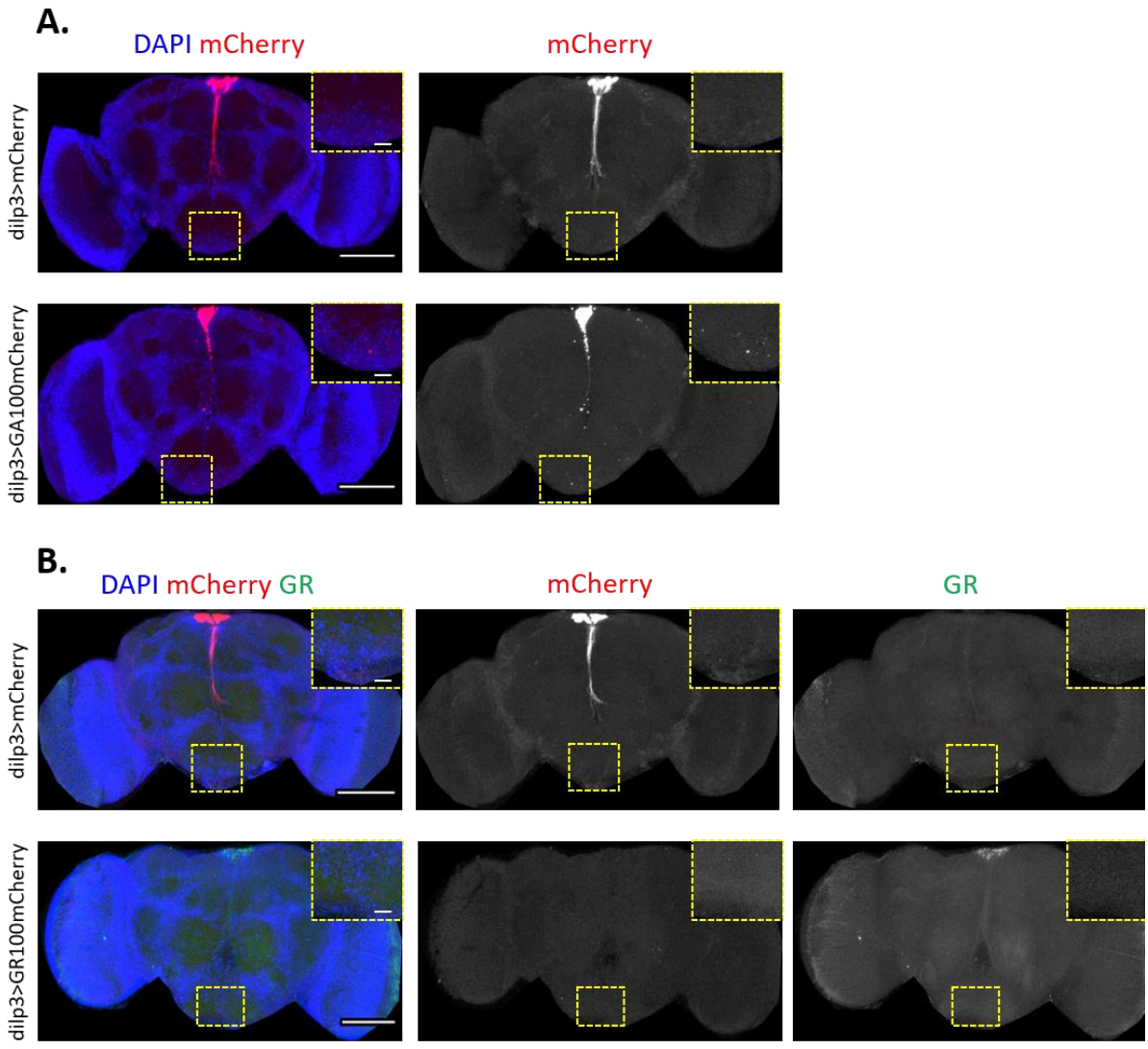

Figure 11. GA100mCherry, but not GR100mCherry, spread outside of MNCs.
A Representative images of 5-day-old fly brains expressing mCherry or GA100mCherry in MNCs under the constitutive driver dilp3-Gal4. The same settings were applied to all genotypes while imaging their mCherry signals. Spreading of GA100mCherry was observed. **B** Representative images of 5-day-old fly brains expressing mCherry or GR100mCherry in MNCs under dilp3-Gal4, and stained with an anti-GR antibody. The same settings were applied to all genotypes while imaging their mCherry and GR signals. No spreading of GR100mCherry was observed. Insets of the highlighted brain regions are shown. Scale bars in images and insets are 100 µm and 10 µm, respectively.

Altogether, our data indicate that, out of the three toxic DPRs, at least for DPRs up to 100 repeats in length, only GA DPRs spread from different neuronal subsets to the rest of the central brain, and the 36 repeat GA DPR may be too short to do so.

3.1.1.3 GA repeat length modulates the aggregation and spread of GA DPRs

Unlike GA100mCherry, GA36mCherry did not form puncta and did not spread, suggesting that transmission is aggregation- and repeat length-dependent. To test this hypothesis further, we generated novel fly lines expressing untagged or C-terminally mCherry-tagged GA200 (Figure 12A, B).

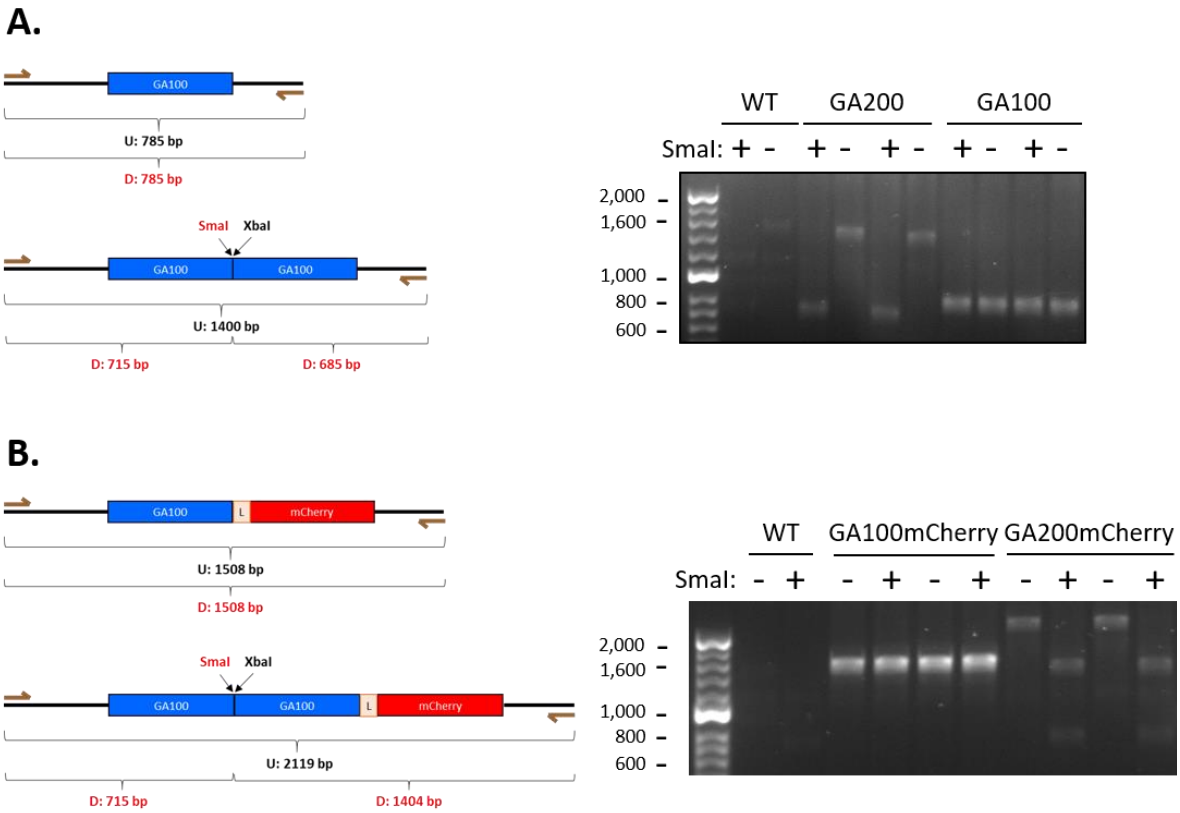

Figure 12. Validation of GA200 and GA200mCherry constructs.
A Schematic and genotyping results showing the construct sizes of UAS-GA100 and UAS-GA200 homozygous flies upon PCR amplification of the DNA sequences of their UAS transgenes with (red, +) or without (black, -) subsequent digestion with SmaI. **B** Schematic and genotyping results showing the construct sizes of UAS-GA100mCherry and UAS-GA200mCherry homozygous flies upon PCR amplification of the DNA sequences of their UAS transgenes with (red, +) or without (black, -) subsequent digestion with SmaI.

Similarly to the previous constructs, a spacer sequence consisting of five Gly-Ser dipeptides was placed between the GA and the mCherry sequences. The GA200 stretch consisted of two identical sequences encoding GA100 and separated by two restriction sites that were used for cloning purposes. The transgenes were inserted into the same genomic locus (attP2 landing

site). To confirm the correct insertion of our constructs, I PCR amplified the UAS transgenes using primers that bind to the flanking sequences of the transgenes, and subsequently digested the amplicons with SmaI, which should only cut the GA200 sequences as depicted in Figure 12A and Figure 12B. Both the digested and the undigested amplicons had the expected sizes (Figure 12A, B), thus confirming the correct insertion of our constructs.

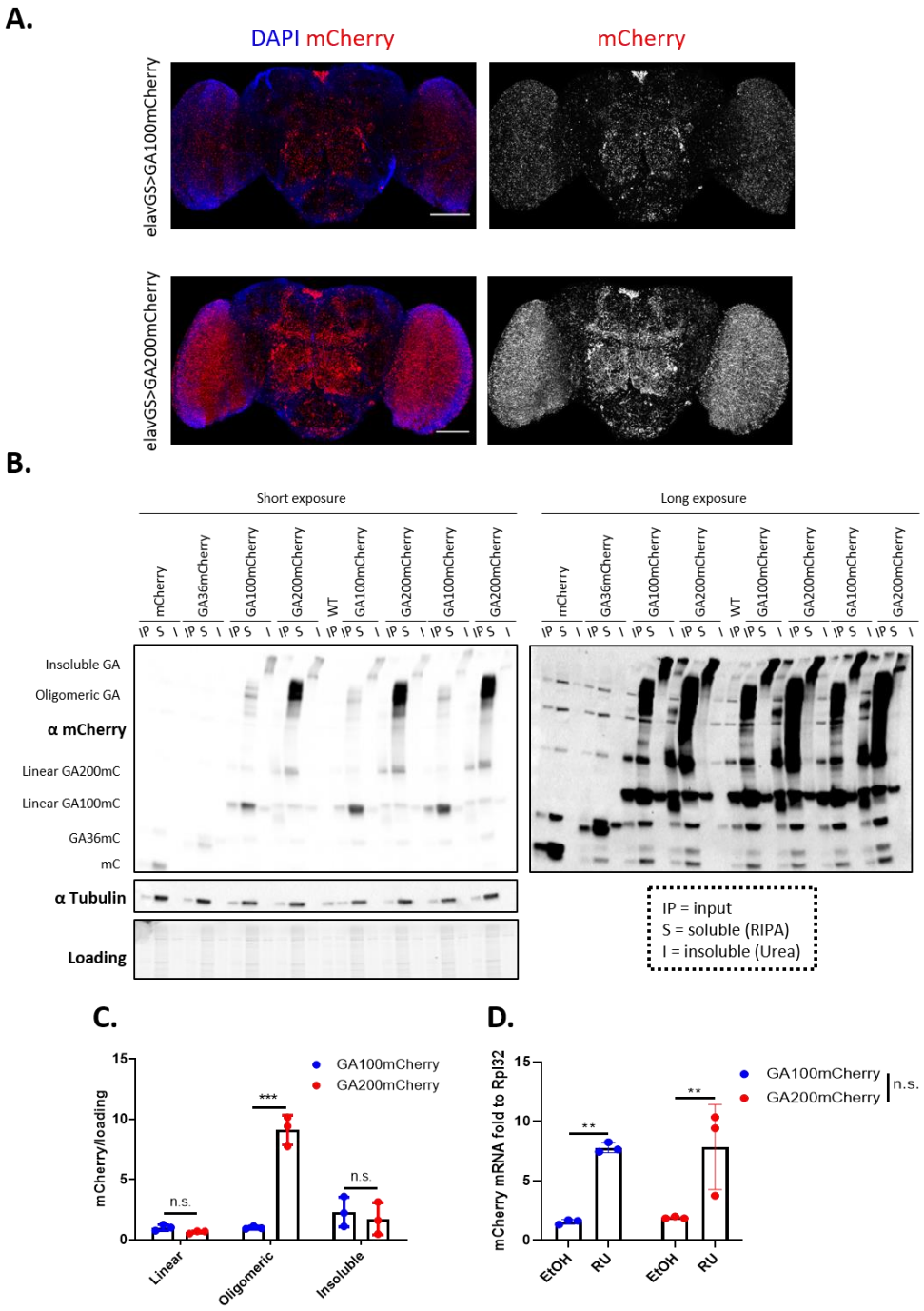

Figure 13. GA200mCherry forms more abundant puncta and oligomeric bands than GA100mCherry upon pan-neuronal expression.

(Below)

Figure 13. GA200mCherry forms more abundant puncta and oligomeric bands than GA100mCherry upon pan-neuronal expression.

A Representative images of 7-day-old fly brains induced to express GA100mCherry or GA200mCherry for 5 days under the pan-neuronal driver elavGS. The same settings were applied to both genotypes while imaging their mCherry signals. Scale bars in images are 100 µm. **B** Immunoblotting results of fractionated protein extracts of heads from flies induced to express mCherry, GA36mCherry, GA100mCherry or GA200mCherry for 5 days with elavGS, and probed with anti-mCherry and tubulin antibodies. 10% of each fraction was loaded. The lack of tubulin in the insoluble fractions proves its purity. IP = input. S = soluble fraction. I = insoluble fraction. **C** Quantification of the linear, oligomeric and insoluble mCherry signals in GA100mCherry- and GA200mCherry-expressing flies after 5 days of expression induction. The mCherry signals were normalized to protein loading (***P < 0.001; t-test, n = 3 sets of 20 fly heads) **D** qRT-PCR results of normalized mCherry transcript levels in heads from elavGS > UAS-GA100mCherry and elavGS > UAS-GA200mCherry flies fed EtOH or the inducing drug RU486 (RU) for 7 hours (treatment: ***P < 0.001; genotype: n.s.; interaction: n.s.; Two-way ANOVA with Bonferroni's multiple comparisons test, n = 3 sets of 20 fly heads).

To further validate and characterize the new lines, I generated flies that pan-neuronally expressed GA100mCherry or GA200mCherry for 5 days and imaged their mCherry signals. While both constructs formed puncta, these accumulated more upon GA200-mCherry expression (Figure 13A), suggesting that GA200 aggregates more than GA100.

To further confirm this, I extracted the soluble and insoluble protein fractions from heads of flies pan-neuronally expressing GA36mCherry, GA100mCherry or GA200mCherry for 5 days, and performed western blotting using an anti-mCherry antibody. The majority of the mCherry signal was detected in the soluble fraction for all of the three polyGAmCherry constructs (Figure 13B). However, unlike GA36mCherry expression, which only gave rise to a soluble and linear product, GA100mCherry and GA200mCherry formed both a soluble linear and a soluble oligomeric product, the latter being significantly more abundant in GA200-mCherry-expressing flies (Figure 13B, C). In addition, I also observed the presence of insoluble GA that did not run through the gel in flies expressing GA100mCherry or GA200mCherry. However, no differences in insoluble GA were observed between these two genotypes (Figure 13B, C). Of note, flies expressing GA100mCherry or GA200mCherry responded equally to feeding of the inducing drug RU486, as I confirmed equal transcript levels by qRT-PCR after 7 hours of drug feeding (Figure 13D). Therefore, GA200mCherry formed more oligomeric aggregates than GA100mCherry despite similar baseline transcription rates.

Next, I measured the spread of GA100-mCherry and GA200-mCherry from ORNs. To exclude the possibility that co-expression of eGFP-tagged synaptotagmin could influence spreading, I only expressed polyGAmCherry. I found accumulation of mCherry puncta of both DPRs outside of ORNs after 3 days of expression induction, with substantially greater spread of the 200 than the 100 GA DPR (Figure 14A, B).

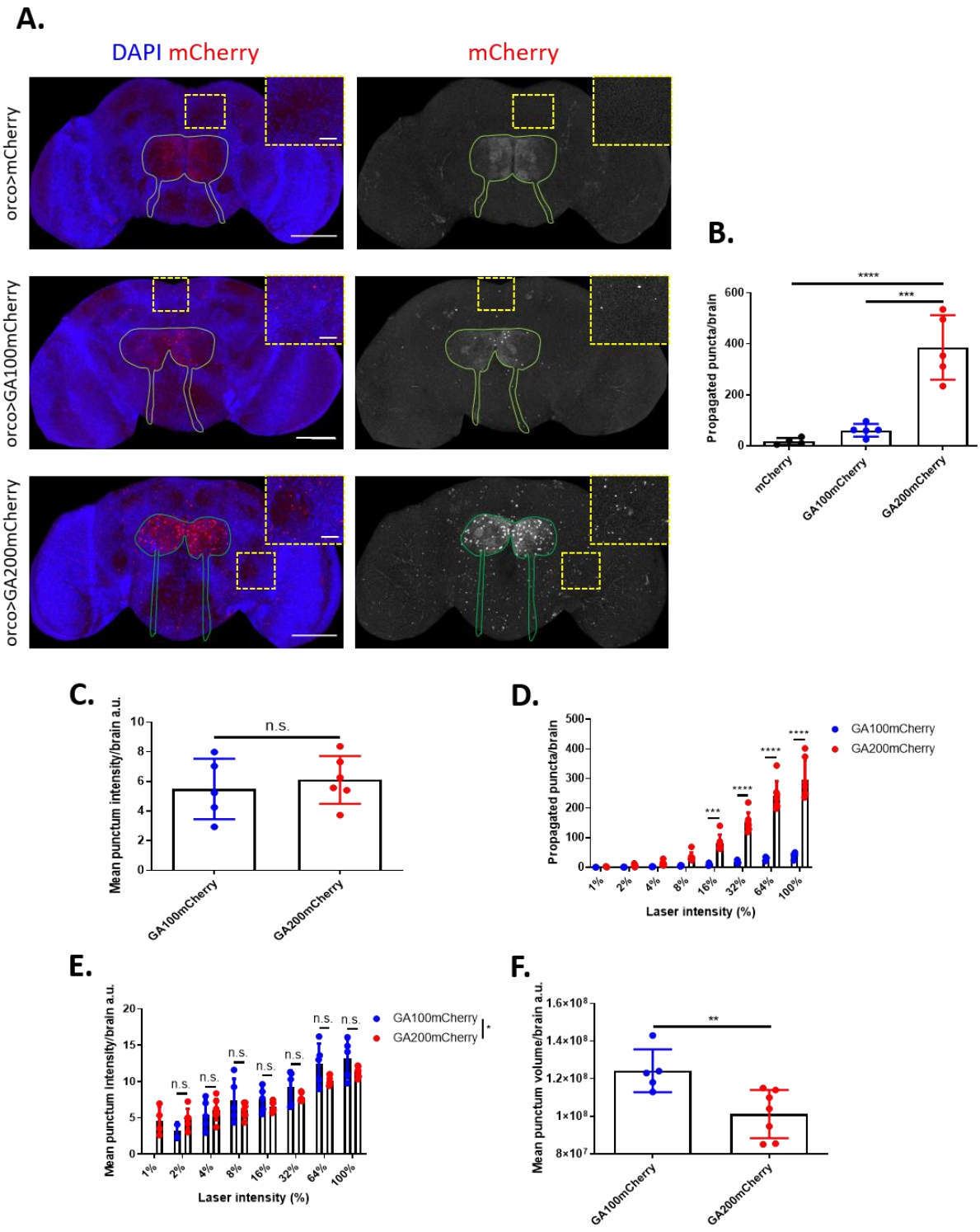

Figure 14. Spreading is higher for longer mCherry-tagged GA repeat proteins.
A Representative images of 5-day-old fly brains expressing mCherry, GA100mCherry and GA200mCherry from ORNs for 3 days under the ORN-specific driver orco-Gal4. The same settings were applied to all genotypes while imaging their mCherry signals. Spreading was greater for GA200mCherry than for GA100mCherry. The mCherry signal was detected using its fluorescence as read-out. Flies expressing only mCherry served as negative control to ensure that mCherry does not spread by itself. The boundaries of the ORN axons and synaptic terminals are highlighted with a solid green line. Insets of the indicated areas are shown to facilitate visualization. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B** Quantification of mCherry puncta detected in the central brain outside of ORNs across genotypes per brain after induction for 3 days (****P < 0.0001 and ***P < 0.001; One-way ANOVA with Tukey's multiple comparisons test, n = 4–5 brains). **C** Quantification of the mean punctum intensity of non-saturated propagated puncta from orco > GA100mCherry and orco > GA200mCherry flies (laser intensity = 4%) (n.s. = not significant; t-test, n = 5-6 brains). **D** Quantification of

detected propagated puncta using increasing laser power settings in orco > GA100mCherry and orco > GA200mCherry flies (genotype: ****P < 0.0001; laser intensity: ****P < 0.0001; interaction: ****P < 0.0001; Two-way ANOVA with Bonferroni's multiple comparisons test, n = 4–7 brains). **E** Quantification of the mean punctum intensity using increasing laser power settings in orco > GA100mCherry and orco > GA200mCherry flies (genotype: *P < 0.05; laser intensity: ****P < 0.0001; interaction: n.s. = not significant; Two-way ANOVA with Bonferroni's multiple comparisons test, n = 4–7 brains). **F** Quantification of the mean punctum volume using 100% as laser power in orco > GA100mCherry and orco > GA200mCherry flies (**P < 0.01; t-test, n = 4-7 brains).

To rule out that the greater spread of GA200mCherry was simply a consequence of increased detection of GA200mCherry puncta due to increased intensity, I measured the mean intensity of the propagated puncta using settings where the puncta were not overexposed, to achieve an accurate intensity measurement. No difference was found between the mean intensity of GA100mCherry and GA200mCherry puncta (Figure 14C), indicating that the differences in the number of propagated puncta were not due to differences in punctum intensity between genotypes. Furthermore, to confirm that the greater detection of GA200mCherry propagated puncta compared to GA100mCherry was not observed due to the specific settings that I used while imaging, I imaged brains with 8 different settings and quantified the number of propagated puncta after expression in ORNs for 3 days. I found a statistically significant interaction between the genotype and the laser intensity used for imaging (Figure 14D), indicating that increasing laser power settings only improved puncta detectability in GA200mCherry-expressing flies. Therefore, GA200mCherry expression in ORNs may trigger the spread of GA aggregates containing a more variable number of molecules than that of GA100mCherry, where all of the propagated aggregates would comprise a number of GA molecules comparable to the biggest GA200mCherry aggregates. This would account for the similar mean intensity of the first detected puncta (Figure 14C), which would presumably be the biggest (i.e., brightest) aggregates. To corroborate this interpretation, I measured the mean intensity of all of the puncta at each of the imaging settings tested and, in agreement with our hypothesis, I found greater mean intensity for GA100mCherry than for GA200mCherry (Figure 14E). This suggests that I mostly detected new and less bright puncta (i.e., initially not detectable) for GA200mCherry as I increased our imaging settings. Finally, to further confirm our interpretation, I measured the mean punctum volume using maximal settings and found an overall smaller punctum volume for GA200mCherry than for GA100mCherry (Figure 14F). Collectively, these analyses indicate that spreading of GA was more pronounced with longer repeats and was independent of eGFP-tagged synaptotagmin co-expression.

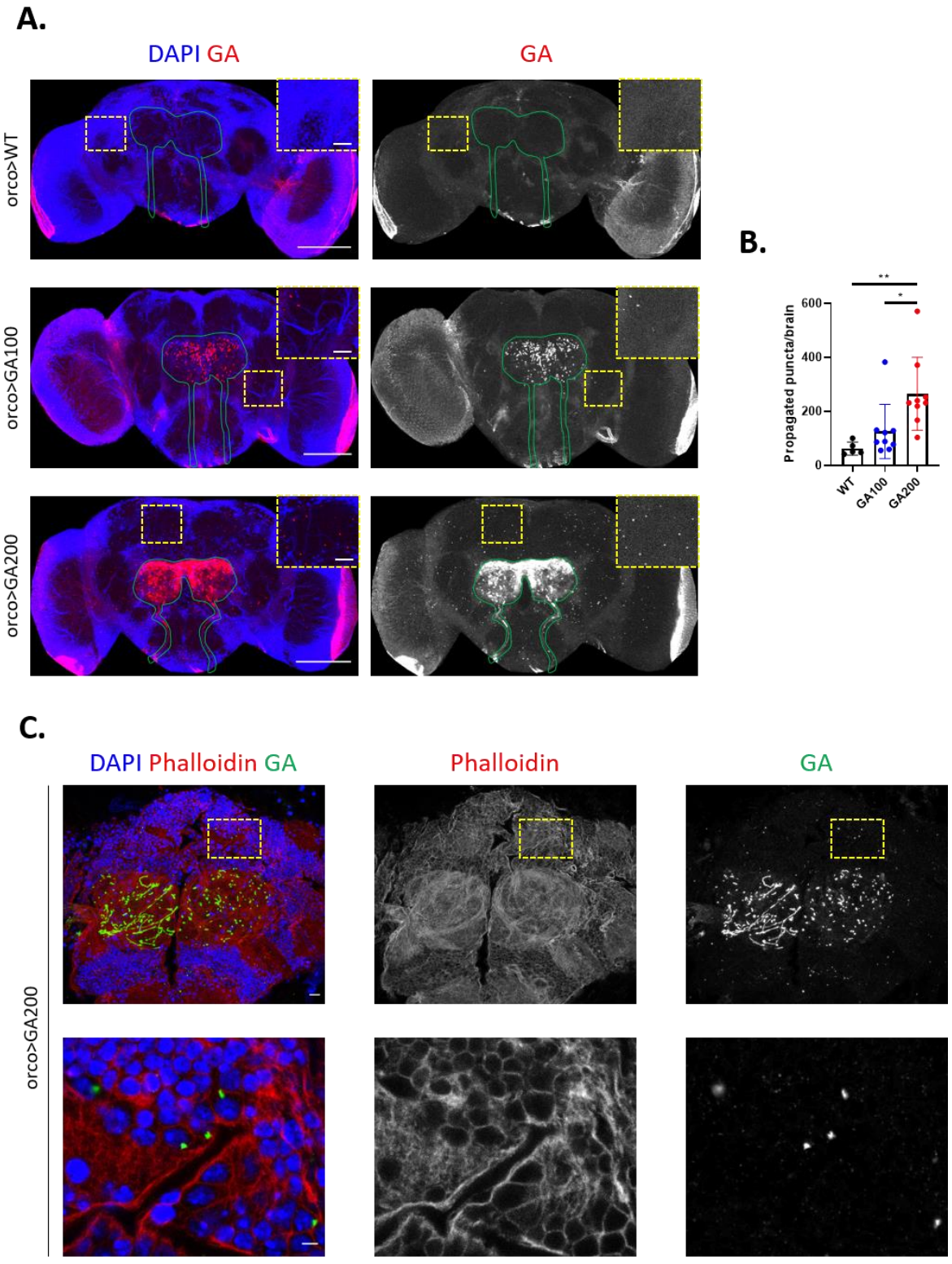

Figure 15. Spreading is increased in longer untagged GA repeat proteins.

A Representative images of 5-day-old fly brains expressing GA100 or GA200 in ORNs for 3 days and probed with an anti-GA antibody. GA200 spread more than GA100. The same settings were used while imaging the GA signal across genotypes. Flies expressing only the driver (WT) served to control for unspecific binding of the anti-GA antibody. Mild unspecific binding to trachea and the lamina of the optic lobes was observed. The boundaries of the ORN axons and synaptic terminals are highlighted with a solid green line. Insets of the indicated areas are shown to facilitate visualization. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B** Quantification of GA puncta detected outside of the ORN boundaries across genotypes per brain after induction for 3 days (**P < 0.01 and *P < 0.05; One-way ANOVA with Tukey's multiple comparisons test, n = 5–9 brains). **C** Representative image and indicated insets of a 5-days-old fly brain expressing GA200 in ORNs for 3 days, and stained with an anti-GA antibody (green) and the rhodamine-conjugated fluorophore

phalloidin (red). Five cells positive for GA intracellular puncta can be observed in the enlarged insets. Scale bars in image and inset are 25 µm and 3 µm, respectively.

Since tags can interfere with protein function (Saiz-Baggetto et al. 2017), I next tested the spread of untagged GA DPRs using GA100 and GA200 expressed in ORNs and a GA-specific antibody (Solomon et al. 2018). In agreement with our results using mCherry-tagged GA constructs, I found that the number of GA puncta detected outside of ORNs dramatically increased with repeat length (Figure 15A, B), further supporting the notion that the propensity of GA to spread is greater in longer GA repeats. To further characterize the propagated puncta, I co-stained brains from flies expressing GA200 in ORNs with fluorescently labelled phalloidin, a dye that strongly binds to actin F and can therefore be used to identify the boundaries of single cells in tissue (Albertson et al. 2013). Using this approach, I detected GA-positive puncta in the cytoplasm of recipient cells, thus indicating that propagated GA puncta are intracellular (Figure 15C).

To determine if GA could also spread from different types of neurons in a repeat length-dependent manner, I initiated expression in optic lobe (OL) neurons. I expressed the GA constructs using the R9D03-Gal4 driver and, consistent with our finding in ORNs, GA also spread in a repeat-length dependent manner from OL neurons (Figure 16A, B).

Altogether, our results show that, in two independent neuronal subsets, longer GA repeats spread in a length-dependent manner.

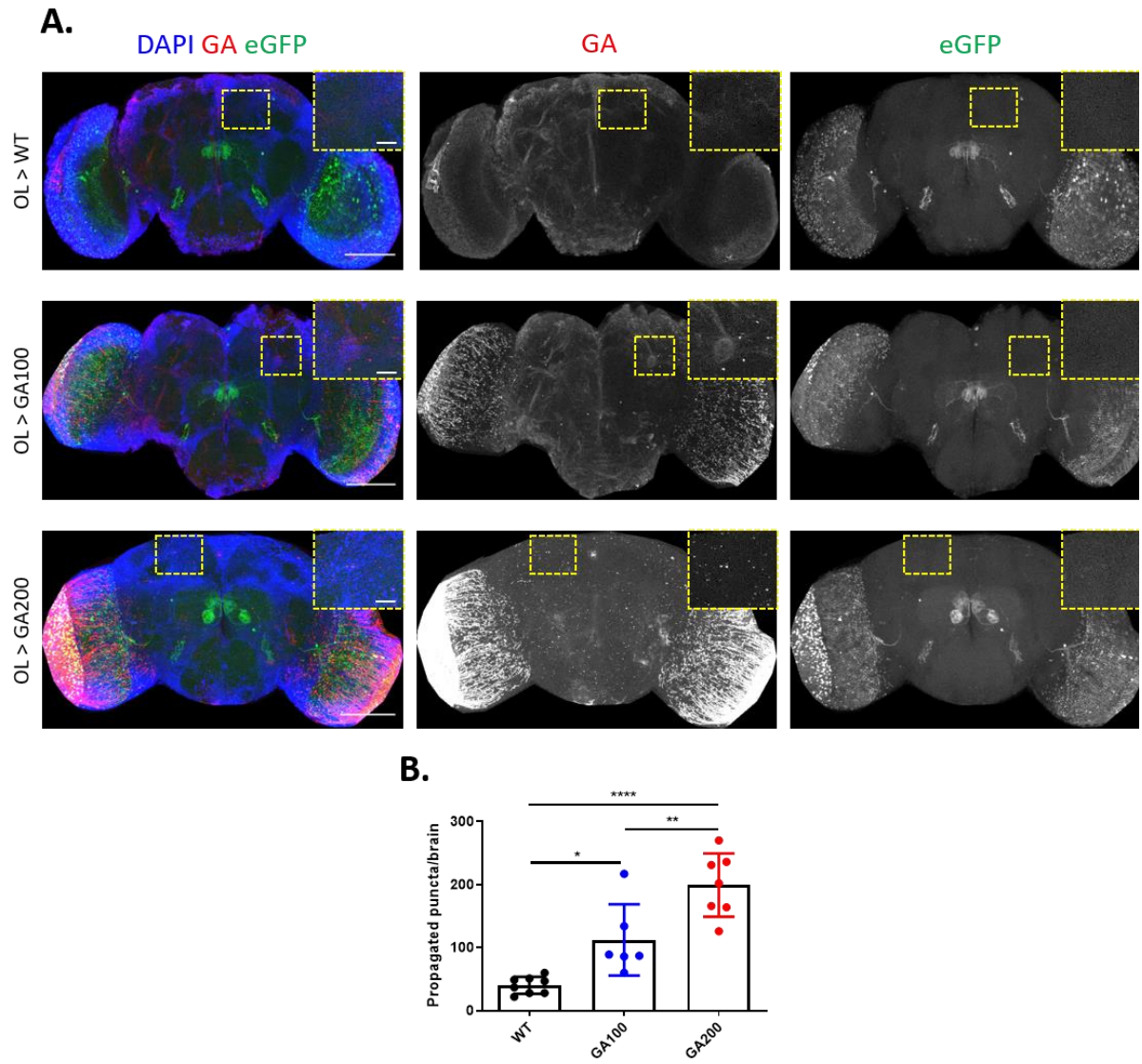

Figure 16. GA spreads in a repeat length-dependent manner from an independent neuronal population.
A Representative images of 5-day-old fly brains from control flies (expressing only the OL-Gal4 driver) and flies expressing GA100 or GA200 in the OLs for 3 days and probed with an anti-GA antibody. GA200 also spread more than GA100 from this brain region. The same settings were used while imaging the GA signal across genotypes. Flies expressing only the driver were used to control for unspecific binding of the anti-GA antibody. Mild, unspecific binding to trachea and the lamina of the optic lobes was observed. EGFP with a nuclear localization signal was co-expressed to identify the cells targeted by the OL-Gal4 driver. **B** Quantification of GA puncta detected in the central brain outside of the targeted cells after expression of the indicated constructs for 3 days (****P<0.0001, **P < 0.01 and *P < 0.05; One-way ANOVA with Tukey's multiple comparisons test, n = 6–8 brains). Insets of the indicated areas are shown to facilitate visualization. Scale bars in images and insets are 100 µm and 10 µm, respectively.

3.1.1.4 GA DPRs exhibit an age-related increase in spreading

Given that ageing is a major risk factor for ALS and FTD (Niccoli, Partridge, and Isaacs 2017), I investigated whether GA spread was affected by the age at which I induced GA expression. I induced ORN-specific expression of untagged GA200 starting in 2-day-old or 30-day-old adult flies for 3 days, and measured the spread outside of ORNs (Figure 17A).

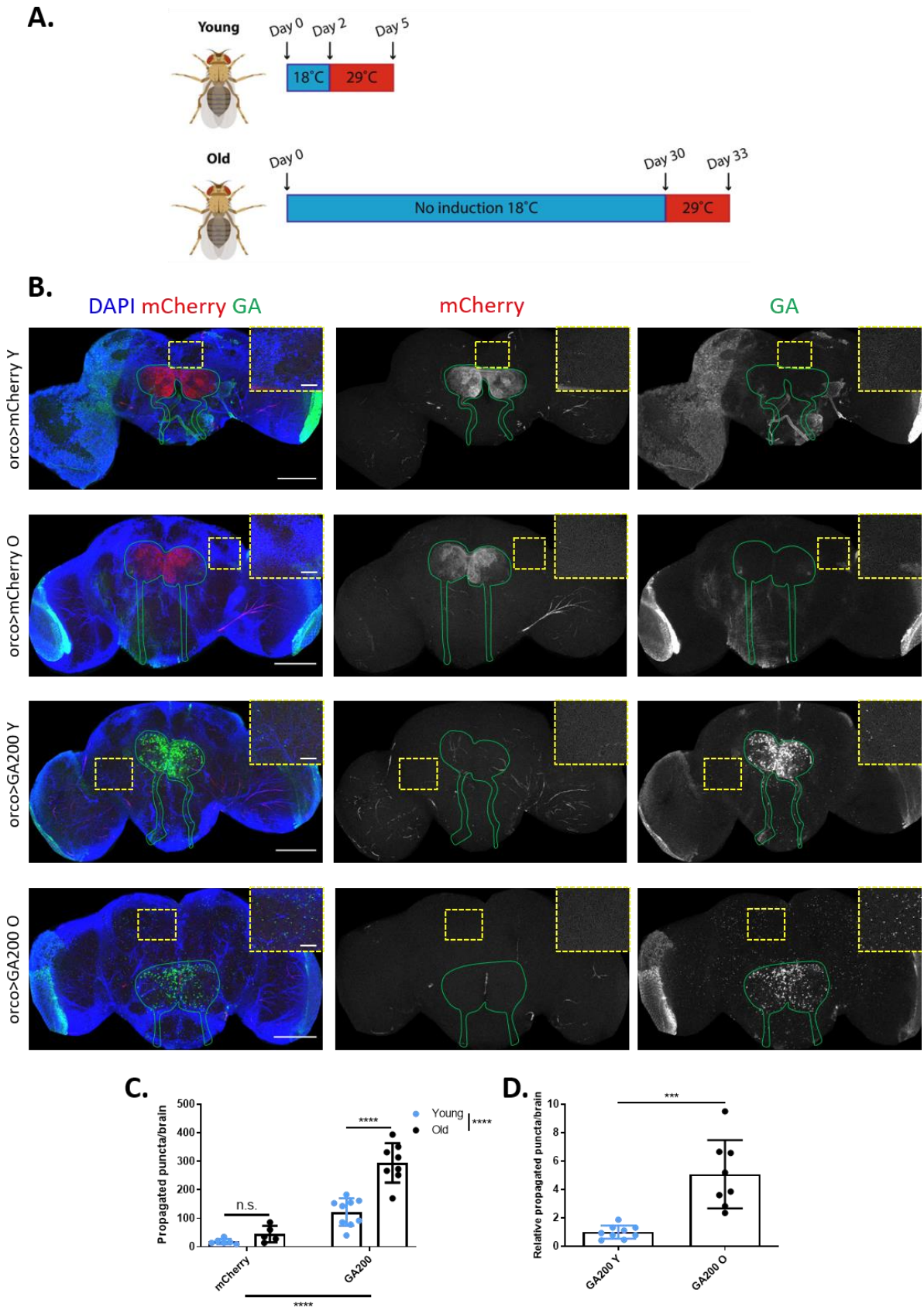

Figure 17. Age-associated factors exacerbate GA spread.

(Below)

Figure 17. Age-associated factors exacerbate GA spread.
A The expression of GA200 was induced for 3 days in ORNs in young (2 days old) and old flies (30 days old), after which GA spread was measured. Fly cartoons were created with BioRender. **B** Representative images of control fly brains expressing mCherry or GA200 in ORNs for 3 days in young (Y) or old (O) flies. Brains were probed with an anti-GA antibody (green). The outline of ORN axons and synaptic terminals is shown in green. Insets of indicated areas highlight differences in the number of propagated dots across conditions. Mild unspecific binding to trachea and the lamina of the optic lobes was observed. Scale bars in images and insets are 100 µm and 10 µm, respectively. **C** Quantification of the total number of GA-positive puncta detected outside of ORNs after 3 days of expression in young and old flies (age: ****$P < 0.0001$; genotype: ****$P < 0.0001$; interaction: ***$P < 0.001$; Two-way ANOVA with Bonferroni's multiple corrections test, $n = 5–9$ brains). **D** Quantification of the number of propagated GA-positive puncta relative to the GA signal in ORNs after 3 days of expression in young and old flies (***$P < 0.001$; t-test, $n = 5–9$ brains).

There was a 3-fold increase in the total number of propagated GA puncta when GA expression was induced at the older age (Figure 17B, C). Given that the accumulation of the peptides could change after expression induction at different ages, I also quantified the cumulative number of propagated GA puncta relative to GA expression in ORNs, as an indicator of whether the proportion of propagated GA compared to the total amount of GA in ORNs changed at different ages. Indeed, I found a larger proportion of propagated GA compared to ORN GA after expression induction in older brains (Figure 17D), thus showing that the increased spread in older brains was not simply due to changes in the accumulation of GA in ORNs upon expression in older brains. Collectively, these results suggest that age-associated factors strongly affect the propagation propensity of GA DPRs.

3.1.2 Consequences of GA transmission in the fly brain

Neuropathological studies have shown signs of proteostasis impairment in the brains of patients with various neurodegenerative diseases, including accumulation of protein deposits whose poly-ubiquitination indicates that they had been labelled for degradation (Shahheydari et al. 2017). Autophagy plays a key role in the removal of misfolded poly-ubiquitinated proteins and dysfunctional organelles in ageing and neurodegeneration (Menzies et al. 2017). Specifically, in the *C9orf72* mutation context, DPRs co-aggregate with the autophagic receptor p62 (Al-Sarraj et al. 2011; Mori, Weng, et al. 2013), which typically binds to poly-ubiquitinated proteins and signals their co-uptake into autophagosomes for subsequent co-degradation of p62 and the protein aggregates in the lysosomes (Menzies et al. 2017). Since accumulation of p62 is, therefore, usually interpreted as a sign of autophagic flux impairment, higher levels of p62 in the brains of *C9orf72* ALS/FTD suggests autophagy failure and impairment of proteostasis upon accumulation of the DPRs.

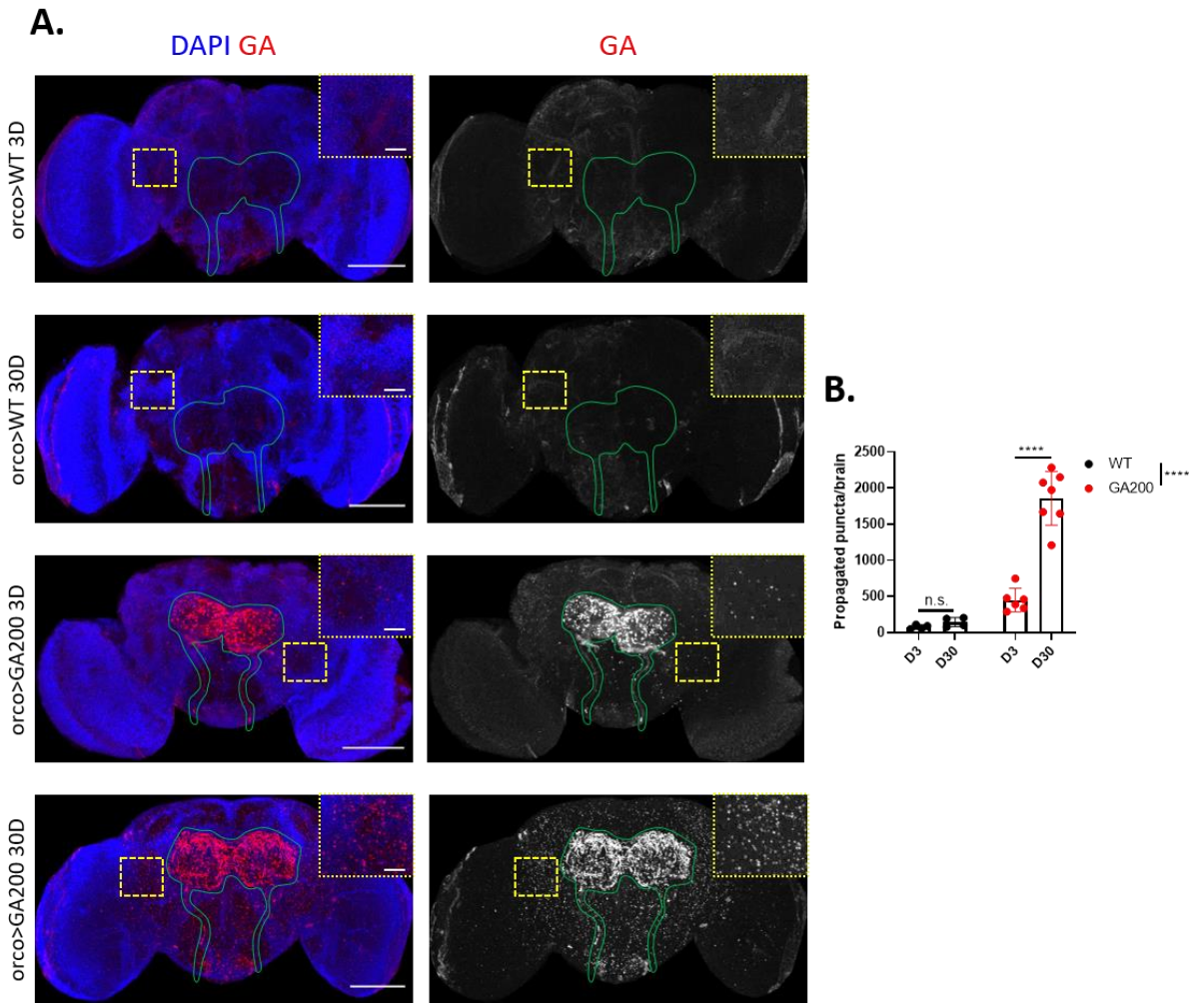

Figure 18. Propagated GA accumulates over time.
A Representative images of control fly brains expressing the driver-only (i.e., orco>WT) or GA200 in ORNs for 3 or 30 days. Brains were probed with an anti-GA antibody (red) and the same laser power settings were used for imaging of all conditions. The outline of ORN axons and synaptic terminals is shown in green. Mild unspecific binding to trachea and the lamina of the optic lobes was observed. Insets of indicated areas highlight differences in the number of propagated dots across conditions. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B** Quantification of the total number of GA-positive puncta detected outside of ORNs after 3 or 30 days of expression (age: ****$P < 0.0001$; genotype: ****$P < 0.0001$; interaction: ****$P < 0.0001$; Two-way ANOVA with Bonferroni's multiple corrections test, n = 4-7 brains).

While many aggregation-prone proteins associated with various neurodegenerative diseases have been shown to spread in animal models, it remains unclear whether this spread leads to a clear response in the recipient brain areas of these models. Given that previous studies have shown that GA expression causes DNA damage and impairs proteostasis (Q. Guo et al. 2018; Y. B. Lee et al. 2017; Y. J. Zhang et al. 2016), I investigated whether GA spread could also trigger these effects in the adult fly brain. First, I tested whether GA-propagated puncta accumulated over time in the recipient tissue. Indeed, 30-day induction of GA200 expression in ORNs led to a greater accumulation of propagated GA puncta compared to expression induction for 3 days

(Figure 18A, B). Next, I performed a TUNEL assay to stain for dsDNA breaks, which is a common feature of apoptotic cells.

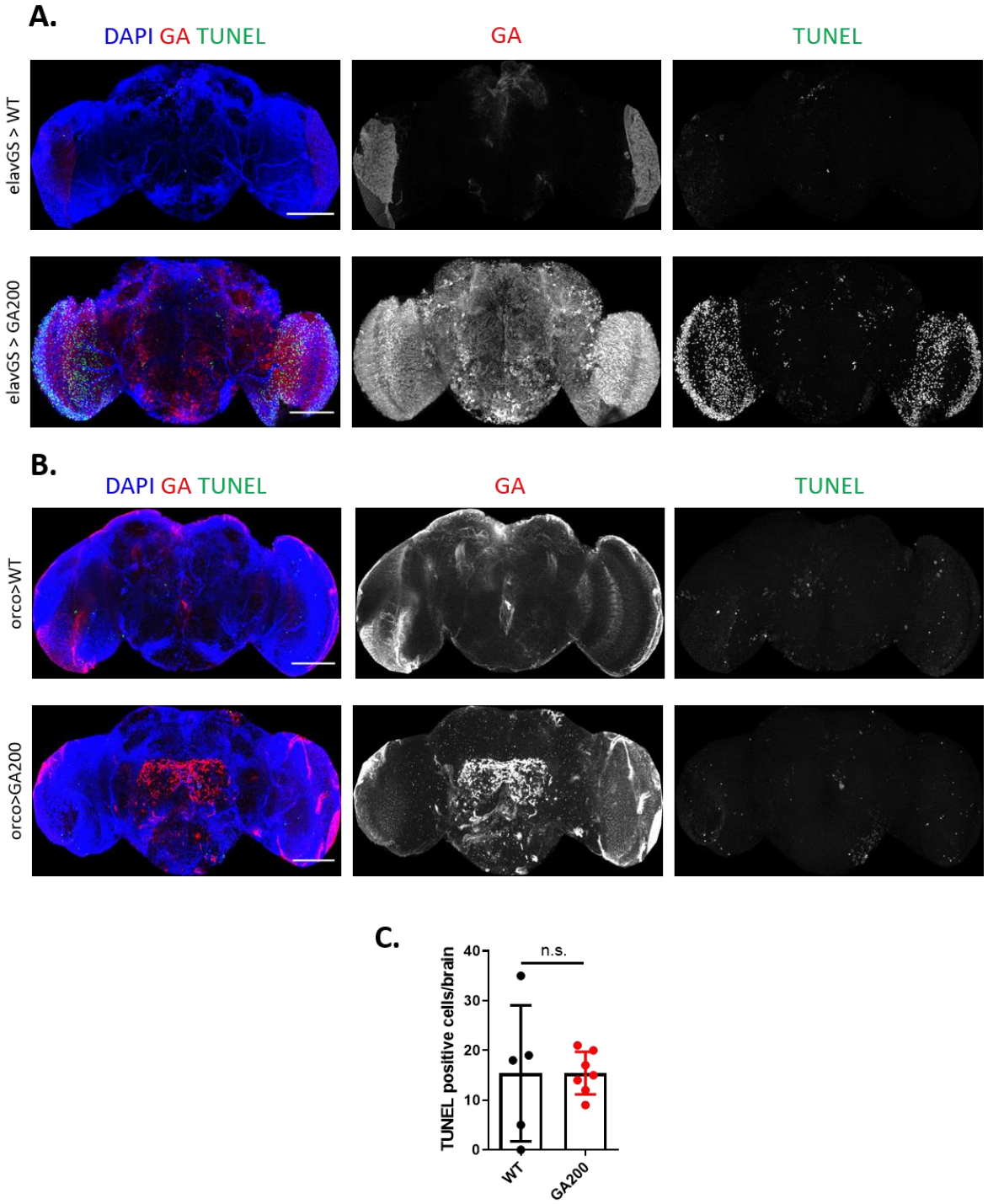

Figure 19. Long-term GA propagation from ORNs does not increase brain DNA damage.
A Representative images of fly brains expressing GA200 under the pan-neuronal driver elavGS after 30 days of expression induction, and stained with an anti-GA antibody (red) and the TUNEL dye (green) to detect dsDNA breaks. DNA damage was more prominent in GA200-expressing brains compared to their driver-only control. **B** Representative images of control fly brains expressing the driver-only (i.e., orco>WT) or GA200 in ORNs for 30 days. Brains were probed with an anti-GA antibody (red) and the TUNEL dye (green) to detect DNA damage. Scale bars in images are 100 µm. **C** Quantification of the total number of TUNEL positive cells at 30 days of expression (n.s. = not significant; t-test, *n* = 5-7 brains).

Pan-neuronal GA200 expression for 30 days strongly increased the number of brain cells containing dsDNA breaks compared to driver-only flies (Figure 19A), indicating that the ability of GA to induce DNA damage is also present in our fly GA models. Of note, DNA damage was particularly apparent in the OLs despite pan-neuronal GA200 expression (Figure 19A), and these neurons have been reported to be particularly vulnerable to the expression of other aggregation-prone proteins (Latouche et al. 2007). However, inducing GA200 spread from ORNs for 30 days did not increase DNA damage in the recipient tissue compared to driver-only flies (Figure 19B, C). This might be due to the relatively low spread of GA200 from ORNs to the OLs (Figure 19B).

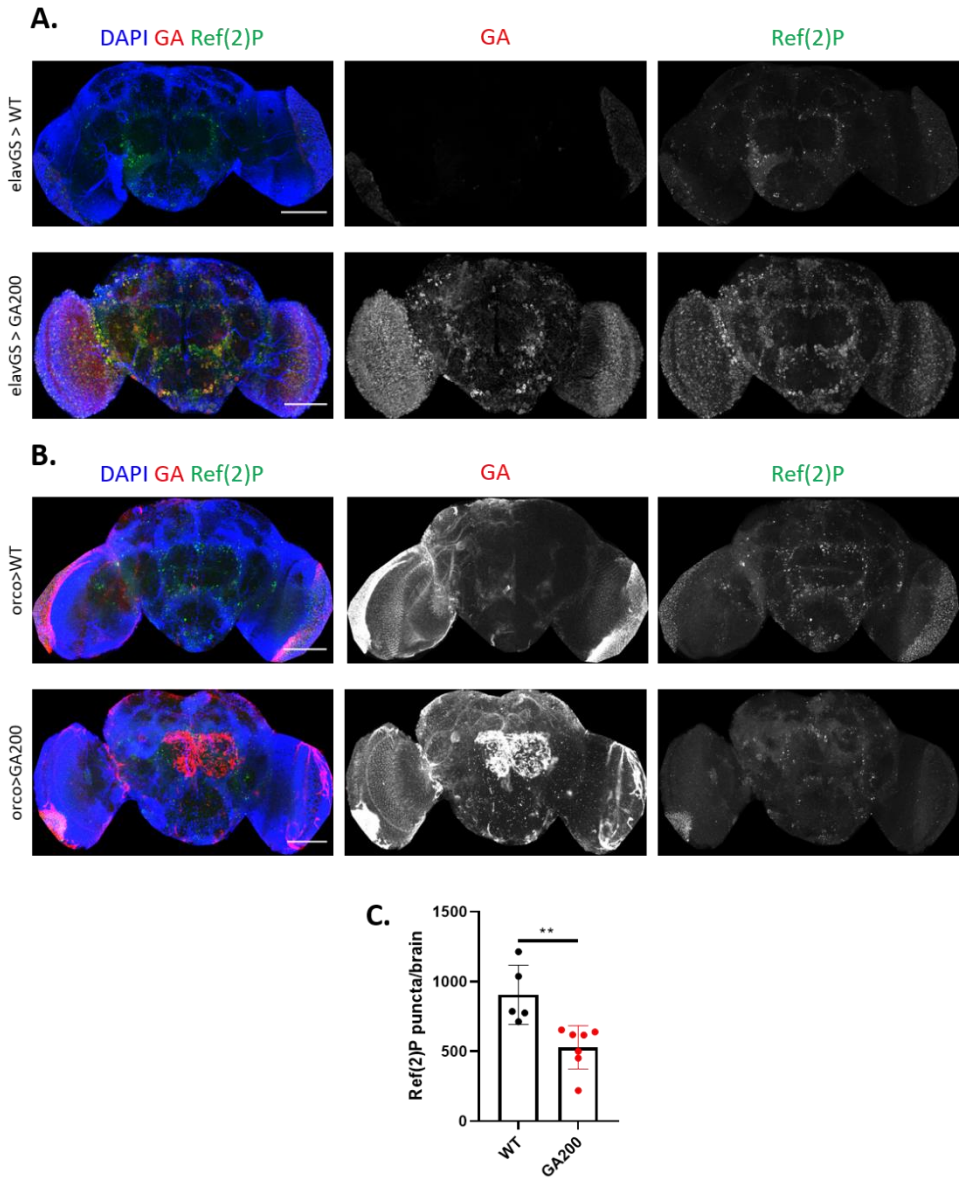

Figure 20. Long-term GA propagation from ORNs lowers brain Ref(2)P puncta.
(Below)

Figure 20. Long-term GA propagation from ORNs lowers brain Ref(2)P puncta.
A Representative images of fly brains expressing GA200 under the pan-neuronal driver elavGS after 30 days of expression induction, and stained with anti-GA (red) and anti-Ref(2)P (green) antibodies. Ref(2)P signal was stronger in GA200-expressing brains compared to their driver-only control. The pattern of the Ref(2)P signal in control brains was punctated, while in GA200 was rather diffuse. **B** Representative images of control fly brains expressing the driver-only (i.e., orco>WT) or GA200 in ORNs for 30 days. Brains were probed with anti-GA (red) and anti-Ref(2)P (green) antibodies. The pattern of the Ref(2)P signal was rather punctated in both genotypes. Scale bars in images are 100 µm. **C** Quantification of the total number of Ref(2)P puncta after 30 days of ORN-specific expression of GA200 (**P<0.01; t-test, *n* = 5-7 brains).

To test for an effect of GA expression on autophagy, I performed adult fly brain stainings against refractory to sigma P (Ref(2)P), which is the fly homolog of p62 (Nezis et al. 2008). Pan-neuronal expression of GA200 for 30 days strongly increased Ref(2)P signal, which, similarly to human p62, often co-localized with GA in neuronal somata (Figure 20A). However, inducing GA200 spread from ORNs for 30 days led to lower Ref(2)P levels in the recipient tissue compared to driver-only flies (Figure 20B, C). While this indicates that the recipient tissue responds to GA spread, lower Ref(2)P levels suggest increased autophagy upon GA transmission. This may result from the relatively low amount of GA in our ORN spread paradigm compared to the pan-neuronal expression model, with low GA levels increasing autophagy activation and high levels blocking it. As elaborated in the discussion section, GA spread may have different consequences in the human context of the *C9orf72* mutation, where autophagy may be impaired due to lower levels of the autophagic C9orf72 protein.

3.1.3 Mechanisms of GA transmission in the fly brain

3.1.3.1 GA200 transmission from ORNs is reduced upon down-regulation of genes involved in exosomal release

The trans-cellular transmission of pathological proteins may occur through various pathways (Figure 2). One previous *in vitro* study found that, compared to the other DPRs, GFP-GA50 was enriched in the exosomal fraction of the culture medium of transfected mammalian cortical neurons, and GFP-GA50-containing exosomes could transmit this DPR to naïve neurons in culture (Westergard et al. 2016).

To explore whether this transmission mechanism plays a role in our *in vivo* paradigm of GA200 spread from ORNs, I modulated the activity or expression of several genes previously reported to play a role in the release of exosomes in *Drosophila* (Koles et al. 2012; Lauwers et al. 2018). To this end, I generated a homozygous orco-Gal4 and UAS-GA200 (expressed from the attP40 site for practical genetic reasons; see methods) fly stock and crossed it with flies homozygous

for UAS constructs expressing a dominant negative (DN) isoform of Rab11 (Rab11DN) or RNA interference (RNAi) constructs against syntaxin 1A (*syntx1* RNAi), heat shock protein 90 (*hsp90* RNAi) or the Hsp90 co-factor stress-induced phosphoprotein 1 (*stip1* RNAi). All flies also contained a copy of tub-Gal80^{ts} to restrict transgene expression to the adult stage. Since co-expressing two UAS transgenes may dilute the Gal4 driver and lead to relatively lower levels of expression of each UAS transgene, I also co-expressed mCherry as a genetic control. As an additional control, I systematically assessed whether the effect of the transgenes on GA spread stemmed from changed GA levels, by normalizing the propagated puncta to GA signal in the ORN synaptic terminals. At 3 days of expression induction, only Rab11DN expression significantly lowered the number of propagated GA puncta (Figure 21A, C) without altering the amount of synaptic GA (Figure 21B). *Stip1* RNAi mildly decreased synaptic GA levels in our preliminary miniscreen (Figure 21B). However, at this age none of our candidates significantly modified the proportion of propagated GA, but a clear trend was again shown by Rab11DN, which did not reach statistical significance, probably due to an outlier brain in flies expressing GA200-only (Figure 21D).

To further determine the contribution of exosomes to the long-term accumulation of propagated GA from ORNs, I also performed stainings at 30 days of expression induction. While none of our exosomal candidates altered synaptic GA at this age, all of them reduced the accumulation of propagated GA, and this effect was particularly strong for Rab11DN, as well as for *stip1* RNAi and *syntx1* RNAi (Figure 21E-G). I also found a clear trend towards lower relative propagated GA upon long-term Rab11DN co-expression (Figure 21H), which further supported a particularly relevant role for this protein. Unexpectedly, mCherry co-expression significantly increased the number of propagated GA puncta and GA relative spread without affecting GA synaptic accumulation (Figure 21E-H). As elaborated in the discussion section, we hypothesize that the latter may be due to mild proteostasis stress inflicted by mCherry, thus leading to the progressive accumulation of propagation-competent GA.

Overall, our preliminary miniscreen revealed that the release of exosomes plays a role in the spread of GA from ORNs and is therefore a conserved mediator of this phenomenon between cultured mammalian cells and the fly brain.

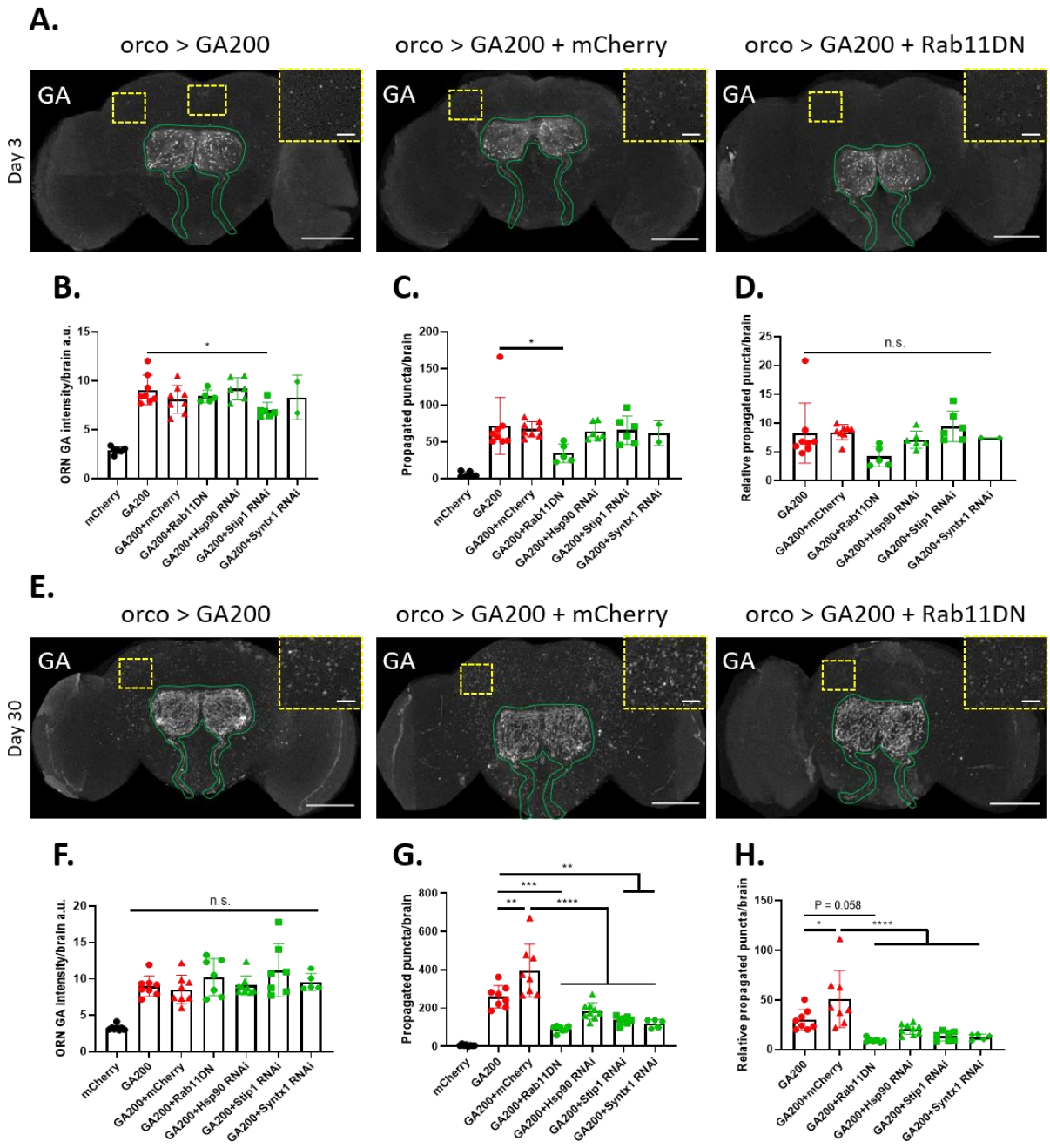

Figure 21. Genetic down-regulation of genes involved in the release of exosomes lowers GA200 spread from ORNs.

A, E Representative images of brains from flies expressing GA200, GA200 and mCherry, or GA200 and Rab11DN in ORNs under the ORN-specific orco-Gal4 driver for 3 **(A)** or 30 **(E)** days. Brains were probed with an anti-GA antibody (white) and the same laser power settings were used for imaging of all conditions. The outline of ORN axons and synaptic terminals is shown in green. Mild unspecific binding to trachea and the lamina of the optic lobes was observed. Insets of indicated areas highlight differences in the number of propagated dots across conditions. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B, F** Quantification of the mean GA intensity in ORNs after 3 **(B)** or 30 **(F)** days of expression of the indicated constructs. mCherry-only was used to control for unspecific binding of the anti-GA antibody and GA200+mCherry was used as a genetic control. Reference genotypes are shown in red, while genotypes expressing a modulator of the release of exosomes are shown in green. Only differences between the reference genotypes and flies expressing an exosomal candidate are shown (*P<0.05; n.s. = not significant; One-Way ANOVA with Tukey's multiple comparisons test, n=2-8 brains). **C, G** Quantification of the total number of propagated GA puncta outside of ORNs after 3 **(C)** or 30 **(G)** days of expression of the indicated constructs (****P<0.0001, ***P<0.001, **P<0.01, *P<0.05; One-Way ANOVA with Tukey's multiple comparisons test, n=2-8 brains). **D, H** Quantification of the relative amount of propagated GA after 3 **(D)** or 30 **(H)** days of expression of the indicated constructs (****P<0.0001, *P<0.05, n.s. = not significant; One-Way ANOVA with Tukey's multiple comparisons test, n=2-8 brains).

3.1.3.2 GA200 transmission from ORNs is affected by modulating genes involved in exocytosis and neuronal activity

While modulating the expression of genes involved in exosomal release reduced GA spread from ORNs, it did not completely block it, thus suggesting that other mechanisms may also be at play.

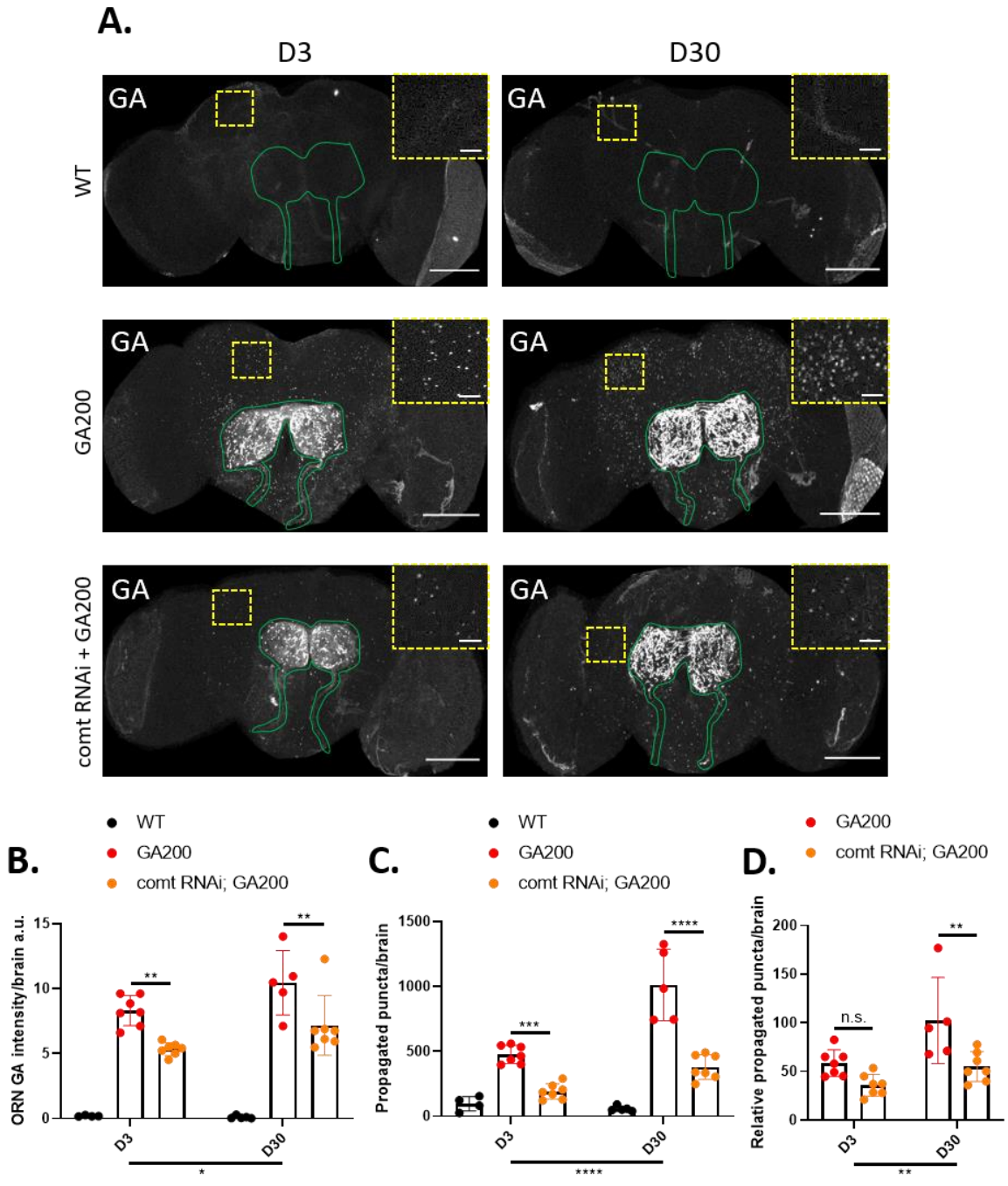

Figure 22. Genetic down-regulation of comatose lowers synaptic and propagated GA200.

(Below)

Figure 22. Genetic down-regulation of comatose lowers synaptic and propagated GA200.Figure 22. Genetic down-regulation of comatose lowers synaptic and propagated GA200.
A Representative images of brains from flies expressing GA200 or GA200 and *comt* RNAi in ORNs under the ORN-specific orco-Gal4 driver for 3 or 30 days. Brains were probed with an anti-GA antibody (white) and the same laser power settings were used for imaging of all conditions. The outline of ORN axons and synaptic terminals is shown in green. Mild unspecific binding to trachea and the lamina of the optic lobes was observed. Insets of indicated areas highlight differences in the number of propagated dots across conditions. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B** Quantification of the mean GA intensity in ORNs after 3 or 30 days of expression of the indicated constructs. Driver-only flies (i.e., orco > WT) were used to control for unspecific binding of the anti-GA antibody. Only differences between GA200 and GA200+*comt* RNAi are shown (age: *$P < 0.05$; genotype: ****$P < 0.0001$; interaction: n.s. = not significant; Two-way ANOVA with Tukey's multiple corrections test, $n = 4$-7 brains). **C** Quantification of the total number of propagated GA puncta outside of ORNs after 3 or 30 days of expression of the indicated constructs (age: ****$P < 0.0001$; genotype: ****$P < 0.0001$; interaction: ****$P < 0.0001$; Two-way ANOVA with Tukey's multiple corrections test, $n = 4$-7 brains). **D** Quantification of the relative amount of propagated GA after 3 or 30 days of expression of the indicated constructs (age: **$P < 0.01$; genotype: ***$P < 0.001$; interaction: n.s. = not significant; Two-way ANOVA with Bonferroni's multiple corrections test, $n = 4$-7 brains).

In agreement with this hypothesis, fewer naïve neurons took up GFP-GA50 after treatment with an exosome-enriched fraction than when treated with the same amount of non-fractioned supernatant (Westergard et al. 2016). Interestingly, a previous study reported that reducing synaptic vesicles by lowering the expression levels of the N-ethylmaleimide-sensitive fusion protein 1 (NSF1) fly homolog comatose (*comt)* strongly decreased the spread of poly-glutamine repeats from ORNs in the fly brain (Babcock and Ganetzky 2015). NSF-1 is an ATPase involved in the disassembly of the protein machinery involved in the docking and fusion of synaptic vesicles at the presynaptic membrane. The disassembly of this protein complex is a pivotal step to maintain physiological levels of synaptic vesicles and efficient neuronal transmission (Kawasaki and Ordway 2009). Therefore, I tested whether reducing Comt levels by *comt* RNAi would have an effect on GA transmission.

I initially assessed this by generating double transgenic flies for UAS-*comt* RNAi and UAS-GA200 (expressed from the attP2 site due to practical genetic reasons; see methods) and crossing them with homozygous orco-Gal4 flies. I stained the brains of the offspring after 3 and 30 days of expression induction. Expression of *comt* RNAi significantly lowered both GA accumulation in ORNs and GA propagated puncta at both ages (Figure 22A-C). However, reduced GA expression did not account for the lower spread of GA at 30 days (Figure 22D), suggesting that synaptic activity, specifically the number of synaptic vesicles, may also play an important role in the long-term propagation of GA.

To further assess the role of synaptic vesicles to GA spread, I performed a similar genetic miniscreen of synaptic activity modulators. I used the same experimental set-up as in our genetic miniscreen of exosomal candidates, but this time I modulated the expression or activity of genes involved in the transport and fusion of synaptic vesicles at the presynaptic

compartment, or in the maintenance of the neuronal resting membrane potential, which determines the physiological fusion of synaptic vesicles. More specifically, our genetic miniscreen included co-expression of the following candidates: i) a DN isoform of the dynamin fly homolog shibire (ShiDN), which, similarly to Comt, plays a role in the recycling of synaptic vesicles and therefore, expression of its DN isoform leads to neuronal silencing (Bengtson and Kitamoto 2001), ii) DN and constitutively active (CA) isoforms of Rab3 (Rab3DN or Rab3CA, respectively), Rab3 playing a key role in the transport of primed synaptic vesicles and proteins to the presynaptic membrane for subsequent membrane fusion (Bae et al. 2016; Graf et al. 2009; Schlüter et al. 2006), iii) *comt* RNAi, iv) a voltage-activated bacterial sodium channel (NaChBac), which hyperexcitates neurons (Ren et al. 2001), v) a DN truncated ether-a-gogo (eagDN) potassium channel subunit, which increases neuronal activity (Broughton, Kitamoto, and Greenspan 2004), and vi) a truncated CA open rectifier potassium channel 1 (Ork1ΔC), which leads to neuronal silencing (Nitabach, Blau, and Holmes 2002). At both 3 and 30 days of expression, blocking the fusion of synaptic vesicles by Rab3DN expression led to a significant accumulation of synaptic GA (Figure 23A, B, E, F). In contrast, decreasing the number of synaptic vesicles by ShiDN expression led to a non-significant but clear trend towards lower synaptic GA at both ages (Figure 23A, B, E, F). Collectively, these results suggested that GA may at least partially accumulate within synaptic vesicles in the ORN terminals.

Moreover, all genetic interventions up- or down-regulating synaptic activity lowered the total and relative amount of propagated GA at day 3 (Figure 23C, D). Surprisingly, our automatic quantification of puncta revealed a strong increase in propagated GA by mCherry co-expression when this miniscreen was performed (Figure 23A, C, D), which I had not detected in our exosomal genetic miniscreen at day 3 (Figure 21A, C). At day 30, expression of ShiDN and Rab3CA, but not that of Rab3DN or *comt* RNAi, significantly lowered the total number of propagated GA puncta compared to GA200-only (Figure 23E, G). When quantifying the relative amount of propagated GA, both Rab3CA and Rab3DN led to the strongest decrease in propagated GA (Figure 23H). Finally, similarly to our exosomal screen, mCherry co-expression increased propagated GA at day 30 (Figure 23E, G, H).

While these preliminary results should be considered with caution, these experiments indicate that Rab3 is a relevant modulator of GA spread from ORNs and that GA may partially accumulate in synaptic vesicles within ORN terminals. Therefore, the role of synaptic activity in GA trans-cellular spread deserves further investigation.

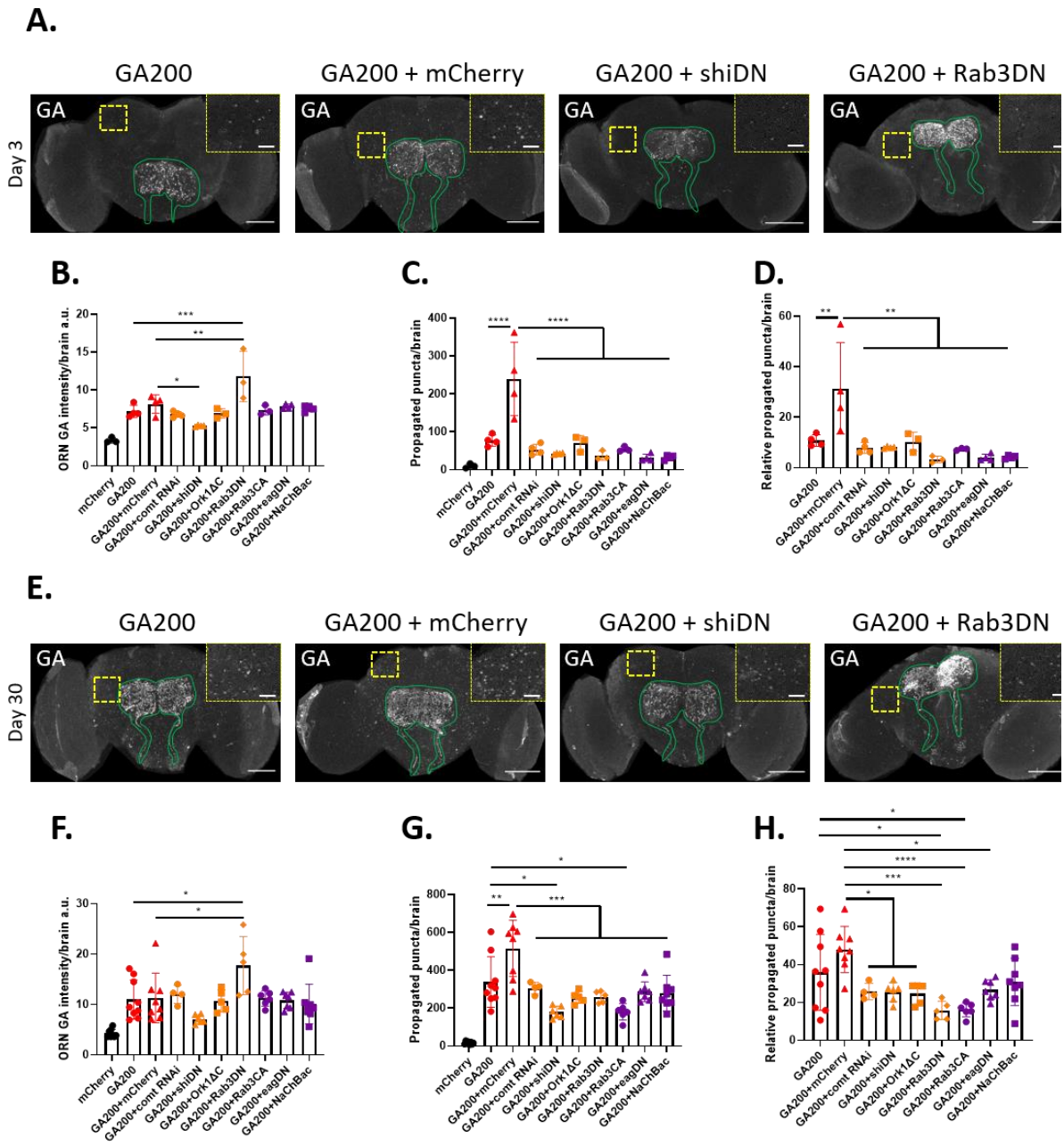

Figure 23. Genetic modulation of genes involved in synaptic transmission affects GA200 spread from ORNs. **A, E** Representative images of brains from flies expressing GA200, GA200 and mCherry, GA200 and ShiDN, and GA200 and Rab3DN in ORNs under the ORN-specific orco-Gal4 driver for 3 **(A)** or 30 **(E)** days. Brains were probed with an anti-GA antibody (white) and the same laser power settings were used for imaging of all conditions. The outline of ORN axons and synaptic terminals is shown in green. Mild unspecific binding to trachea and the lamina of the optic lobes was observed. Insets of indicated areas highlight differences in the number of propagated dots across conditions. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B, F** Quantification of the mean GA intensity in ORNs after 3 **(B)** or 30 **(F)** days of expression of the indicated constructs. mCherry-only was used to control for unspecific binding of the anti-GA antibody, and GA200+mCherry was used as a genetic control. Reference genotypes are shown in red, genotypes expressing a repressor of neuronal activity are shown in orange, and genotypes expressing an enhancer of neuronal activity are shown in purple. Only differences between the reference genotypes and flies expressing a modulator of neuronal activity are shown (***P<0.001, **P<0.01, *P<0.05; One-Way ANOVA with Tukey's multiple comparisons test, n=3-9 brains). **C, G** Quantification of the total number of propagated GA puncta outside of ORNs after 3 **(C)** or 30 **(G)** days of expression of the indicated constructs (****P<0.0001, ***P<0.001, **P<0.01, *P<0.05; One-Way ANOVA with Tukey's multiple comparisons test, n=3-9 brains). **D, H** Quantification of the relative amount of propagated GA after 3 **(D)** or 30 **(H)** days of expression of the indicated constructs (****P<0.0001, ***P<0.001, **P<0.01, *P<0.05; One-Way ANOVA with Tukey's multiple comparisons test, n=3-9 brains).

3.2 EFFECT OF TAGS ON GA TOXICITY, AGGREGATION PATTERN AND CELLULAR RESPONSES IN THE FLY BRAIN

3.2 Effect of tags on GA toxicity, aggregation pattern and cellular responses in the fly brain

A key question in the *C9orf72* mutation field is which of the five DPRs causes the greatest toxicity in ALS/FTD patients and would therefore afford the most promising target for clinical benefit. In humans, only postmortem analyses have been possible as of now, and these have been hindered by difficulties in the generation of antibodies against DPRs, presumably due to their low immunogenicity (Zhou et al. 2020). However, various antibodies are currently available and these have overall revealed an apparent discrepancy between the burden of DPR aggregates, particularly abundant in the cerebellum and hippocampus, and the neurodegeneration severity, which is greatest in the cortex and spinal cord (I. R. A. Mackenzie et al. 2015). Since available antibodies only detect insoluble deposits in postmortem examination, insoluble DPR aggregates may either not represent the most toxic species or only deposits with specific morphology, composition or subcellular location may cause toxicity, which remains largely unexplored (Freibaum and Taylor 2017).

When compared head-to-head, most studies have concluded that the arginine-rich DPRs cause more toxicity than GA, which may be only mildly toxic (Mizielinska et al. 2014). However, to study DPR toxicity in cellular or animal models, most laboratories have used genetic engineering to develop constructs consisting of a codon-optimized sequence specific for each DPR, preceded or followed by one of the many commonly used tags, including FLAG (May et al. 2014; Yang et al. 2015), GFP (May et al. 2014; Wen et al. 2014) and mCherry (Darling et al. 2019; K. H. Lee et al. 2016). While epitope tagging can be very useful to detect and characterize proteins for which no antibodies are available, it can also add undesired artifacts. Worrisomely, most preclinical studies have not included an untagged control in their investigations. Therefore, I explored whether GA toxicity and cellular responses may have been underestimated due to the use of tags.

3.2.1 Tags reduce GA100-mediated lifespan shortening upon pan-neuronal expression

To address whether commonly used tags affect GA toxicity *in vivo*, we generated novel fly lines that express either untagged or C-terminally GFP-, mCherry-, or FLAG-tagged, codon-optimized GA with 100 repeats (hereafter GA100, GA100GFP, GA100mCherry and

GA100FLAG) from a UAS transgene. Given the large size of the GFP and mCherry tags, and similarly to our previous DPR-mCherry constructs, I placed a spacer sequence consisting of five Gly-Ser dipeptides between the GA100 sequence and the fluorophores. All transgenes were integrated into the same genomic locus, the attP2 landing site, to ensure equal transcriptional expression.

Expression of GA100 from the attP40 landing site in adult female flies, driven in neurons by the inducible elavGS driver, shortens fly lifespan (Mizielinska et al. 2014). I therefore measured the response of lifespan to induction of expression of the untagged and tagged GA100 driven by elavGS in adult female flies. Controls were fed vehicle-only (EtOH). All fly lines had been backcrossed into the same genetic background before conducting experiments. In agreement with the previous study, untagged GA100 strongly shortened fly lifespan (Figure 24B). In contrast, lifespan was mildly extended upon expression of GA100GFP or GA100mCherry (Figure 24D, F), while GA100FLAG expression led to a modest, but significant, reduction (Figure 24G). The lifespan extension observed upon GA100GFP and GA100mCherry expression by RU treatment was also seen in flies induced to express GFP-only (Figure 24C) or mCherry-only (Figure 24E), as well as driver-only flies treated with RU (Figure 24A). RU treatment to female flies has previously been reported to extend lifespan (Tower et al. 2017), which may explain the mild lifespan extension in driver-only, GFP-only, GA100GFP, mCherry-only and GA100mCherry flies upon RU treatment. Of note, the differential effects of the GA-tagged and -untagged constructs on lifespan were not due to genetic differences unrelated to the presence or expression of the transgenes, because no lifespan changes were found across non-induced genotypes carrying a UAS transgene (Figure 24H). When directly compared to each other, untagged GA100-expressing flies were shorter lived than any of the tagged lines, and GA100FLAG-expressing flies exhibited a significantly shorter lifespan than those expressing GA100GFP or GA100mCherry (Figure 24I). Therefore, commonly used tags reduced the toxic effects of GA on fly lifespan, and the longer tags completely abolished GA-mediated lifespan shortening. These results show that GA toxicity is strongly influenced by commonly used tags *in vivo*.

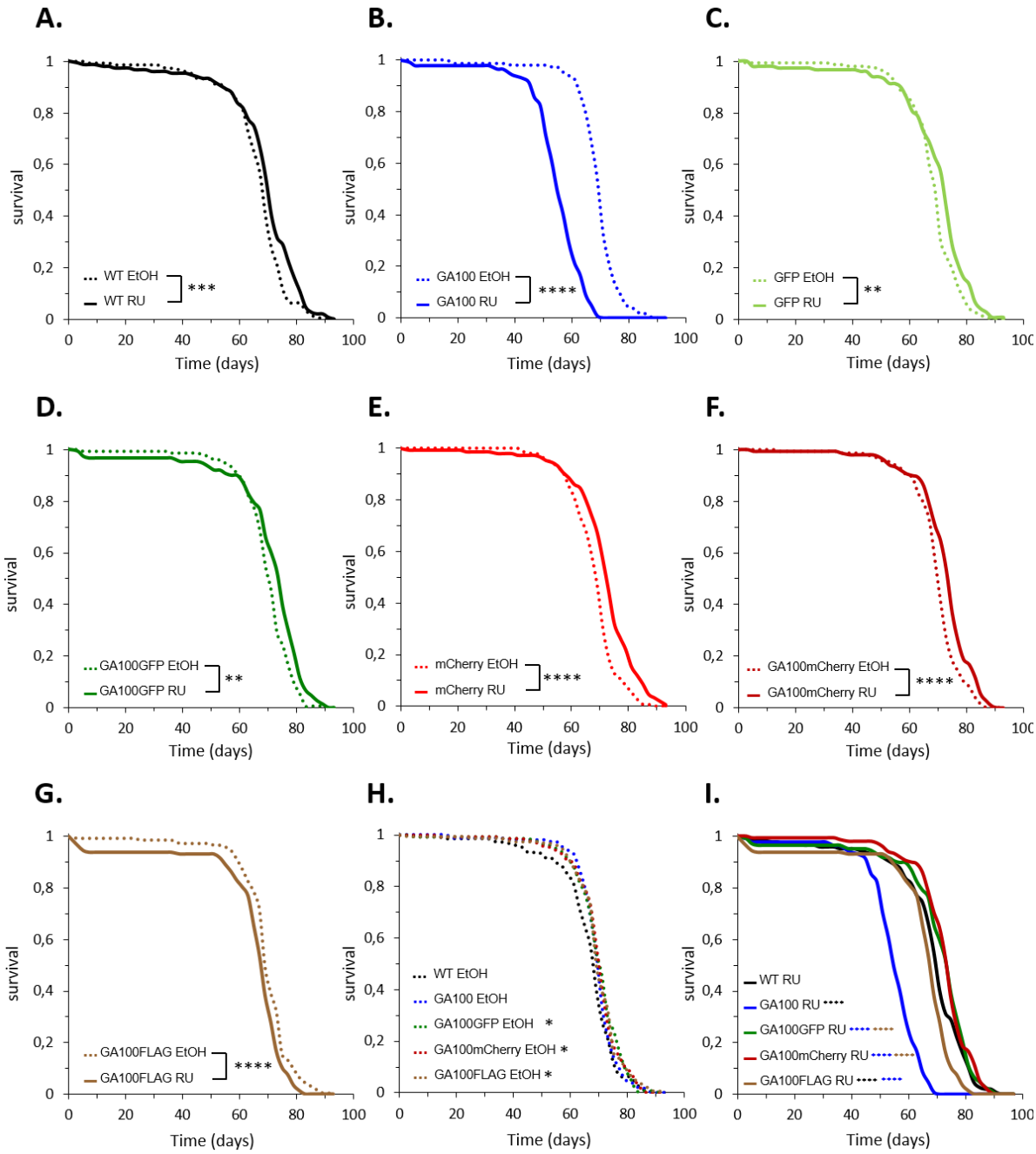

Figure 24. GA100 toxicity is not recapitulated by tagged GA100.

A-G Lifespan curves of flies non-induced (EtOH) or induced (RU) to express the indicated UAS transgenes under the pan-neuronal driver elavGS. Only expression of GA100 **(B)** and GA100FLAG **(G)** shortened fly lifespan (****P<0.0001, ***P<0.001, **P<0.01; log-rank test, n=150 female flies). **H** Comparison of the lifespan curves of all of the non-induced genotypes. The same EtOH lifespan curves shown in A-G are plotted together. Carrying one copy of any of the UAS transgenes did not shorten lifespan and no differences were observed among UAS lines. Black asterisks refer to the statistical comparison of the lifespan curve of each genotype to that of driver-only flies (i.e., WT) (*P<0.05; log-rank test, n=150 female flies). **I** Comparison of the lifespan curves of all of the induced genotypes. The same RU lifespan curves shown in A-G are plotted together. Expression of GA100 was more toxic than that of any of its tagged counterparts, and expression of GA100FLAG was more toxic than that of GA100GFP or GA100mCherry, which were not different from the driver-only. Black, blue and brown asterisks refer to the statistical comparison of the lifespan curve of each genotype to that of driver-only, GA100 or GA100FLAG flies, respectively (****P<0.0001; log-rank test, n=150 female flies).

3.2.2 Large tags affect GA aggregation profile without interfering with puncta formation in the fly brain

The large toxicity differences between tagged and untagged GA indicated that the tags altered the molecular properties of GA DPRs. To probe the basis of this effect, I first extracted proteins from heads of flies induced to express each construct for 24 hours pan-neuronally, and ran them in a gel to assess aggregation. Despite using reducing conditions and in agreement with our previous results (Figure 13B), I found that GA100mCherry and GA100GFP formed bands reminiscent of oligomeric peptides, which were not observed for either untagged GA100 or GA100FLAG (Figure 25A). In addition, the overall GA intensity in GA100mCherry and GA100GFP samples was stronger than in the other samples (Figure 25B). However, these apparent differences in concentration could be influenced by solubility differences among the proteins that would favour the transfer or the migration of the less aggregated GA species. To compare the overall accumulation of each construct more reliably, I subjected the same protein extracts to dot blot and probed the extracts with an anti-GA antibody. GA100GFP was the most abundant species, and no significant differences were found between untagged GA100, GA100mCherry and GA100FLAG after 24 hours of expression induction (Figure 25C). These results indicate that the presence of large tags at the C-terminus of GA100 modifies its aggregation profile and C-terminal GFP makes it more stable.

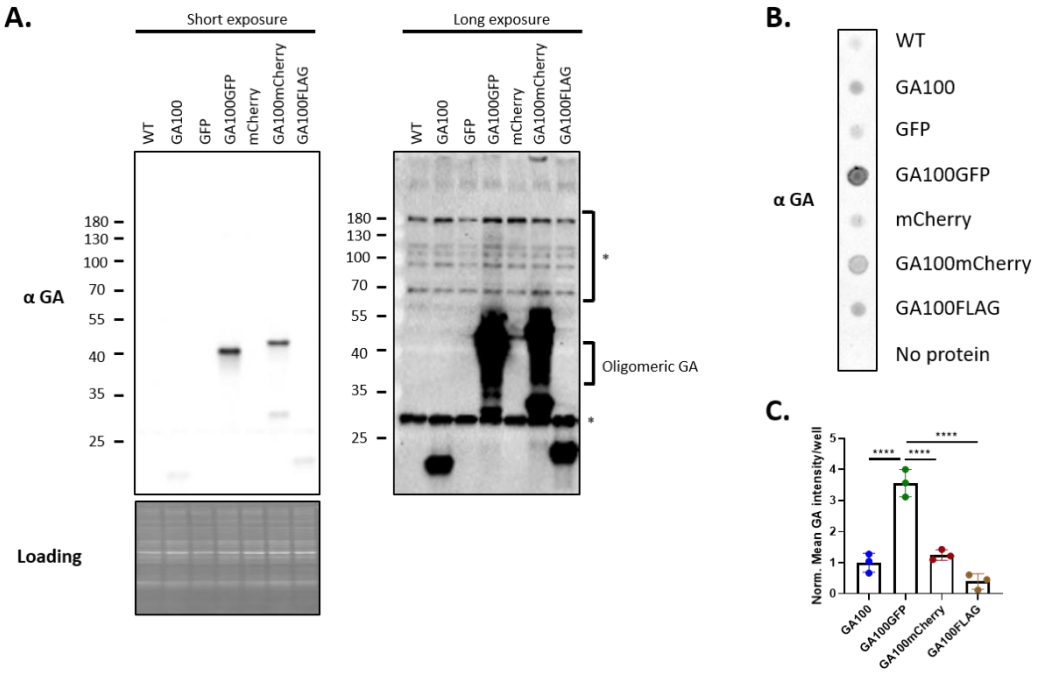

Figure 25. Large tags affect the aggregation pattern of GA100.
(Below)

Figure 25. Large tags affect the aggregation pattern of GA100.
A Immunoblotting of head protein extracts from flies induced to express each of the indicated constructs under the pan-neuronal driver elavGS for 24 hours. GA100GFP and GA100mCherry formed oligomeric bands not observed upon GA100 or GA100FLAG expression. Extracts were probed with anti-GA antibody. Equal protein amounts were loaded. Each lane contained extracts from 20 fly heads. * = unspecific bands. **B** Representative dot blot results using the same protein extracts after probing with an anti-GA antibody. Flies expressing the driver-only (i.e., WT), GFP or mCherry were used to control for unspecific binding of the antibody to non-GA DPR epitopes. No-protein wells were also included to account for signal from the membrane and dot blot procedure. **C** Quantification of the mean GA intensity per well of the dot blot experiment in B. Each dot represents the mean of two technical replicates of the same biological sample in different positions of the dot blot membrane. The mean GA intensity signal from WT, GFP and mCherry was substracted from that of the GA-expressing conditions, which were subsequently normalized to GA100 (****P<0.0001; One-Way ANOVA with Tukey's multiple comparisons test, n=3 sets of 20 fly heads).

GA DPRs form cytoplasmic puncta in human postmortem tissue (Mori, Arzberger, et al. 2013), and these have also been detected by expressing tagged GA and using tag-specific antibodies or fluorescence in model systems (Wen et al. 2014; Yang et al. 2015). To further investigate if tags could affect the types of aggregates formed by GA, I performed brain stainings of flies induced to express each construct pan-neuronally for 24 hours with elavGS and visualized GA using an anti-GA antibody, as well as the endogenous GFP or mCherry fluorescence. In line with previous reports, flies expressing both untagged GA100 and GA100FLAG exhibited abundant puncta-shaped GA-positive aggregates across the whole brain (Figure 26A). Interestingly, while GA100GFP and GA100mCherry also formed abundant puncta detectable through the fluorescence of the GFP and mCherry tags, respectively, these were not labelled by the same anti-GA antibody (Figure 26A, B). In fact, the anti-GA antibody exclusively detected diffuse cytoplasmic GA signal in GA100GFP and GA100mCherry brains (Figure 26B), suggesting that, unlike for untagged GA and GA100FLAG aggregates, the conformation of the GA100GFP and GA100mCherry puncta did not allow the binding or penetration of the anti-GA antibody. Importantly, GFP-only and mCherry-only did not form puncta (Figure 26A), indicating that the GFP-positive and mCherry-positive spots observed for GA100GFP and GA100mCherry flies also comprised GA, which in turn triggered their aggregation.

Collectively, our results indicated that, while large tags do not prevent GA puncta formation, they interfere with the aggregation profile and type of aggregates formed by GA.

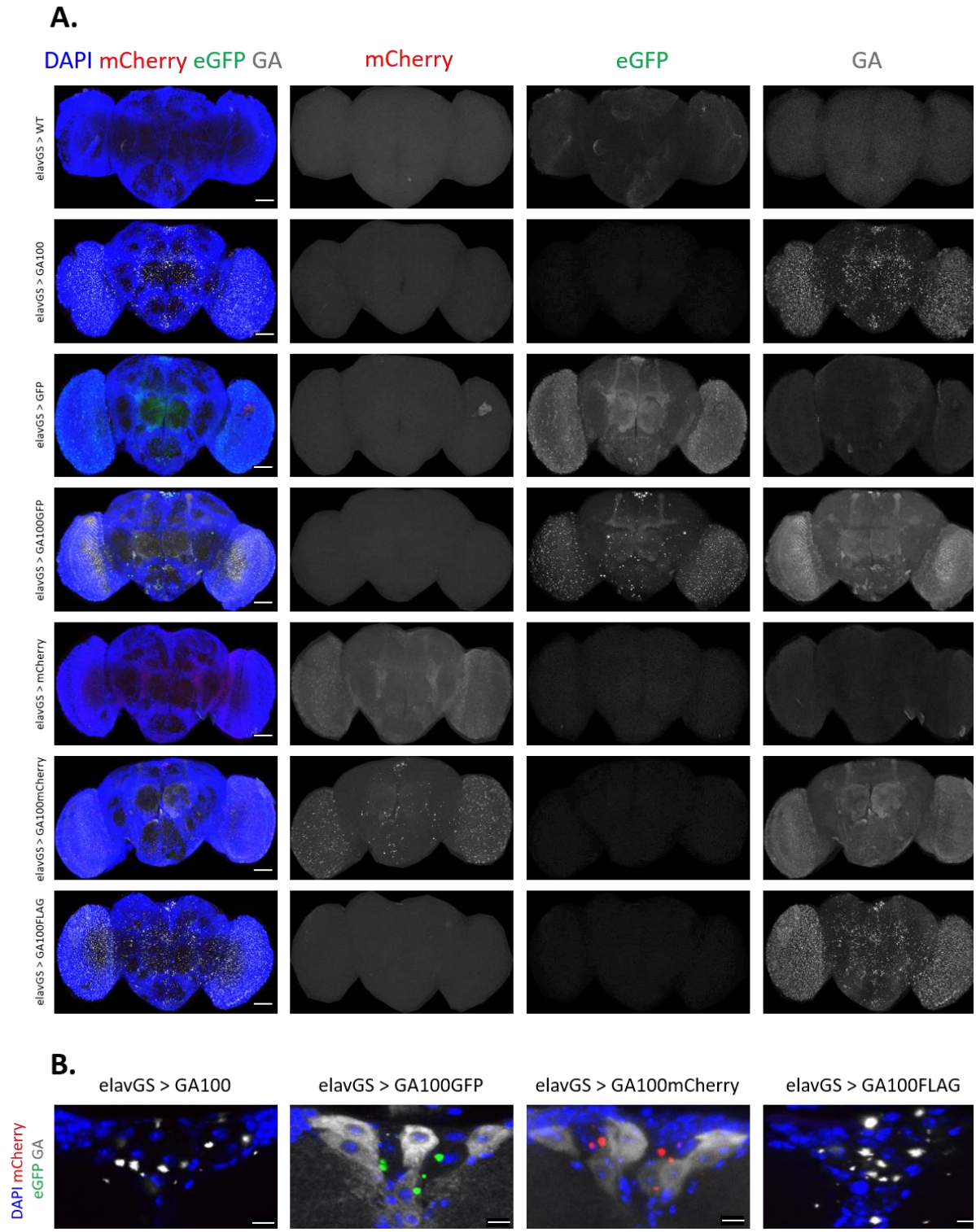

Figure 26. GA100GFP and GA100mCherry form puncta that are not recognized by an anti-GA antibody.
A Representative images of adult fly brains expressing the indicated constructs under the pan-neuronal driver elavGS for 24 hours. Brains were probed with an anti-GA antibody, and their mCherry (red), GFP (green) and GA (gray) signals were imaged using the same laser power settings across conditions. GA100 and GA100FLAG formed puncta recognized by an anti-GA antibody, while GA100GFP and GA100mCherry formed puncta detected from their GFP and mCherry signals, respectively, but not by an anti-GA antibody. Scale bar in images is 100 µm. **B** Representative images of the MNC region of the indicated genotypes from the brains shown in A. The GA antibody specifically detected puncta upon GA100 and GA100FLAG expression, while in GA100GFP and GA100mCherry flies it detected diffuse GA. Scale bar in images is 5 µm.

3.2.3 Large tags interfere with cellular consequences of GA expression in the fly brain

I next investigated whether brain cells would react differentially to untagged or tagged GA. Several studies have expressed GA DPRs in model systems and investigated the cellular responses. One common observation is that GA DPRs lead to increased levels of the p62 in mammalian systems (May et al. 2014; Nguyen et al. 2019; Y. J. Zhang et al. 2016), which may be an indication of cells defending themselves against toxicity but failing to clear the DPR aggregates.

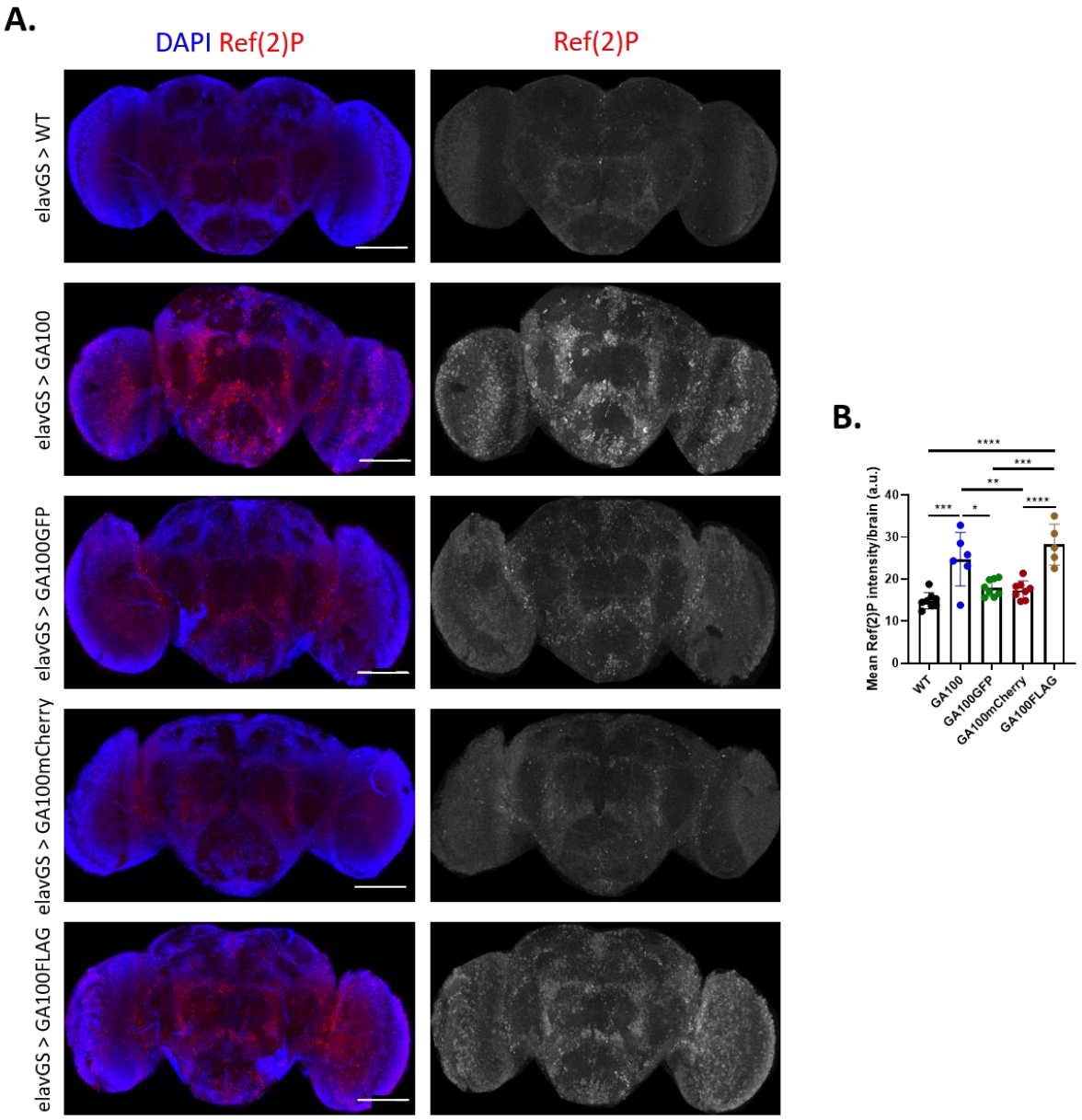

Figure 27. Brain cells do not up-regulate Ref(2)P upon long-term GA100GFP or GA100mCherry expression.
A Representative images of adult fly brains upon expression of the indicated constructs under the pan-neuronal driver elavGS for 25 days. Brains were probed with an anti-Ref(2)P antibody (red) and the same laser power settings were used for imaging. Scale bar in images is 100 µm. **B** Quantification of the mean Ref(2)P intensity signal per brain (****P<0.0001, ***P<0.001, **P<0.01, *P<0.05; One-way ANOVA with Tukey's multiple comparisons test, n=5-8 brains).

I had previously corroborated this GA effect in flies upon pan-neuronal expression of GA200 (Figure 20A). I next investigated whether GA100 affected p62 levels in a tag-dependent manner *in vivo* by inducing the pan-neuronal expression of untagged or tagged GA100 for 25 days with elavGS and performing brain stainings against Ref(2)P. Long-term expression of untagged GA100 or GA100FLAG triggered a strong increase in Ref(2)P levels compared to driver-only flies of the same age (Figure 27A, B). However, this effect was not observed upon GA100GFP and GA100mCherry expression, which led to largely unchanged Ref(2)P levels compared to control flies (Figure 27A, B). This suggests that cells do not require Ref(2)P up-regulation to cope with GA100GFP and GA100mCherry expression, but other interpretations are also possible, as detailed in the discussion section.

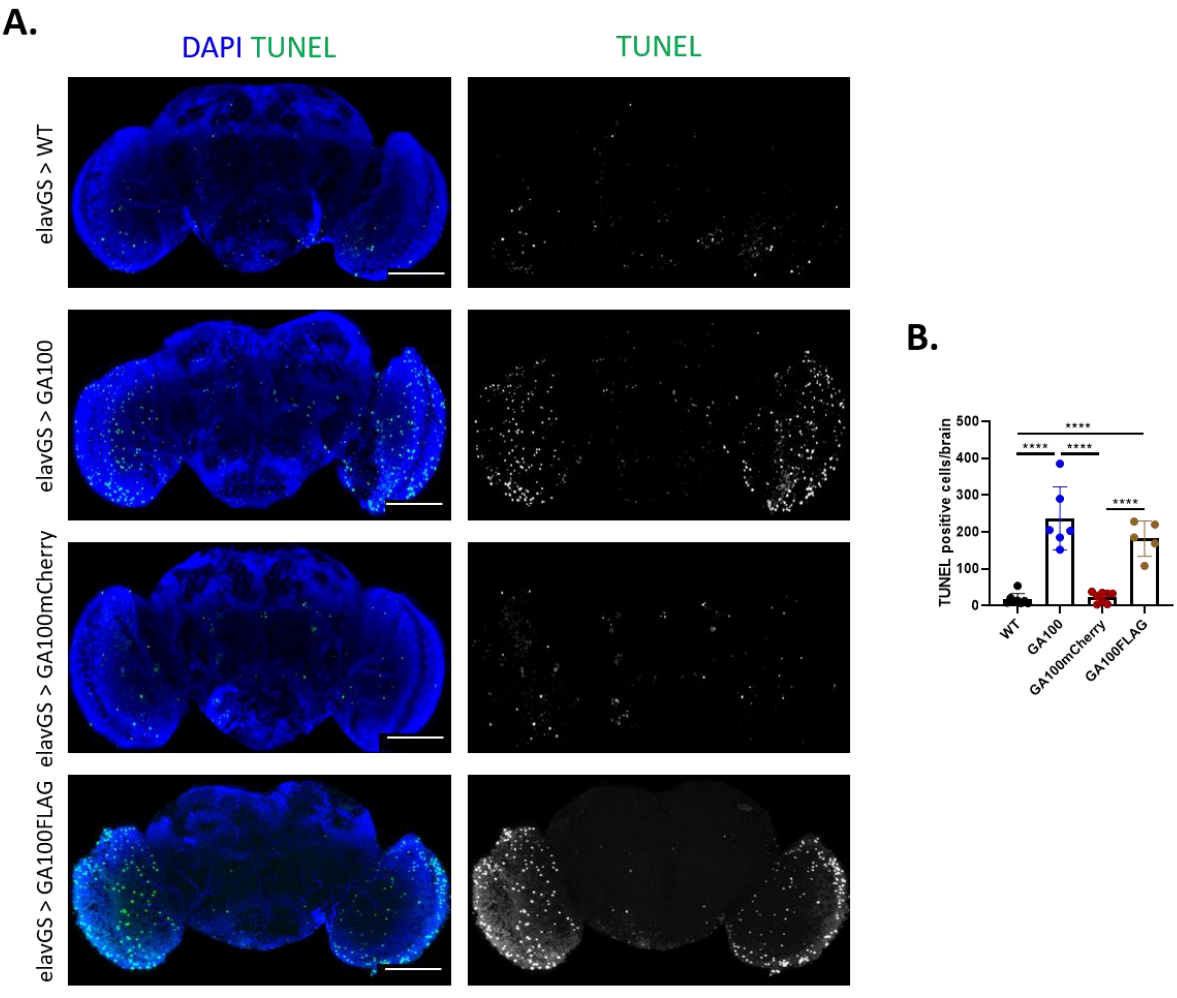

Figure 28. Long-term GA100mCherry expression does not increase DNA damage.
A Representative images of adult fly brains upon expression of the indicated constructs under the pan-neuronal driver elavGS for 25 days. Brains were stained with the TUNEL dye (green) and the same laser power settings were used for imaging. Scale bar in images is 100 µm. **B** Quantification of the number of TUNEL positive cells per brain (****P<0.0001; One-way ANOVA with Tukey's multiple comparisons test, n=5-8 brains).

Another common effect of GA expression in mammalian systems is increased DNA damage (Nihei et al. 2020), which I previously corroborated in flies upon pan-neuronal expression of GA200 (Figure 19A). I therefore assessed whether GA tagging would have an effect on GA-mediated DNA damage. I performed TUNEL labelling of DNA fragmentation in brains of flies induced to pan-neuronally express each construct with elavGS for 25 days and quantified the number of TUNEL-positive cells per brain. Since the TUNEL excitation and emission spectra overlap with those of GFP, I focused on untagged GA100, GA100mCherry and GA100FLAG for these experiments. Control flies showed scarce TUNEL labelling, which was strongly increased by expression of untagged GA100 and GA100FLAG (Figure 28A, B). As noted earlier, DNA fragmentation was primarily observed in the OLs. In contrast, GA100mCherry expression did not increase DNA fragmentation compared to control flies (Figure 28A, B).

Overall, these results show that at least some of the effects caused by untagged GA100 are not recapitulated by GA100 carrying large tags.

3.3 ROLE OF REPEAT LENGTH IN GA TOXICITY, AGGREGATION PATTERN AND CELLULAR RESPONSES IN THE FLY BRAIN

3.3 Role of repeat length in GA toxicity, aggregation pattern and cellular responses in the fly brain

In addition to using tagged constructs, most preclinical studies in the *C9orf72* mutation field have used constructs expressing relatively short repeats, typically ranging from 10 to 100 repeats, to study DPR-mediated toxicity. However, *C9orf72* mutation carriers usually have several 100s-1000s of repeats, whose specific length seems to be both tissue- and age-dependent, with repeats typically expanding over time and being particularly long in some disease-relevant areas, such as the frontal cortex (van Blitterswijk et al. 2013). Albeit unproven, patients are expected to generate DPRs of a similar repeat length to that of their genomic mutation. Given that several studies have shown that RAN translation efficiency increases with longer repeats (Mori, Weng, et al. 2013; Zu et al. 2011) and that DPRs can have different subcellular localizations (Kwon et al. 2014; Yang et al. 2015) or cellular effects (Callister et al. 2016) depending on their repeat length, it is therefore unclear how clinically relevant our current knowledge on DPR-mediated neurotoxicity is.

Given that GA is the most abundant DPR and has predominantly been associated with mild toxicity, I investigated whether GA DPRs cause length-dependent toxicity and molecular responses in the fly brain.

3.3.1 GA400 DPRs are more toxic than GA100 or GA200 DPRs

3.3.1.1 Generation and validation of GA400 fly lines

To characterize the effect of repeat length on GA-mediated toxicity, I used flies expressing GA100, GA200 or GA400 from the same genomic locus (the attP2 landing site) to ensure equal expression levels. To generate GA400 flies, we used our previously described GA200 construct and cloned a pure GA200 stretch between the SmaI and XbaI restriction sites (Figure 29A). I generated and validated three independent lines (#1, #2 and #3) injected with the UAS-GA400 plasmids. First, I isolated head proteins from flies induced to pan-neuronally express GA100, GA200 or GA400 for 7 hours. I detected a single linear band for GA100, a stronger higher band as well as a weaker band close to the GA100 size for GA200 (the latter presumably resulting from aberrant translation or cleavage of GA200), and a completely aggregated smear that did not run through the gel for GA400 lines #2 and #3 (Figure 29B). GA400 line #1 had a similar

molecular weight as GA200, suggesting incorrect insertion of the UAS-GA400 transgene (Figure 29B). While this indicated a clear solubility difference between GA400 lines #2 and #3 and the rest of the lines, I could not determine the correct insertion of the whole GA400-coding sequence, as I did not observe a linear protein band of a molecular weight equivalent to twice the size of GA200 for any of the GA400 constructs.

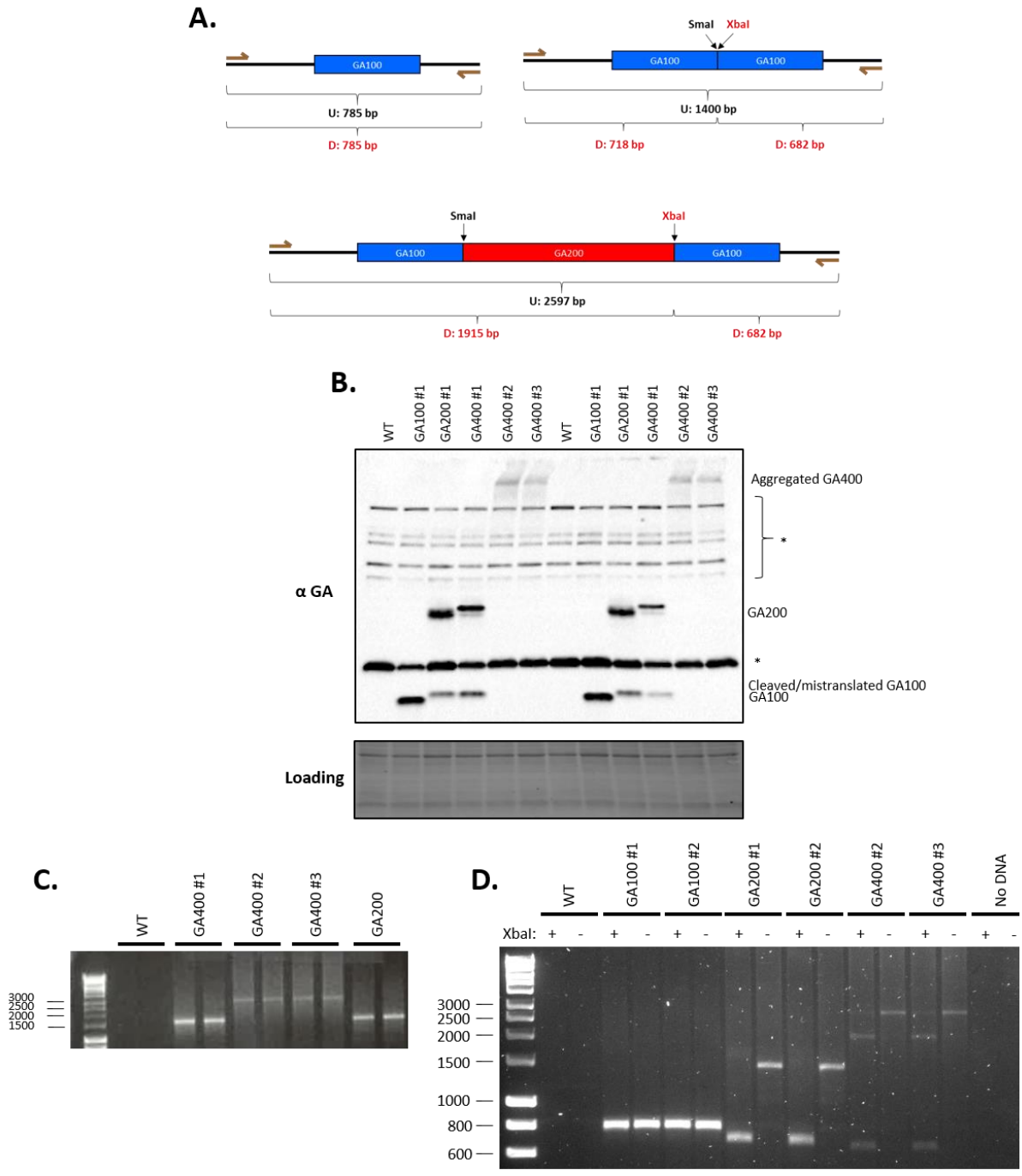

Figure 29. Validation of the GA400 construct.

(Below)

Figure 29. Validation of the GA400 construct.
A Schematic showing the construct sizes of UAS-GA100, UAS-GA200 and UAS-GA400 flies upon PCR amplification of the DNA sequences of their UAS transgenes with (red) or without (black) subsequent digestion with XbaI. **B** Immunoblotting of head protein extracts from flies expressing each of the indicated constructs under the pan-neuronal driver elavGS for 7 hours. Extracts were probed with an anti-GA antibody and equal protein amounts were loaded. Three UAS-GA400 lines from independent injections and back-crossing rounds were analyzed. Extracts from two different biological replicates per genotype are shown. Each lane contains protein extracts from 20 fly heads. * = unspecific bands. **C** Genotyping results after PCR amplification of the UAS transgenes of flies homozygous for the indicated UAS transgenes. Only GA400 lines #2 and #3 showed the expected size for the full GA400 sequence. **D** Genotyping results confirming the construct sizes of two independent lines of each UAS construct upon PCR amplification of the DNA sequences of their UAS transgenes with (+) or without (-) subsequent digestion with XbaI.

To confirm the correct insertion of our constructs, I PCR amplified the UAS transgenes using primers that bind to the flanking sequences of the transgenes. GA400 lines #2 and #3 showed an amplicon of the expected size for GA400, while the GA400 line #1 amplicon was very similar to that of GA200 (Figure 29C), indicating that only GA400 lines #2 and #3 carried the whole-length GA400 construct. To further verify the insertion of the whole GA400 sequence, I performed new PCRs using the same primer pair, and subsequently digested the amplicons with XbaI, which should only cut the GA200 and GA400 sequences, as depicted in Figure 29A. I included two independent lines of GA100 and GA200 in this experiment. Both the digested and the undigested amplicons had the expected sizes for both GA100 and GA200 lines, as well as for GA400 line #2 and line #3 (Figure 29D), thus confirming the correct insertion of the UAS-GA400 plasmids in these two lines. Therefore, I used GA100 line #1, GA200 line #1 and GA400 line #2 for subsequent experiments.

3.3.1.2 GA400 DPRs are more toxic than GA100 and GA200 DPRs upon expression in neuronal tissue

Before comparing the toxicity mediated by each of our GA constructs, I backcrossed all of the lines into the same background for 9 generations. Then, I generated flies with low (18°C) or high (25°C) eye-specific expression of each construct under the constitutive GMR-Gal4 driver and imaged their eye morphology. To validate our assay, I included GR100 in our experiments, as previous studies have shown that eye-specific expression of GR causes a strong rough eye phenotype (Mizielinska et al. 2014; Wen et al. 2014). I corroborated these findings in our experimental set-up (Figure 30A, C). Eye-specific expression of GA100 or GA200 did not cause any overt morphology abnormalities either at low (Figure 30A) or high (Figure 30C) expression levels. In contrast, expression of GA400 at low levels led to a consistent disarray and fusion of eye ommatidia, which is a sign of toxicity (Figure 30A). This phenotype was further

exacerbated at high expression levels, as flies expressing GA400 at 25°C showed a strong rough eye phenotype (Figure 30C).

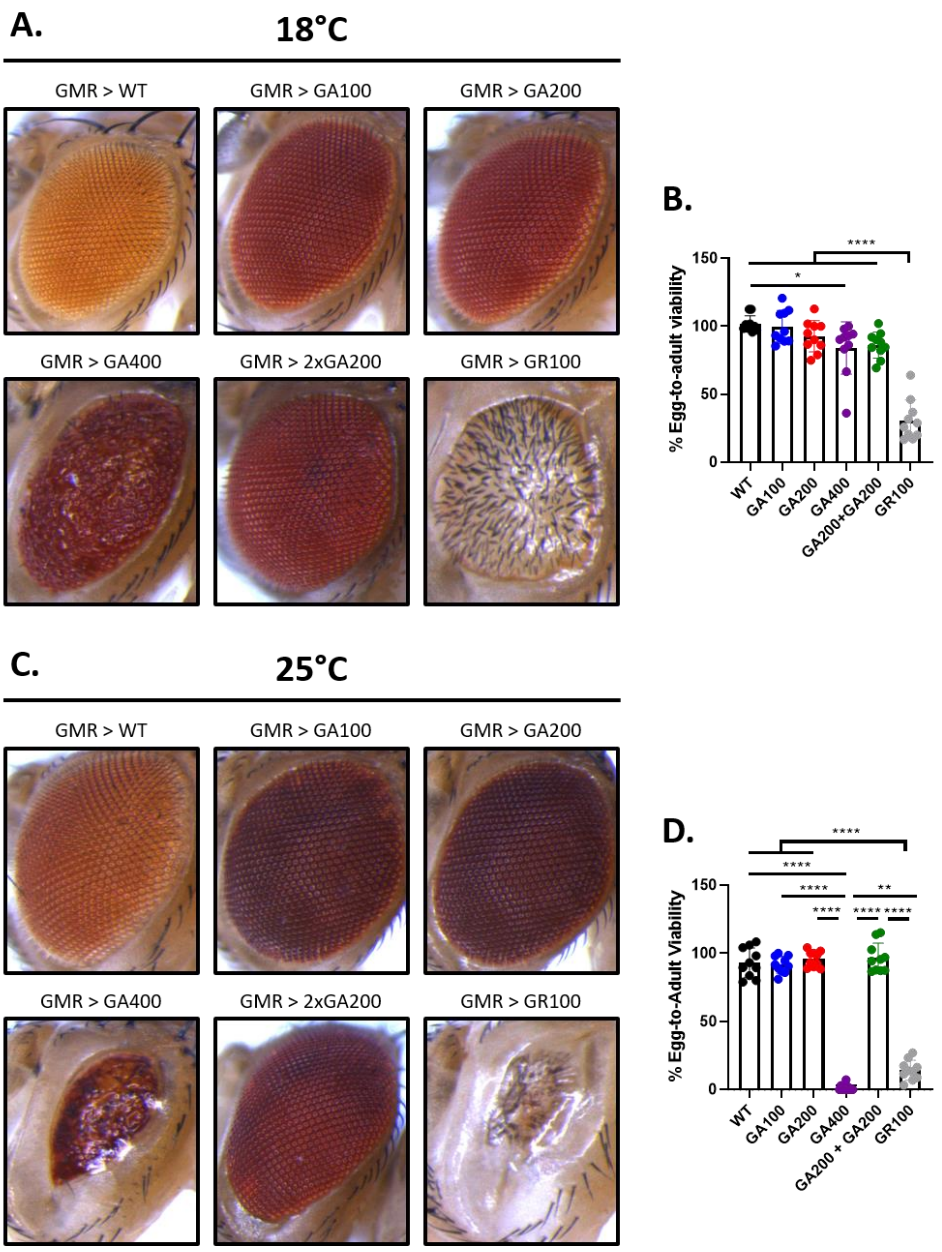

Figure 30. Eye-specific expression of GA400 disrupts eye morphology.
A Representative eye images of 2-day-old flies expressing the indicated constructs under the eye-specific constitutive driver GMR-Gal4 after developing at 18°C. Mild morphology abnormalities were observed in the eyes of GA400-expressing flies, which were not recapitulated by shorter GA DPR constructs or two copies of the UAS-GA200 transgenes (i.e., 2xGA200). GR100 was used as a positive control of toxicity. **B** Developmental assay results showing the percentage of eggs expressing the indicated constructs that reached adulthood after developing at 18°C. Each dot represents one independent vial (****P<0.0001, *P<0.05; One-way ANOVA with Tukey's multiple comparisons test, n=10 independent vials). **C** Representative eye images of 2-day-old flies expressing the indicated constructs under GMR-Gal4 after developing at 25°C. Severe morphology abnormalities were observed in the eyes of GA400-expressing flies, which were not recapitulated by shorter GA DPR constructs or 2xGA200. GR100 was used as a positive control of toxicity. **D** Developmental assay results showing the percentage of eggs expressing the indicated constructs that reached adulthood after developing at 25°C. Each dot represents one independent vial (****P<0.0001, **P<0.01; One-way ANOVA with Tukey's multiple comparisons test, n=10 independent vials).

To further confirm these toxicity differences, I performed a developmental assay where the percentage of eggs that developed to adults was recorded. A previous study had shown that GR100 impairs development in a dose-dependent manner (Mizielinska et al. 2014), which I corroborated in our analysis (Figure 30B, D). Neither GA100 nor GA200 impaired development upon eye-specific expression at low or high levels compared to driver-only flies (Figure 30B, D). However, GA400 expression at low levels mildly decreased survival (Figure 30B), whereas at high levels it dramatically impaired development (Figure 30D). GA400 could cause toxicity by simply interacting with more of the same proteins that GA100 and GA200 interact with due to GA400 having more GA epitopes, and/or by interacting with a different set of proteins from its shorter counterparts due to a different molecular structure. To distinguish between these two mechanisms, I generated flies expressing two copies of GA200 (2xGA200), one from the attP40 site and another one from the attP2 site. Interestingly, 2xGA200 did not cause a rough eye phenotype or impaired development at low or high expression levels (Figure 30A-D), suggesting that GA400 toxicity is not simply explained by comprising more GA epitopes and likely interferes with distinct mechanisms compared to those affected by GA100 and GA200.

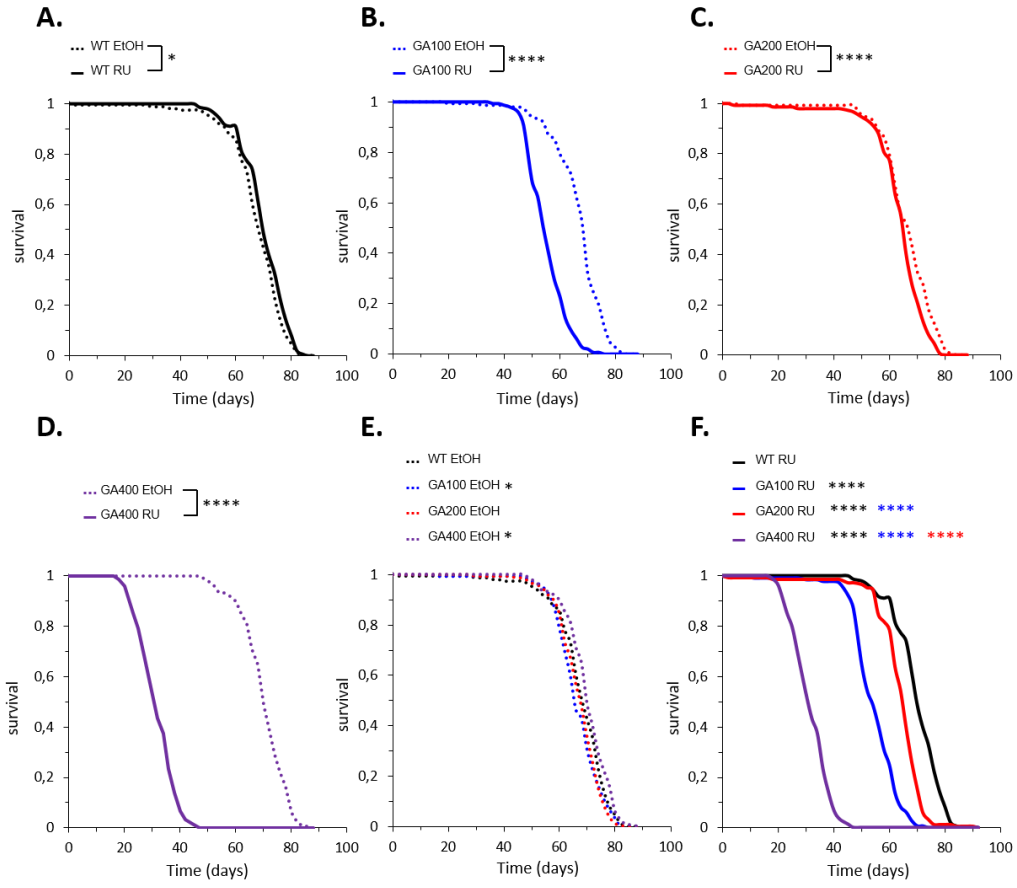

Figure 31. GA400 shortens lifespan more than GA100 and GA200.

(Below)

Figure 31. GA400 shortens lifespan more than GA100 and GA200.
A-D Lifespan curves of flies non-induced (EtOH) or induced (RU) to express the indicated UAS transgenes under the pan-neuronal driver elavGS. Expression of all of the GA constructs significantly shortened fly lifespan (****P<0.0001, *P<0.05; log-rank test, n=150 female flies). **E** Comparison of the lifespan curves of all of the non-induced genotypes. The same EtOH lifespan curves shown in A-D are plotted together. No differences were found among UAS lines. Black asterisks refer to the statistical comparison of the lifespan curve of each genotype to that of driver-only flies (i.e., WT) (*P<0.05; log-rank test, n=150 female flies). **F** Comparison of the lifespan curves of all of the induced genotypes. The same RU lifespan curves shown in A-D are plotted together. Expression of GA400 was more toxic than that of GA100, which was in turn more toxic than GA200. GA200 was also mildly toxic. Black, blue and red asterisks refer to the statistical comparison of the lifespan curve of each genotype to that of driver-only, GA100 or GA200 flies, respectively (****P<0.0001; log-rank test, n=150 female flies).

To further assess the effect of repeat length on GA-mediated toxicity, I generated flies expressing each construct pan-neuronally in the adult stage with elavGS. Induction of all GA constructs by RU feeding shortened fly lifespan compared to their uninduced isogenic controls (Figure 31B-D), which was not observed in driver-only flies (Figure 31A) in agreement with our previous results (Figure 24A). Flies expressing GA400 had a significantly shorter lifespan than those expressing GA100 or GA200 (Figure 31F). This was not due to major genetic differences other than transgene expression, as uninduced GA100, GA200 and GA400 flies showed a very similar lifespan curve (Figure 31E). Surprisingly, pan-neuronal GA100 expression shortened lifespan to a larger degree than that of GA200 (Figure 31F), further suggesting that GA DPRs cause toxicity in a repeat length-dependent manner that cannot simply be accounted for by the number of GA epitopes in each construct.

A common feature shown by fly models of neurodegenerative diseases is the formation of neurodegenerative vacuoles, which are defined as round, dark spots within the neuropil that result from cell death or axonal degeneration (Sunderhaus and Kretzschmar 2016). Therefore, I quantified the number and size of degenerative vacuoles in the brains of flies induced to express each GA construct for 30 days. To detect the vacuoles, I performed a phalloidin staining and focused on a posterior area of the fly brain next to the mushroom body calyx (Figure 32A), where pan-neuronal expression of Aβ had previously been shown to increase vacuole number (Coelho et al. 2018). GA400-expressing brains showed significantly larger and more numerous vacuoles than driver-only flies, which was not observed in brains expressing GA100 or GA200 (Figure 32B, C).

Collectively, our results show that GA400 is more toxic than GA100 and GA200, and this is probably due to GA400 acquiring GOF properties that may enable it to affect distinct pathways compared to its shorter counterparts

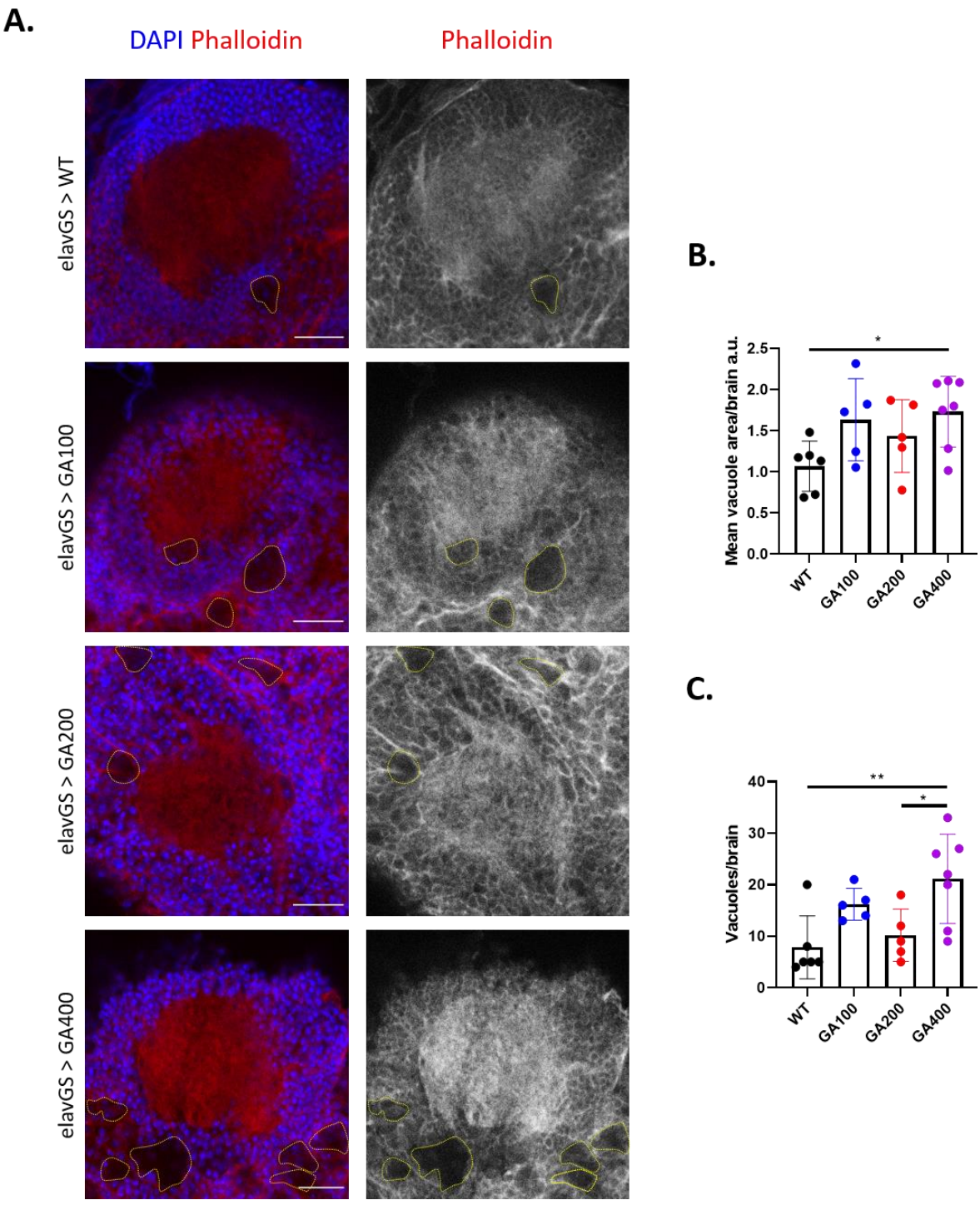

Figure 32. Long term GA400 expression increases neurodegenerative vacuoles.
A Representative images of adult fly brains expressing each of the indicated constructs under the pan-neuronal driver elavGS for 30 days. Brains were stained with the cell membrane-binding phalloidin dye, and a posterior area of the fly brain next to the mushroom body calyx was imaged using the same laser power intensities. Scale bar in images is 20 µm. **B** Quantification of the mean vacuole area upon expression of the indicated constructs. Each dot represents the mean of the two mushroom body calyces of the same fly brain (*P<0.05; One-way ANOVA with Tukey's multiple comparisons test, n=5-7 brains). **C** Quantification of the mean vacuole number upon expression of the indicated constructs. Each dot represents the mean of the two mushroom body calyces of the same fly brain. Vacuole identification was performed in a blinded manner (**P<0.01, *P<0.05; One-way ANOVA with Tukey's multiple comparisons test, n=5-7 brains).

3.3.2 The subcellular location and aggregation propensity of GA DPRs is repeat-length-dependent

I next tested whether the large toxicity differences among our GA constructs could be related to repeat length-dependent changes in subcellular location and aggregation propensity. To this end, I first stained the brains of flies induced to express each construct pan-neuronally for 5, 15 or 30 days with an anti-GA antibody and examined the subcellular location of the GA DPRs. At 5 days of expression, both GA100 and GA200 exhibited a rather punctated pattern both in the somata and axons of neurons all across the brain (Figure 33A), which was in agreement with our previous results showing that both GA100 and GA200 are transported along the axons of ORNs upon ORN-specific expression (Figure 15A). However, GA400 had a rather diffuse pattern and was mostly detected in the somata of neurons all across the brain (Figure 33A).

In addition, while the location of GA400 did not change during ageing, I observed increased cytoplasmic signal in flies expressing GA100 or GA200 at older ages, although the majority of the GA signal was still mainly punctated in these genotypes (Figure 33A). To further confirm that GA400 is not transported along axons and to test whether this is due to the increased number of GA epitopes, I generated flies induced to express GA200, 2xGA200 or GA400 in ORNs for 33 days and stained their brains with an anti-GA antibody. Even after long-term expression induction, no GA signal was detected in the axonal terminals of ORNs or outside of this region in GA400-expressing flies (Figure 34A-C). However, compared to expression of one copy of the GA200 transgene, expression of 2xGA200 further increased both GA accumulation in the synaptic terminals of ORNs and GA spread out of this neuronal population (Figure 34A-C). The increased spread of 2xGA200 compared to GA200 remained significant after normalizing it to GA expression (Figure 34D), suggesting that the mechanisms of GA200 spread can be affected in a synergistic manner by GA200 molecules. Overall, these results indicate that the subcellular location of GA DPRs changes in a repeat length-dependent fashion independent of GA expression levels.

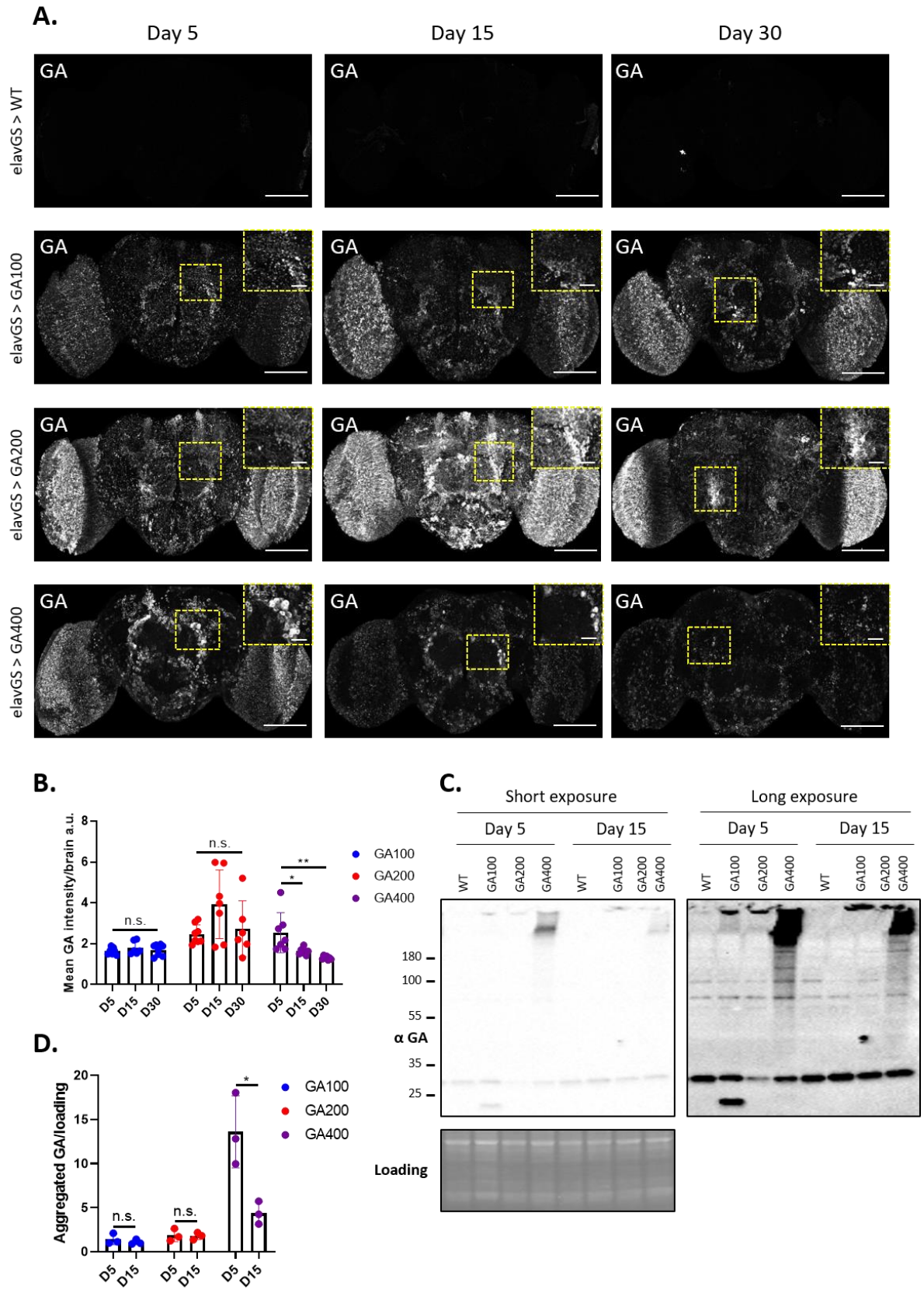

Figure 33. GA subcellular location and aggregation propensity are repeat length-dependent.

(Below)

Figure 33. GA subcellular location and aggregation propensity are repeat length-dependent.
A Representative images of adult fly brains upon expression of the indicated constructs under the pan-neuronal driver elavGS for 5, 15 or 30 days. Brains were probed with an anti-GA antibody (white) and the same laser power settings were used for imaging. GA400 signal was mostly diffuse and located in somata, while that of GA100 and GA200 looked rather punctated and was located both in somata and axons. Scale bar in images and insets are 100 µm and 10 µm, respectively. **B** Quantification of the mean GA intensity signal per brain, genotype and age. Brains of the same genotype were compared to each other. GA signal did not change at older ages in GA100- or in GA200-expressing flies, but it decreased upon GA400 expression **P<0.01, *P<0.05, n.s. = not significant; One-way ANOVA with Tukey's multiple comparisons test, n=6-8 brains). **C** Immunoblotting of head protein extracts from flies expressing each of the indicated constructs for 5 or 15 days under elavGS. Extracts were probed with an anti-GA antibody, and equal amounts of protein were loaded per condition. GA propensity to aggregate was directly proportional to its repeat length. **D** Quantification of the amount of aggregated GA at the top of the membrane normalized to protein loading. Aggregated GA is lower in 15-days-old than in 5-days-old flies expressing GA400 (*P<0.05; t-test, n=3 sets of 20 fly heads).

While GA100 and GA200 protein levels did not change over time, I detected progressively less GA400 signal at older ages (Figure 33B). I confirmed these expression differences by performing western blotting on heads of flies of the same genotypes and induced to express each construct for 5 or 15 days (Figure 33C, D). These time-dependent differences may result from the progressive loss of neurons in GA400-expressing flies, which would be in agreement with the increased detection of neurodegenerative vacuoles after long-term expression of GA400 (Figure 32C). However, I cannot rule out that GA400 expression leads to changes in feeding behaviour after long-term expression, which would lead to lower transcript levels too. Regarding the effect of GA repeat length on aggregation propensity, I found that GA400 was completely insoluble and did not run through the gel either after 7 hours (Figure 29B), 5 days or 15 days of expression induction (Figure 33C). In contrast, GA100 exhibited a non-aggregated linear band both after 7 hours (Figure 29B) and 5 days of expression (Figure 33C), but became fully aggregated at 15 days (Figure 33C). Finally, while GA200 showed non-aggregated linear bands after 7 hours (Figure 29B), it was fully aggregated after 5 and 15 days of expression (Figure 33C). These results indicate that the aggregation propensity of GA DPRs increases with longer repeats.

Collectively, these results show that repeat length affects both the subcellular location and aggregation propensity of GA DPRs. Since GA400 caused toxicity earlier than GA100 and GA200, the initial differences in the location of GA DPRs could play a relevant role in determining their overall toxicity. Furthermore, given that a previous study reported that GA toxicity is completely dependent on its ability to aggregate (Y. J. Zhang et al. 2016), the effect of repeat length on toxicity could at least be partially correlated with its effect on aggregation.

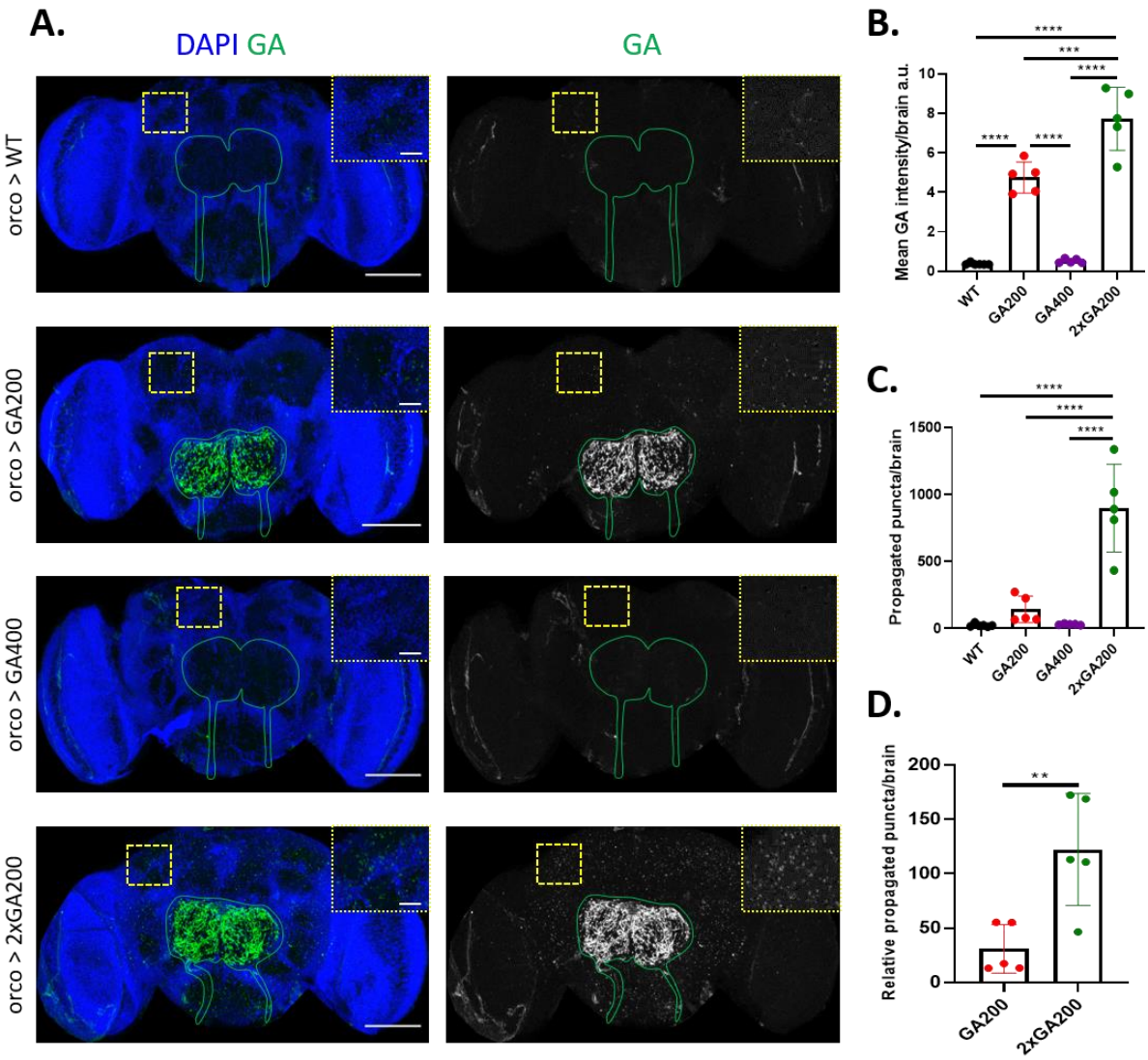

Figure 34. GA400 does not spread from ORNs.
A Representative adult brains from flies induced to express the indicated constructs under the ORN-specific orco-Gal4 driver for 33 days. Brains were probed with an anti-GA antibody (green) and the same laser power settings were used across conditions. The boundaries of the ORN axons and synaptic terminals are highlighted with a solid green line. Insets of the indicated areas are shown to facilitate visualization. Scale bars in images and insets are 100 µm and 10 µm, respectively. **B** Quantification of the mean GA signal in ORN synaptic terminals. Two copies of GA200 lead to greater GA accumulation than one copy thereof. GA400 is not detected in ORN terminals (****P<0.0001, ***P<0.001; One-way ANOVA with Tukey's multiple comparisons test, n=5-6 brains). **C** Quantification of the total number of propagated puncta from ORNs. Two copies of GA200 spread more than one copy thereof. GA400 does not spread from ORNs (****P<0.0001; One-way ANOVA with Tukey's multiple comparisons test, n=5-6 brains). **D** Quantification of the relative amount of propagated GA between flies expressing one or two copies of GA200. Larger accumulation of GA signal in flies expressing two copies of GA200 does not account for their greater spread (**P<0.01; t-test, n=5 brains).

3.3.3 The repeat length of GA DPRs affects their cellular responses in a non-linear manner

3.3.3.1 GA expression modulates Ref(2)P levels in a repeat length-dependent manner

As previously discussed, GA expression has been reported to increase p62 levels in various experimental models (May et al. 2014; Y. J. Zhang et al. 2016), whose clinical relevance is

supported by studies showing GA-p62 colocalization in patient tissue (Schludi et al. 2015). I previously showed that long-term pan-neuronal expression of GA100 in flies leads to increased levels of the p62 fly homolog Ref(2)P, which was not triggered by non-toxic tagged GA100 constructs (Figure 27A, B). To test whether this compensatory response to fight toxicity is influenced by GA repeat length, I stained brains with an anti-Ref(2)P antibody in flies pan-neuronally expressing GA100, GA200 or GA400 for 5, 15 or 30 days, and quantified the Ref(2)P signal per brain. I found an age-dependent accumulation of Ref(2)P in all genotypes, including driver-only flies (Figure 35A, B), which is in agreement with previous reports (Aparicio, Rana, and Walker 2019; Nezis et al. 2008). While Ref(2)P intensity increased from day 5 to day 15 in all GA-expressing flies, it remained unchanged in controls (Figure 35B). This may be due to Ref(2)P showing a much more punctated pattern in this genotype compared to GA-expressing flies at day 15, which may dilute the detection of subtler changes between day 5 and 15. Therefore, I also quantified Ref(2)P puncta in driver-only flies at these ages and found a marked time-dependent increase in Ref(2)P puncta (Figure 35C).

In addition, unlike GA100 or GA200, GA400 expression increased Ref(2)P accumulation compared to control conditions even 5 days after transgene induction (Figure 35A, D). After 15 and 30 days of GA expression, Ref(2)P accumulated significantly more in all GA-expressing flies (Figure 35A, D). However, while GA100 caused a progressive accumulation of Ref(2)P levels from day 5 to day 30, these plateaued at 15 days for flies expressing GA200 or GA400 (Figure 35B). I further measured these expression changes by performing western blotting on heads of flies induced to pan-neuronally express each construct for 5 or 15 days. In agreement with our immunostaining results, only GA400 expression for 5 days markedly increased Ref(2)P levels compared to driver-only flies (Figure 35E, F). At 15 days of expression induction, all GA-expressing conditions exhibited similarly increased Ref(2)P levels compared to driver-only flies (Figure 35E, F). Altogether, these results suggest that brain-specific GA expression up-regulates Ref(2)P, with longer repeats activating this response earlier.

I next tested whether the increase in Ref(2)P expression was a consequence of GA-mediated proteasome inhibition and subsequent accumulation of poly-ubiquitinated proteins, as proteasome impairment has been reported to increase Ref(2)P and autophagy in an attempt to maintain proteostasis (Velentzas et al. 2013). In addition, GA175-GFP forms densely packed ribbons that sequester numerous proteasome complexes in primary rat neurons (Q. Guo et al. 2018) and inhibits overall proteasome activity (Khosravi et al. 2020).

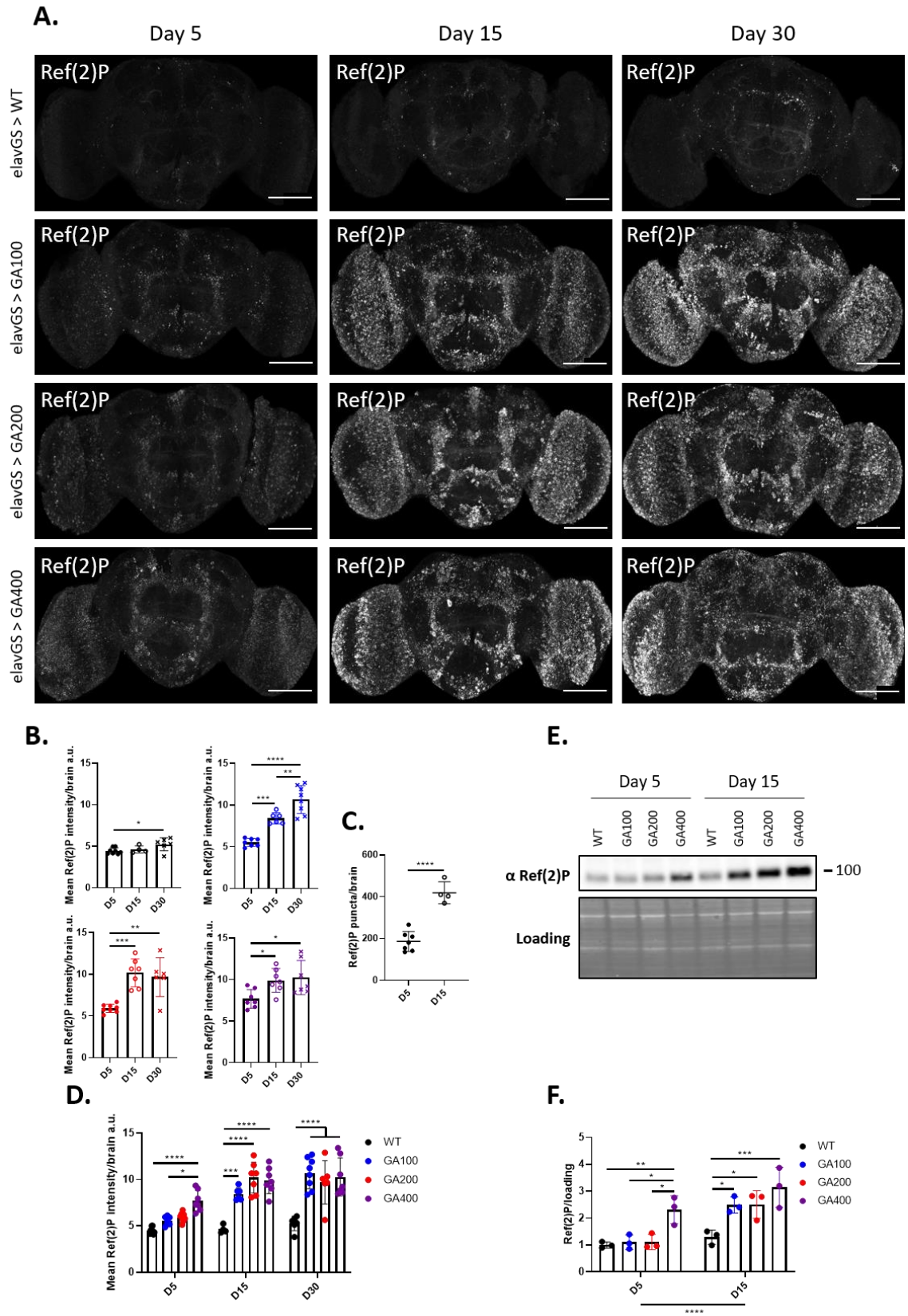

Figure 35. Brain cells up-regulate Ref(2)P upon GA400 expression earlier than upon GA100 or GA200 expression induction.

(Below)

Figure 35. Brain cells up-regulate Ref(2)P upon GA400 expression earlier than upon GA100 or GA200 expression induction.

A Representative images of adult fly brains upon expression of the indicated constructs under the pan-neuronal driver elavGS for 5, 15 or 30 days. Brains were probed with an anti-Ref(2)P antibody (white) and the same laser power settings were used for imaging. Scale bar in images is 100 μm. **B** Quantification of the mean Ref(2)P intensity signal per brain, genotype and age. Each dot represents one independent brain. Brains of the same genotype were compared to each other. Ref(2)P accumulated over time in all conditions, but this occured to a larger extent in GA-expressing flies. Ref(2)P levels reached a plateau on day 15 in GA200- (red) and GA400- (purple) expressing flies, whereas it progressively accumulated in driver-only (black) and GA100 (blue) flies from day 5 to day 30 (****P<0.0001, ***P<0.001, **P<0.01, *P<0.05; One-way ANOVA with Tukey's multiple comparisons test, n=4-8 brains). **C** Quantification of Ref(2)P puncta in whole brains of driver-only flies of 5 or 15 days of age (****P<0.0001; t-test, n=4-7 brains). **D** Quantification of the mean Ref(2)P intensity signal per genotype and age. The results on panel B are compared to each other on this plot to facilitate visualization. GA400 expression caused the earliest increase in Ref(2)P levels (age: ****P<0.0001; genotype: ****P<0.0001; interaction: ***P<0.001; Two-way ANOVA with Tukey's multiple comparisons test, n=4-8 brains). **E** Immunoblotting of head protein extracts from flies expressing the indicated constructs for 5 or 15 days under elavGS. Extracts were probed with an anti-Ref(2)P antibody, and equal amounts of protein were loaded per condition. **F** Quantification of Ref(2)P protein levels from immunoblotting shown in D. GA400 expression increases Ref(2)P levels by day 5. Ref(2)P levels rose between day 5 and 15 in all conditions, but this was particularly prominent in GA-expressing flies (age: ****P<0.0001; genotype: ****P<0.0001; interaction: n.s. = not significant Two-way ANOVA with Tukey's multiple comparisons test, n=4 sets of 20 fly heads).

To test whether GA-mediated increase in Ref(2)P occurs in response to increased poly-ubiquitinated proteins, I performed western blotting on heads of flies induced to pan-neuronally express GA100, GA200 or GA400 for 5 or 15 days. I did not detect an overall change in poly-ubiquitinated proteins in GA-expressing heads at any of the two ages tested (Figure 36A, B). Of note, I detected less poly-ubiquitinated proteins in 15-days-old heads than in 5-days-old heads (Figure 36A, B). This may be associated with the age-dependent increase in Ref(2)P levels that I observed in all conditions (Figure 35B, C), suggesting that Ref(2)P up-regulation may contribute to maintaining proteostasis.

To test whether the early increase in Ref(2)P in GA400-expressing flies was due to proteasome inhibition, I extracted the protein content of heads from flies induced to pan-neuronally express each construct for 5 days, ran it through a native gel in conditions that allowed the maintenance of proteasome assembly and activity, and subsequently incubated the gel with a fluorogenic substrate. The cleavage and subsequent fluorescence emission of this substrate are directly correlated with the chymotrypsin activity of the proteasome (Tain et al. 2017; Vernace et al. 2007). This in-gel assay enables the quantification of the activity of the 20S core particle of the proteasome, which carries out the hydrolyzing functions of the proteasome, as well as that of the 26S proteasome, which comprises one or two 19S particles that modulate the degradation of poly-ubiquitinated substrates within the 20S particle (Ben-Nissan and Sharon 2014). As expected, I detected two independent bands that reflected the activity of the 20S and 26S proteasome complexes in our head protein extracts (Figure 36C). However,

no changes in the activity of the fully assembled 26S proteasome complex or the 20S core were observed (Figure 36C-E).

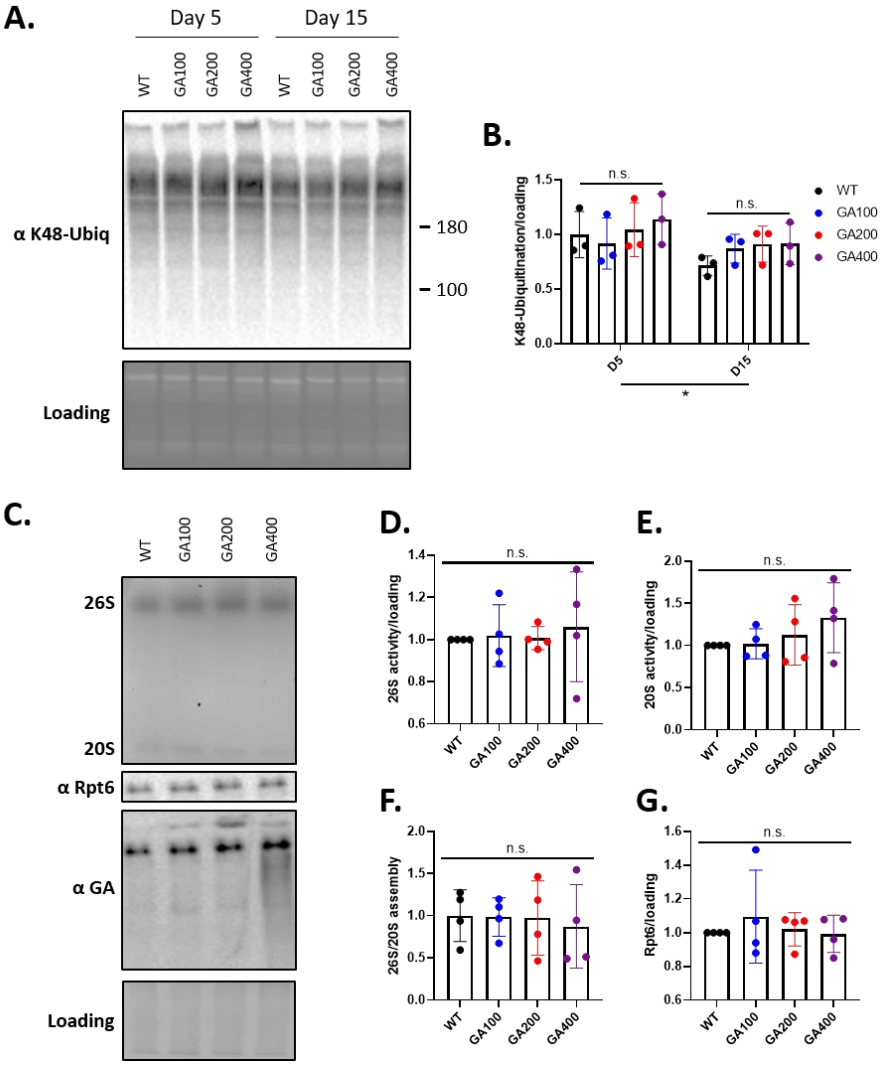

Figure 36. GA expression does not trigger an early increase in poly-ubiquitinated proteins or a decrease in proteasome activity.

A Immunoblotting of head protein extracts from flies expressing the indicated constructs for 5 or 15 days under the pan-neuronal driver elavGS. Protein extracts were probed with an anti-K48-linked poly-ubiquitin antibody, this PTM targeting proteins for proteasomal degradation. Equal amounts of protein were loaded. **B** Quantification of the K48-poly-ubiquitin signal per genotype and age. Values were normalized to K48-poly-ubiquitin levels in driver-only flies (i.e., WT) at day 5. GA expression does not raise levels of poly-ubiquitinated proteins either after 5 or 15 days of expression. However, these overall mildly decrease from day 5 to day 15 (age: *P<0.05; genotype: n.s. = not significant; interaction: n.s.; Two-way ANOVA with Tukey's multiple comparisons test, n=3 sets of 20 fly heads). **C** In-gel proteasome activity assay using head protein extracts from flies expressing the indicated constructs for 5 days under elavGS. Protein extracts were run in a native gel and incubated with a fluorogenic substrate cleavable by the proteasome. The two observed bands are indicative of the activity of the fully assembled 26S proteasome complex and of the 20S core particle. Protein extracts were transferred to a membrane and probed with anti-Rpt6 and anti-GA antibodies. Coomassie staining was used as loading control. Specific GA bands were observed above the Rpt6 band in GA100- and GA200-expressing flies, and a strong smear in association with the Rpt6 was observed for GA400-expressing flies. **D, E** Quantification of the 26S **(D)** and 20S **(E)** activities per genotype. Each biological replicate was normalized to the driver-only control of the same dissection round. **F** Quantification of the 26S/20S assembly per genotype after normalizing each condition to the mean of driver-only control flies. No changes were found across genotypes. **G** Quantification of the Rpt6 signal per genotype after normalizing Rpt6 levels to protein loading and the driver-only control of the same dissection round. No changes were found across genotypes (n.s. = not significant; One-way ANOVA with Tukey's multiple comparisons test, n=4 sets of 5 fly heads).

I also quantified the percentage of 26S vs 20S proteasome activity to calculate the assembly efficiency in each condition. However, no differences were noted in this parameter either (Figure 36F). Finally, I tested whether GA expression led to increased levels of the 19S component Rpt6 (Tsakiri et al. 2019), which could account for the lack of proteasome activity changes despite GA-mediated sequestration of proteasomes. To this end, I transferred the protein content from the native gels to a membrane and performed immunoblotting. However, no changes in Rpt6 levels were observed after 5 days of expression of any of the GA DPRs (Figure 36C, G). Nonetheless, following probing with an anti-Rpt6 antibody, I probed the membrane with an anti-GA antibody and found a systematic smear around the Rpt6 band of the 26S proteasome complex only in GA400-expressing extracts (Figure 36C). The running of Rpt6 and GA400 at apparently similar molecular weights on a native gel suggest a potential physical interaction specific for this GA DPR and the 26S proteasome that should be further investigated.

Collectively, our results show that Ref(2)P up-regulation is a common brain response against GA DPRs and that longer GA DPRs activate this compensatory response earlier than shorter GA DPRs. Our data from fly heads suggests the increase in Ref(2)P levels was not a consequence of excessive accumulation of poly-ubiquitinated proteins or lower proteasome activity. However, this might result from the use of heads, which contain tissues other than the brain that could dilute the GA-specific effects in the brain. In addition, I found a potential interaction between GA400 and the 26S component Rpt6, which may be associated with the early increase in Ref(2)P levels in GA400-expressing flies.

3.3.3.2 GA100 and GA200 exacerbate DNA damage more than GA400

As explained earlier, expression of GA DPRs exacerbates DNA damage (Nihei et al. 2020), with DNA fragmentation being a classical event upstream and downstream of caspase activation during apoptosis. I previously showed that long-term expression of GA100 leads to increased DNA damage, which is not recapitulated by expression of non-toxic forms of tagged GA100 (Figure 28A, B). Therefore, to test whether the repeat length of GA DPRs affects their mediation of DNA damage, I performed a TUNEL assay using brains from flies pan-neuronally expressing GA100, GA200 or GA400 for 5, 15 or 30 days, and quantified the number of TUNEL positive cells. All genotypes, including driver-only flies, exhibited an age-dependent accumulation of TUNEL positive cells (Figure 37A, B). Cell death was also predominantly

observed in OLs. Interestingly, at 15 days GA200-expressing flies displayed significantly more apoptotic cells than GA100 or GA400, which was further exacerbated at 30 days of expression (Figure 37C). Finally, GA100 expression was associated with a greater increase in TUNEL-positive cells over time than was GA400 expression (Figure 37C).

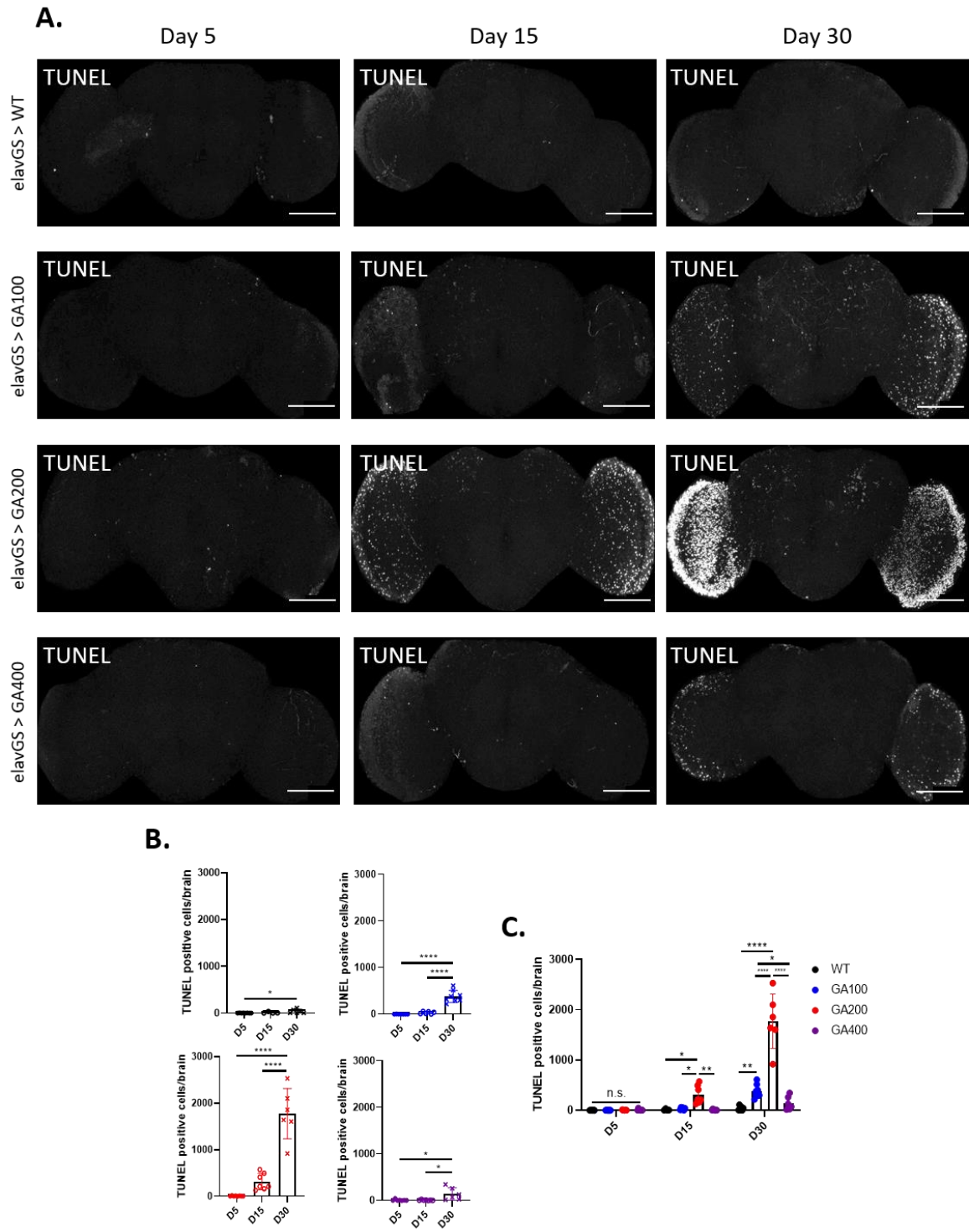

Figure 37. GA-mediated DNA damage does not match GA toxicity.

(Below)

Figure 37. GA-mediated DNA damage does not match GA toxicity.
A Representative images of adult fly brains upon expression of the indicated constructs under the pan-neuronal driver elavGS for 5, 15 or 30 days. Brains were stained with the TUNEL dye (white) to label cells with dsDNA breaks, and the same laser power settings were used for imaging. Scale bar in images is 100 μm. **B** Quantification of the number of TUNEL positive cells per brain, genotype and age. Each dot represents one independent brain. Brains of the same genotype were compared to each other. TUNEL-positive cells accumulated over time in all conditions, but this occurs to a larger extent in GA-expressing flies, especially upon GA200 expression (red) (****P<0.0001, *P<0.05; One-way ANOVA with Tukey's multiple comparisons test, n=4-8 brains). **C** Quantification of the number of TUNEL positive cells per brain, genotype and age. The results on panel B are compared to each other on this plot to facilitate visualization. GA200 increased DNA damage by day 15, while GA100 caused this effect by day 30. GA400 did not significantly increase DNA damage (age: ****P<0.0001; genotype: ****P<0.0001; interaction: ***P<0.001; Two-way ANOVA with Tukey's multiple comparisons test, n=4-8 brains).

These results show that GA DPRs of various repeat lengths can induce DNA damage upon brain-specific expression *in vivo*. However, GA100 and GA200 caused more severe DNA damage than GA400. Therefore, different read-outs of toxicity are differentially affected by GA DPRs of different sizes, indicating that they are mediated by different mechanisms from each other.

4. Discussion

4.1. Spread of GA DPRs in the fly brain

Recent studies have shown that some proteins that typically aggregate in the brains of patients with various neurodegenerative diseases show the ability to be transmitted inter-cellularly in model organisms (Peng, Trojanowski, and Lee 2020). These findings have led to the hypothesis that protein transmission could underpin the clinical progression of such patients (Brettschneider et al. 2015). For instance, the clinical progression of ALS and FTD may be explained by the progressive spreading of TDP43 pathology across conserved neuronal circuits relevant to these diseases (Brettschneider et al. 2013, 2014). In this study, I have explored whether the toxic DPRs derived from a repeat expansion in the *C9orf72* gene can spread *in vivo* using the brain of adult *Drosophila melanogaster* flies as a model. After showing the particular ability of GA DPRs to spread, I have characterized the contribution of clinically relevant factors, namely repeat length and age, to propensity to spread *in vivo*. Finally, I have explored whether this phenomenon causes non-cell-autonomous responses, as well as investigated the molecular mechanisms that may underlie GA transmission *in vivo*.

4.1.1 GA DPRs spread rapidly in a repeat length- and age-dependent manner in the fly brain

To investigate whether the toxicity-associated *C9orf72* DPRs could also spread under *in vivo* conditions, I first used flies expressing, exclusively in ORNs or in MNCs, mCherry-tagged constructs of GA, GR and PR, to avoid an antibody bias, as antibodies specific for different DPRs could have different sensitivities. I only found evidence of trans-neuronal spread for GA DPRs, which is in line with a previous cell culture study, where GA, but not GR or PR, was found to spread from cell to cell (Zhou et al. 2017). This particular ability of GA to spread may also contribute to its greater detection in patient tissue (Mori, Weng, et al. 2013). Future studies should address what mechanisms are activated by GA, and not by GR or PR, that could be associated with its release from DPR-expressing neurons and/or its uptake by recipient neurons. While expression of the arginine-rich DPRs has been mostly associated with translation inhibition (Moens et al. 2019; Y. J. Zhang et al. 2018), impairment of NCT (Freibaum et al. 2015; Jovičič et al. 2015; K. Zhang et al. 2015), RNA processing (Conlon et al. 2016; Cooper-Knock et al. 2014; Kwon et al. 2014; Y. B. Lee et al. 2013), and dysfunction of stress granule dynamics (Tao et al. 2015; Y. J. Zhang et al. 2018), GA expression has been strongly

correlated with impairment of proteostasis (Q. Guo et al. 2018; Y. J. Zhang et al. 2016). I observed that the co-expression of eGFP-tagged synaptotagmin with mCherry-tagged GR and PR led to decreased eGFP signal, which is likely due to their well-known inhibitory effect on translation (Kanekura et al. 2016). In contrast, mCherry-tagged GA co-expression increased eGFP signal. While proteostasis impairment by proteasome inhibition has been shown to exacerbate the release of toxic proteins (J. G. Lee et al. 2016), in this study I could not find evidence of GA-mediated proteasome activity inhibition when assessing its chymotrypsin-like activity in whole head extracts after pan-neuronal GA expression. It remains, therefore, unclear why GA co-expression leads to increased eGFP signal in the synaptic terminals of ORNs, but this could be related to increased synaptic transport of synaptotagmin-eGFP or increased proteostasis stress independent of proteasome inhibition upon GA expression. Another reason why GA, but not GR or PR, may spread could be related to its ability to travel along axons. In contrast to GA, I did not detect GR or PR in the axons or synaptic terminals of ORNs. Consistently, a recent study found that GA DPRs are actively transported along neuronal dendrites and axons in cultured mammalian neurons (Jensen et al. 2020). Furthermore, GA, unlike GR and PR, has been shown to form oligomeric amyloids (Chang et al. 2016), which have been associated with greater propensity to spread (Lasagna-Reeves et al. 2012). This feature could also account for the greater propensity of GA to spread compared to GR and PR. Overall, our findings further support previous *in vitro* studies by showing that GA DPRs have a unique ability to propagate, which occurs upon expression from different neuronal subsets *in vivo*.

Interestingly, unlike mHtt, which strongly spread to a pair of large posterior neurons in the posterior protocerebrum upon expression in ORNs (Babcock and Ganetzky 2015), the spreading pattern of GA DPRs was not overtly specific to a single neuronal subset, either upon expression from ORNs, MNCs or OLs. Co-stainings with the cell membrane dye phalloidin revealed that the propagated GA signal partially accumulated in the shape of aggregate-like puncta in the cytoplasm of nearby recipient cells. However, a large amount of the propagated GA signal was not detected in close proximity to cell nuclei, but rather in brain areas rich in neuronal processes, which are difficult to attribute to a single neuronal subset. The large prevalence of propagated GA puncta in neuronal processes may hinder the detection of a clear spreading pattern of GA DPRs from ORNs.

Both the morphology and spreading propensity of GA DPRs is dependent on repeat lengths. In the adult fly brain, GA36mCherry shows a diffuse staining and does not spread. In contrast, GA100mCherry and GA200mCherry look aggregated and spread in a repeat length-dependent manner. The effect of repeat length on spread was corroborated using untagged GA100 and GA200. In the case of untagged GA DPRs, differences in GA transmission could be confounded by the GA antibody having more epitopes to bind to in the longer constructs, which would lead to stronger propagation signal being detected for GA200 than for GA100. However, this is unlikely because I also saw greater spread for mCherry-tagged GA200 compared to mCherry-tagged GA100 when directly recording their mCherry fluorescence signal, which is only influenced by the overall abundance of each DPR as both GA100mCherry and GA200mCherry carry one mCherry tag per poly-GA molecule. In addition, detectability sensitivity differences are unlikely to underlie our findings, as detection of GA100mCherry propagated puncta remained largely unchanged despite imaging with a wide range of laser power settings, which, in contrast, did reveal increased GA200mCherry propagation. Furthermore, GA400 did not spread from ORNs, which further suggests that the greater transmission observed for GA200 DPRs than for GA100 DPRs is not simply due to GA200 having twice as many GA epitopes as GA100. Given that the amino acid composition of GA100, GA200 and GA400 DPRs is largely the same, I hypothesize that GA200 molecules form aggregates of a morphology more compatible with spreading than GA100 and GA400. When directly comparing the aggregation pattern of GA100mCherry and GA200mCherry, I found that both formed similar amounts of insoluble aggregates, but GA200mCherry formed significantly more soluble oligomers than GA100mCherry. In addition, propagated GA200mCherry puncta were overall smaller than those formed by GA100mCherry puncta, supporting that structural differences may affect propagation propensity. A previous study found that tau oligomers spread pathology more efficiently than aggregated tau fibrils in mice (Lasagna-Reeves et al. 2012). Therefore, the increased levels of oligomeric GA upon GA200mCherry expression could underlie its greater propensity to spread. While I did not detect such oligomeric species in untagged GA, other structural differences not detectable by our experimental approach may underlie the greater propensity to spread of GA200 compared to GA100 and GA400. Moreover, I detected greater synaptic levels of untagged or mCherry-tagged GA200 than of its GA100 counterparts, suggesting that repeat length may modulate the transport or mobility of GA DPRs. In fact, a recent study showed that GFPGA400 is less mobile than GFPGA25 in

cortical and in motor primary neurons. However, no mobility differences were found between GFPGA100 and GFPGA200 (Jensen et al. 2020). These contrasting results between studies may be due to differences in the GA constructs used, as GFP-tagged GA DPRs may show different mobility patterns compared to mCherry-tagged and untagged GA.

Ageing is a major risk factor for ALS and FTD (Niccoli, Partridge, and Isaacs 2017). I found that GA spread was greater when its expression was induced at old age, suggesting that ageing-associated factors, such as impaired proteostasis, promote DPR spread, and therefore that GA transmission is likely to occur at a greater rate at old ages. These findings are in agreement with a recent study that reported a 2-fold increase in the spreading rate of tau upon injection of human tau-expressing viruses in aged murine brains compared to young ones (Wegmann et al. 2019). For instance, spread-competent GA aggregates may be more easily degraded when they are formed at a young age, when the proteostasis machinery works efficiently, but accumulate during the progressive decline of proteostasis during ageing. The excessive accumulation of GA aggregates could trigger their mislocalization to exocytotic vesicles, thus mediating their extracellular release. Alternatively, and as further discussed below, I found that GA spread is attenuated by genetically down-regulating genes involved in the release of exosomes and maintenance of synaptic vesicles. Therefore, age-related changes in these pathways may also contribute to the greater spread of GA upon expression in old tissue. To our knowledge, neuronal activity changes in ORNs during ageing remain unexplored, but ageing decreases the synaptic activity of *Drosophila* motor (Azpurua, Mahoney, and Eaton 2018) and clock neurons (Curran et al. 2019), making increased neuronal firing in old ORNs an unlikely mediator of the age-related increase in GA transmission. In contrast, autophagosomes, which typically accumulate during ageing following lysosomal dysfunction (Wyss-Coray 2016), can fuse with multivesicular bodies, the precursors of exosomes, and these can in turn fuse with the plasma membrane, this secretion pathway being known as exophagy. Rab11 has been reported to mediate exophagy (Y. Da Chen et al. 2017) and in our study co-expression of Rab11DN led to decreased GA spread. Therefore, I hypothesize that during ageing exophagy may also be exacerbated to compensate for the lower lysosomal activity and maintain proteostasis, which may contribute to exacerbated GA spread upon expression in old tissue. Finally, since phagocytic clearance by glia decreases in old brains (Purice, Speese, and Logan 2016), I cannot rule out that lower phagocytosis of GA-containing recipient cells by older glia may contribute to the accumulation of propagated GA.

Collectively, in this study (Morón-Oset et al. 2019) we have shown that GA DPRs can rapidly spread *in vivo* from various neuronal subsets, which is strongly influenced by their repeat length and the age of the GA-expressing neurons.

4.1.2 GA transmission is associated with lower Ref(2)P levels in recipient tissue

One current limitation of most studies investigating the cell-to-cell transmission of pathological proteins is the lack of reported consequences of this phenomenon in the recipient tissue, which is key to fully assess its contribution to toxicity (Walsh and Selkoe 2016). GA DPRs can sequester proteins involved in DNA damage repair (Nihei et al. 2020) and inhibit the proteasome (Khosravi et al. 2020). In our spread paradigm, ORN-specific expression of GA200 did not change DNA damage in the surrounding tissue. In contrast, pan-neuronal expression of GA200 increased dsDNA breaks, mostly in the OLs. This anatomical area has been reported to be particularly prone to accumulate DNA damage upon expression of other aggregation-prone pathological proteins, such as poly-glutamine-containing ataxin 7 (Latouche et al. 2007). However, propagated GA aggregates were rarely detected in the OLs, which may explain the lack of overt DNA damage associated with GA spread. Since DNA damage is often an upstream or downstream event during apoptosis (Hou et al. 2019), it is interesting to note that mice expressing high levels of aggregated GA in the spinal cord displayed no overt signs of cell death despite showing a severe locomotor phenotype (Schludi et al. 2017). This suggests that the ability of GA to cause cell death, and potentially DNA damage, may be brain region-dependent.

While I found lower accumulation of cells bearing Ref(2)P aggregates in flies where GA spread had occurred, long-term pan-neuronal expression of GA200 led to strong Ref(2)P up-regulation. Differences in the amount of accumulated GA200 in each expression paradigm might account for these opposite effects. Low levels of accumulated GA200 upon its trans-cellular spread from ORNs may mildly stress recipient cells, thus leading to a compensatory response that would efficiently clear Ref(2)P-positive aggregated proteins. In contrast, high levels of accumulated GA200 upon its pan-neuronal expression would strongly stress brain cells, thus leading to a more dramatic up-regulation of Ref(2)P that may not be overtly degraded due to the constant large resupply of newly translated GA200 molecules. Indeed, Ref(2)P protein levels rise during fly ageing according to others (Bartlett et al. 2011; Nezis et

al. 2008) and our own data, suggesting that high Ref(2)P levels are indicative of cellular stress. To ascertain whether GA spread causes or not toxicity, it would be interesting to test whether GA200-mediated lifespan shortening, DNA damage increase and Ref(2)P up-regulation are rescued by interfering with its trans-cellular spread by pan-neuronally co-expressing GA200 along with *comt* RNAi or Rab11DN, which lowered GA200 spread from ORNs. Of note, while our current data do not show that propagated GA impairs proteostasis in our ORN spread paradigm, the effects of GA spread may be different in the human context of the *C9orf72* mutation, where the levels of the autophagy-relevant C9orf72 protein are significantly lower (Saberi et al. 2018), or a context where recipient cells are more stressed. For instance, GA transmission in cell culture increases expression of mutant *C9orf72* RNA and RAN translation in a non-cell-autonomous manner, which is indicative of a vicious cycle of boosted repeat RNA and DPR expression upon GA spread (Zhou et al. 2017). Indeed, RAN translation is exacerbated upon cellular stress (Green et al. 2017). Furthermore, one recent *in vitro* study found that transmitted GA can inhibit the proteasome of recipient neurons and trigger the accumulation of poly-ubiquitinated proteins, including TDP43 (Khosravi et al. 2020). While the authors did not investigate non-cell-autonomous effects of GA on p62, the mammalian homolog of Ref(2)P, it is tempting to speculate that the lower proteasome activity upon GA uptake may activate the autophagic pathway in an attempt to maintain proteostasis. Since TDP43 is degraded by the proteasome (Igaz et al. 2009; Khosravi et al. 2020), autophagy up-regulation would not prevent TDP43 cytoplasmic accumulation, but, if efficient, it would help maintain proteostasis, thus causing lower accumulation of p62. However, I did not find evidence of proteasome inhibition by expression of GA DPRs. Thus, there might be alternative explanations, like increased phagocytic activity of glial or immune cells following GA spread. A previous study in *Drosophila* found that expression of the AD-related protein Aβ42 causes the recruitment of fly macrophages to the brain (S. C. Wu et al. 2017), which may also be recruited upon GA spread. This response may result in the phagocytosis of stressed cells or in the release of cytokines that attenuate stress in the recipient tissue, which may, in turn, be detrimental in the long term.

Alternatively, given that GA200 was the least toxic GA DPR despite being the most propagation-prone, I can currently not rule out that the trans-cellular spread of GA200 may cause some advantage. For instance, it could help lower GA levels in neurons whose function is most essential for lifespan. Glutamatergic neurons are particularly vulnerable to expression

of arginine-rich DPRs in flies (W. Xu and Xu 2018). However, this remains currently unexplored for GA. Furthermore, I cannot exclude the possibility that GA expression may elicit a cell-autonomous response in ORNs, which may in turn affect Ref(2)P protein levels in the surrounding tissue independent of GA spread.

Overall, our data suggest that recipient cells respond to GA transmission from ORNs by showing milder signs of age-related proteostatis stress. However, I can currently not conclude whether or not GA spread in our paradigm is toxic and therefore, a more detailed characterization of the non-cell-autonomous effects of GA DPRs in an *in vivo* model more similar to the human *C9orf72* mutation setting warrants further investigation.

4.1.3 GA transmission depends on exocytosis of exosomes and synaptic vesicles

In addition to understanding the consequences of the cell-to-cell transmission of pathological proteins, it is essential to elucidate the mechanism(s) underlying this event to develop interventions that can efficiently block the spread. Several mechanisms have been proposed to mediate transmission of pathological proteins, including release via exosomes and exocytosis upon fusion of synaptic vesicles. However, the great majority of our current understanding of these mechanisms is based on data from *in vitro* models, the majority of which has so far not been corroborated *in vivo*. I performed a genetic miniscreen using our *in vivo* model of GA spread by co-modulating the expression of several genes involved in the release of exosomes, as well as in the fusion and fission of synaptic vesicles. In agreement with a previous study showing that GFPGA50 can be transmitted via exosomes in neuronal cultures (Westergard et al. 2016), down-regulating the expression of the exosomal genes *hsp90*, *stip1* and *syntx1A* by RNAi reduced GA transmission from ORNs at 30 days of expression induction. Additionally, over-expression of Rab11DN strongly blocked GA spread both at 3 and 30 days of expression initiation. The earlier effect of Rab11DN may be due to RNAi constructs needing a longer time to down-regulate their targets in a significant manner. In the future, these genetic interventions should be tested after initiating GA expression in other neuronal subsets to confirm that the involvement of exosomes in GA transmission is not restricted to our ORN paradigm. These data further support the currently growing evidence in favour of exosomes playing a key role in the progression of human diseases, such as tumour metastasis (Y. Guo et al. 2019).

Higher neuronal activity increases tau release and spread in mice (J. W. Wu et al. 2016), and genetically interfering with the recycling of synaptic vesicles lowers mHtt spread from ORNs in *Drosophila* (Babcock and Ganetzky 2015), indicating that changes in neuronal firing and/or physiological maintenance of synaptic vesicles are common modulators of the transmission of pathological proteins. Therefore, I tested whether GA spread from ORNs is affected by similar interventions. I consistently found greater synaptic accumulation of GA and lower spread upon Rab3DN co-expression, which blocks the fusion of synaptic vesicles at the presynaptic membrane (Schlüter et al. 2006). Co-expressing ShiDN, which is required for the recycling of synaptic vesicles after neurotransmitter release to maintain homeostatic levels of synaptic vesicles (Bengtson and Kitamoto 2001), markedly lowered synaptic GA and also mildly reduced spread. A similar reduction in synaptic and propagated GA was found upon co-expression of *comt* RNAi, which is also involved in synaptic vesicle recycling (Kawasaki and Ordway 2009). However, the *comt* RNAi effect on accumulated GA at ORN terminals was not observed in our genetic miniscreen, where the UAS-GA200 transgene was expressed from the attP40 site located at position chromosome 2L. This discrepancy may be due to changes in GA200 levels upon transgene expression from different loci, which may affect the relative contribution of each pathway to mediate GA transmission. Indeed, strong regulatory elements, which UAS promoters are known to be, can modulate the transcriptional expression of other genes on different chromosomes in *trans*, which is influenced by the genomic position of each of the elements (King, Johnson, and Bateman 2019). A currently unclear genetic interaction may also account for the consistently greater spread found upon long-term mCherry co-expression. Alternatively, the latter might be due to mild proteostasis stress inflicted by co-expressing an exogenous protein like mCherry, thus reducing the clearance of propagation-competent GA. In addition, long-term expression of Rab3CA also strongly reduced GA spread without affecting its synaptic levels. Rab3CA can rescue the morphological and functional defects of *rab3* mutant fly larvae at the neuromuscular junction (S. Chen et al. 2015), suggesting that it is also functional in ORNs, at least in the short term. However, its full activity requires other effector proteins, such as Rab3 guanine nucleotide exchange factor (Rab3-GEF) (Bae et al. 2016). I therefore hypothesize that long-term expression of Rab3CA may exhaust effector proteins required for the tethering of synaptic vesicles to the pre-synaptic membrane, thus failing to sustain GA spread over time. Furthermore, co-expression of both inhibitory and excitatory ionic channels led to an overall mild reduction of GA spread. While the

interpretation of these results is currently unclear, they do suggest that neuronal activity may modulate GA transmission in a complex manner. In fact, GFPGA itself modulates synaptic activity in neuronal cultures by reducing the protein levels of the synaptic protein SV2, thereby lowering overall synaptic output (Jensen et al. 2020). Future studies should address whether increasing neuronal activity in mammalian GA-expressing neurons or neuron-derived human induced pluripotent stem cells carrying the *C9orf72* mutation affects GA intercellular transfer. This could be of particular relevance in ALS, where hyperexcitability and excitotoxicity play a key role in pathogenesis (Starr and Sattler 2018).

Collectively, these results suggest that GA may also accumulate within synaptic vesicles at axonal terminals and be at least partially transmitted upon fusion of these. I propose a scenario where GA may be located in various vesicle pools, including exosomes and synaptic vesicles, from which this DPR may be susceptible to being released and subsequently taken up by nearby or synaptically connected neurons. Ongoing electron microscopy analyses should shed light on the location of synaptic GA. Of note, expression analyses need to be performed to test the effect of expressing GA200 from specific loci and in concert with the most promising genetic modulators in order to correctly interpret our current results.

4.2. Tags affect GA toxicity, aggregation pattern and cellular responses

While understanding the contribution of non-cell-autonomous toxicity by GA DPRs may be relevant to develop strategies that block the clinical progression of *C9orf72* mutation carriers, interventions targeting cell-autonomous pathways affected by DPRs also hold great promise. One key question in this regard is which DPR causes greatest toxicity and would therefore provide the largest clinical benefit if therapeutically targeted. Out of the five *C9orf72* DPRs, experimental models support that arginine-rich DPRs are the most toxic (Mizielinska et al. 2014; Wen et al. 2014). GA has been associated with mild toxicity (Mizielinska et al. 2014), yet it is the most prevalent DPR in patient tissue (Mori, Weng, et al. 2013). In this study, I addressed whether GA toxicity may have been underrated by the use of tagged and too short constructs by most laboratories. Regarding the former, our data show that GA-mediated toxicity, aggregation and elicited cellular responses are affected by tags.

4.2.1 Tags reduce GA100-mediated lifespan shortening upon pan-neuronal expression

I investigated the effect of tags on GA toxicity by generating and characterizing flies expressing GA100, GA100GFP, GA100mCherry and GA100FLAG from the attP2 landing site. Pan-neuronal expression of untagged GA100 efficiently shorted fly lifespan, in agreement with a previous study using the same driver and construct expressed from the attP40 site (Mizielinska et al. 2014). GA100FLAG expression caused mild but less toxicity than untagged GA100, and the lifespans of GA100GFP- and GA100mCherry-expressing flies were not different from that of driver-only control flies. These results indicate that GA toxicity is strongly reduced if expressed as a fusion protein, especially if the tag to which it is linked is relatively long. It will be interesting to test whether the same effects are observed with other toxicity assays. In contrast to GA, I found that eye-specific expression of GR36mCherry and PR36mCherry at 25°C caused more severe eye toxicity than that of untagged GR36 and PR36, respectively. However, I could not identify major toxicity differences between untagged or mCherry-tagged GR100 or PR100, despite PR100mCherry being more strongly detected than untagged PR100. This might be due to the relatively low sensitivity of our assay to identify small toxicity differences upon major retinal degeneration. However, while the effect of different tags on the toxicity of arginine-rich DPRs was not systematically tested in this study, I speculate that tags could affect different DPRs in a distinct manner, which may at least partially account for the lower toxicity associated with GA. In fact, our results suggest that different tags may have differential effects on the same DPR, which may also be dependent on the repeat length, N- or C-terminal position of the tag, as well as the presence, length and nature of a linker. In agreement with this, a previous study found that N-terminally GFP-tagged PR36 shortens fly lifespan and impairs climbing to a lesser extent than untagged PR36 (W. Xu and Xu 2018), which is opposite to our findings for C-terminally mCherry-tagged PR36. This highlights the need of an untagged control when assessing toxicity of fusion DPRs.

It is worth highlighting that at least a fraction of human DPRs are translated with a C-terminal sequence after the repeat and prior to the stop codon. Depending on the reading frame, this sequence consists of 30-55 amino acids (Mori, Arzberger, et al. 2013). While it is unclear how prevalent this C-terminal sequence is in human DPRs, a very recent study found that it strongly reduces GR toxicity in flies and mammalian neuronal cultures (F. He et al. 2020), thus suggesting that the toxicity of arginine-rich DPRs in humans may be milder than anticipated by experimental models. Furthermore, DPRs may be expressed as chimeric repeats, which

influences their molecular behaviour (McEachin et al. 2020). These studies suggest that using pure DPRs may also not fully recapitulate the complex nature of human DPRs, whose understanding is pivotal to develop more clinically relevant models.

4.2.2 Large tags modify GA aggregation

After observing a strong decrease in toxicity upon expression of tagged GA, I investigated whether this correlated with changes in GA aggregation. I found that the aggregation pattern of GA100GFP and GA100mCherry was reminiscent of soluble oligomers, which I repeatedly did not observe upon expression of untagged GA100 or GA100FLAG. In addition, while GA100GFP and GA100mCherry formed discrete puncta upon brain expression that were detected by imaging their GFP and mCherry signals, respectively, an anti-GA antibody did not recognize those aggregates, but only revealed a non-aggregated signal in the brains of GA100GFP- and GA100-mCherry expressing flies. In contrast, the same anti-GA antibody revealed numerous aggregates in fly brains expressing untagged GA100 and GA100FLAG, which did not exhibit any diffuse signal. These results suggest major structural differences when GA is linked to GFP and mCherry. When immunostaining GR100mCherry and PR100mCherry with anti-GR- or anti-PR-specific antibodies, I did not detect major differences in the morphological pattern between tagged and untagged constructs of the same repeat length. However, GR and PR looked less aggregated than GA in our hands, in agreement with the previously reported lower propensity of arginine-rich DPRs to aggregate (Chang et al. 2016; Yang et al. 2015). These aggregation propensity differences may render GA more susceptible to structural abnormalities upon fusion with large tags. In line with this, a recent study found that GFP tagging of both full-length human tau, which is intrinsically prone to aggregate, and an even more aggregation-prone mutant form of tau, dramatically lowers their aggregation propensity and modifies the morphology of the deposits that they form (Kaniyappan et al. 2020). Therefore, the presence of GFP or mCherry may solubilize and stabilize GA, which may partially explain why GA100GFP was more abundantly detected than untagged GA100 by dot blot and western blot. However, this was not recapitulated by GA100mCherry, which was similarly abundant to untagged GA100. These results suggest that not all tags affect the molecular behaviour of GA in the same manner. However, even though all of the lines were generated to express each construct from the same genomic locus,

expression analyses should be conducted to corroborate equal transcriptional regulation in the future.

It is worth pointing out that human studies have so far only tried to anatomically correlate presence of DPR inclusions and neurodegeneration, without distinguishing among pathology types in a large cohort of patients (Schludi et al. 2015). My results show that GA aggregates alone are not good predictors of toxicity. I speculate that aggregates with specific molecular features may be more informative. In support of this, more GR deposits modified by dimethylarginine in a symmetric, but not in an asymmetric configuration, correlate with slower disease course and longer survival (Gittings et al. 2020). This emphasizes that different aggregates containing the same DPR co-exist and may have variable effects on toxicity.

4.2.3 Large tags interfere with cellular responses triggered by GA100

The major toxicity differences between untagged and tagged GA100 strongly suggested changes in their cellular effects. I explored whether tagged GA also triggered Ref(2)P up-regulation and increased DNA damage, which I had previously observed upon pan-neuronal expression of GA200. I found that both effects were recapitulated by untagged GA100 and GA100FLAG. However, neither were triggered by GA100mCherry, and Ref(2)P up-regulation was not observed upon GA100GFP expression. The apparent lack of Ref(2)P up-regulation upon GA100GFP and GA100mCherry expression suggests that brain cells are not stressed in the presence of these constructs. Alternatively, they might not be degraded by autophagy, thus not requiring Ref(2)P up-regulation for their efficient degradation, or be very efficiently disposed of by autophagy, which would also lead to a steady degradation of Ref(2)P along with GA100GFP and GA100mCherry. While several interpretations are compatible with our findings, our data show that large tags affect the cellular responses elicited by GA, which may be due to the nature of the aggregates formed by each protein isoform.

Of note, while GA100FLAG expression up-regulated Ref(2)P and increased DNA damage to a similar extent to untagged GA100, it was significantly less toxic according to our lifespan data. This suggests that other cellular responses may be differentially affected by untagged GA100 and GA100FLAG, which would better account for their large toxicity differences. Alternatively, since these parameters were only assessed after 25 days of expression, I cannot rule out that GA100 may trigger these responses earlier than GA100FLAG.

Collectively, our results show that artificial tags, especially long ones, can have a strong influence in the toxicity, aggregation and cellular consequences of GA expression. I cannot rule out that these strong differences may be specific to our constructs or experimental model, as others have reported toxicity and strong cellular responses in other experimental systems upon expression of tagged constructs (Jensen et al. 2020; Khosravi et al. 2020). However, our data strongly suggest that artificial tags that are not expressed in humans change the behaviour of GA, and probably that of GR and PR. Therefore, I hope our findings serve as a proof of principle and encourage researchers to carefully assess the effect of artificial tags in their future studies. In the meantime, human studies should also address the frequency and contribution of the native DPR C-termini and chimeric DPRs to human pathology to develop the most accurate models.

4.3. Repeat length affects GA toxicity, aggregation pattern and cellular responses

Unlike other microsatellite disorders, where affected genes carry repeats of a relatively short size (e.g., HD patients carry 40-100 CAG repeats) (Nguyen, Cleary, and Ranum 2019), G4C2 repeats in the *C9orf72* mutation show a very heteromorphic length, ranging from a few tens to several thousands (Van Mossevelde, van der Zee, et al. 2017). As detailed in the introduction, published human studies trying to correlate repeat length and disease onset or severity have come to conflicting conclusions (Fournier et al. 2019; Gijselinck et al. 2016). Nevertheless, they have helped draw an even more complex picture by showing that the repeat length can significantly vary within different tissues of the same individual (van Blitterswijk et al. 2013) and during ageing (Fournier et al. 2019).

While great endeavors have been devoted to improving sizing of G4C2 DNA repeats, the real length of the human C9orf72 DPRs remains unclear. However, since RAN translation is favoured by longer repeats (Nguyen, Cleary, and Ranum 2019) and a C-terminal region is detected in, at least some, human DPRs (Mori, Arzberger, et al. 2013), G4C2 RAN translation seems to progress normally once initiated, suggesting that DPRs may be of a size similar to the hexanucleotide repeat expansion. This poses the challenge of generating models that express DPRs of a clinically relevant length, which is complex due to the repetitive nature of these peptides.

In an effort to understand whether the relatively short length of the DPRs used in previous studies has underrated GA toxicity, I have systematically compared the toxicity, aggregation and cellular responses of relatively pure GA DPRs with 100, 200 or 400 repeats of length upon expression in the fly brain from the same genomic locus.

4.3.1 GA400 is more toxic than GA100 and GA200

GA400 caused more toxicity than GA100 or GA200 upon assessment through four different measurements, which shows a clear gain of toxicity for GA400. Since this gain of toxicity was not recapitulated by two copies of GA200, toxicity differences are unlikely to be simply due to the number of GA epitopes comprised in each construct. Therefore, I speculate that the interaction partners and cellular responses elicited by GA DPRs do not simply increase in a linear manner with longer GA DPRs, but may significantly change due to major conformational differences of each repeat length.

As further explained below, it remains unclear why GA200 is less toxic than GA100. G4C2 repeat length may not always hold a linear correlation with toxicity, which may account for the apparent complexity to associate these two parameters in humans. In fact, a previous study investigating CTG repeats located in the dystrophia myotonica protein kinase (*DMPK*) gene in myotonic dystrophy type 1 patients, only found a strong correlation between a longer repeat length and earlier age of onset below 250 repeats (Savić et al. 2002). It is therefore possible that compensatory mechanisms that are more rapidly activated by intermediate sequences suffice to attenuate toxicity of relatively short repeats, but these are overridden from a specific threshold onwards. In the case of the *C9orf72* mutation, my results suggest that longer repeats become more toxic from at least 400 repeats, although future studies should address whether GA gain of toxicity is maintained in even longer DPRs. A pre-print article has shown that there is a repeat length-dependent toxicity gain between mice expressing 500 and 800 G4C2 repeats (Pattamatta et al. 2020). Even though I only evaluated GA-mediated toxicity, the previous study points to toxicity gain being conserved after 400 repeats.

While my results show that GA400 is more toxic than its shorter counterparts, it did not recapitulate the same toxicity levels typically shown by GR100 and PR100 upon pan-neuronal expression in adulthood, as flies expressing GR100 and PR100 typically die by 10 days of age

(Mizielinska et al. 2014) and the longest-lived GA400 flies reached 40 days of age. This might be due to a particularly prominent vulnerability of flies to arginine-rich DPRs, which may or not be directly relevant in ALS/FTD. In fact, a recent study simultaneously compared GFPGA175 and GFPPR175 congenic mice with similar transgene levels and showed that GA-expressing mice better recapitulated ALS symptoms and transcriptional profile than PR mice (LaClair et al. 2020). Furthermore, while GA400 is more toxic than GA100 and GA200, GR and PR may not show a similar toxicity gain as their length increases *in vivo*. For instance, GFPGR200 reduced in half survival of primary cortical neurons, which was not further exacerbated by GFPGR400 (Wen et al. 2014).

Since GA is more readily RAN translated (Green et al. 2017; Tabet et al. 2018) and detected in human post mortem tissue than GR and PR (Mori, Arzberger, et al. 2013), my results showing that GA expression can be associated with high levels of toxicity further support the need to characterize this DPR at clinically relevant sizes. In addition, given that repeat length varies across tissues (van Blitterswijk et al. 2013) and time (Fournier et al. 2019) within the same patient and RAN translation is error-prone (McEachin et al. 2020; Zu et al. 2011), I propose that DPRs of several repeat lengths may co-exist in the same tissue and individual. This heterogeneity may be particularly evident and relevant in patients carrying thousands of repeats, which may pose an additional clinical challenge, as my data show that repeat length strongly affects GA-mediated neurotoxicity.

4.3.2 The subcellular location and aggregation propensity of GA DPRs is repeat length-dependent

The present study shows for the first time that repeat length strongly modulates the subcellular location of GA DPRs *in vivo*, which may directly contribute to toxicity differences. In support of this, cytoplasmic GA175-Myc causes TDP43 cytoplasmic accumulation in primary murine neurons, whereas expressing GA175-Myc with a nuclear localization signal abolishes this effect (Khosravi et al. 2017). My findings that GA400, the most toxic GA DPR, is cytoplasmatic suggest that this subcellular location may be particularly relevant in the context of GA toxicity.

In addition, the subcellular location of GA DPRs is likely to be affected by their tridimensional structure and therefore their aggregation propensity. I found that the electrophoretic mobility of GA DPRs decreases in a linear repeat length-dependent manner. This is a strong indication

of conformational differences among GA DPRs of a distinct size. The fact that GA400, the most toxic GA DPR, was also the most insoluble one is in agreement with a previous study showing that the toxicity and aggregation propensity of GFPGA50 in mice were simultaneously abolished if expressed with proline interruptions (Y. J. Zhang et al. 2016). However, aggregation propensity alone does not correlate with toxicity in my models, as GA100 was more toxic than GA200 despite exhibiting a slower aggregation rate. Therefore, I speculate that other structural differences not detected by assessing electrophoretic mobility may account for differences in GA toxicity.

In conclusion, the present study indicates that the cytoplasmic location of GA DPRs is more predictive of GA toxicity than its propensity to aggregate. I therefore support future neuropathological studies investigating the morphology of GA aggregates residing specifically in the cytoplasm, as these may reveal conformational differences relevant for human pathogenesis.

4.3.3 GA DPRs may cause toxicity by affecting repeat length-dependent mechanisms

Several mechanisms have been shown to be disrupted by GA DPRs, including proteasome activity (Khosravi et al. 2020; Nguyen et al. 2019), DNA damage repair (Nihei et al. 2020), NCT (Khosravi et al. 2017) and synaptic activity (Jensen et al. 2020). However, constructs with different tags and repeat lengths were used in different studies, and the effect of repeat length in pure GA constructs on various pathways has never been systematically investigated *in vivo*.

I found that GA DPRs cause progressive Ref(2)P accumulation in a repeat length-dependent manner, which may be directly associated with their progressive cytoplasmic subcellular location. While GA has been reported to bind to and inhibit the proteasome (Q. Guo et al. 2018; Khosravi et al. 2020) and Ref(2)P levels increase upon pharmacological inhibition of the proteasome in flies (Velentzas et al. 2013), I found no differences in the amount of poly-ubiquitinated proteins or proteasome activity between GA-expressing and non-expressing conditions. These results may be explained by the use of heads instead of brains in my experiments, which dilute the GA effects on brain cells. Of note, I detected, only in GA400-expressing flies, a strong and consistent GA smear around the same location as the 26S subunit Rpt6, which is a preliminary indication that GA400 may physically interact with the fully

assembled proteasome. Co-immunoprecipitation studies, along with the analysis of brains and larger sample sets, should clarify whether GA also interferes with proteasome function in flies and whether repeat length affects this interaction. Nonetheless, I cannot rule out that GA expression may up-regulate Ref(2)P to boost autophagy independent of the proteasome, in an attempt to dispose of GA aggregates that may be more efficiently cleared by the autophagic pathway. In fact, protein aggregates are preferentially cleared by autophagy, which is known as aggrephagy (Lamark and Johansen 2012). Alternatively, the elevated accumulation of Ref(2)P in GA-expressing brains may reflect autophagy blocking, as p62 becomes degraded in autolysosomes to facilitate degradation of ubiquitinated substrates (Pankiv et al. 2007) and its accumulation has also therefore been correlated with autophagy blockade (Gottlieb et al. 2015) Therefore, it will also be of interest to characterize the efficiency of the autophagic flux upon expression of GA DPRs of different repeat lengths to better interpret my Ref(2) data.

While Ref(2)P up-regulation was increased in a linear manner by GA DPRs of an increasing size, DNA damage was not, as this molecular insult was barely observed upon GA400 expression, but it was strongly increased by GA100 and, to an even larger extent, by GA200. Mechanistically, this may be related to the differential ability of GA DPRs to interact with DNA repair proteins, such as pATM and hnRNPA3 (Nihei et al. 2020). Along with my findings that GA100FLAG causes similar levels of DNA damage to GA100 despite being less toxic, these results suggest that GA-mediated DNA damage does not strongly contribute to GA-mediated toxicity. Alternatively, and especially in the case of GA400, brain cells showing DNA damage may undergo degeneration and subsequent phagocytosis by glia, which, in the long term, could contribute to the larger number of degenerative vacuoles observed after long-term GA400 expression. However, I propose that this scenario is unlikely, as I did not detect any signs of DNA damage at earlier ages in GA400-expressing brains. Instead, my data rather support the notion that increased DNA damage may initiate a partially protective response, at least at early disease stages, as this molecular signature is particularly prominent in the least toxic GA DPRs. In this regard, a recent study (Coelho et al. 2018) found that brain-specific expression of the neurotoxic Aβ42 peptide triggers the culling of dysfunctional cells, which is relevant to maintain brain homeostasis, as well as healthspan and lifespan. I hypothesize that this mechanism may be activated in GA200-expressing flies, presumably following accumulation of cells with high levels of DNA damage. In GA400 flies, this pathway would not be activated at early stages, thus preventing the early removal of dysfunctional cells.

Altogether, these findings show that repeat length of GA DPRs profoundly affects their toxicity, aggregation, subcellular location and cellular responses. These data suggest that GA DPRs may activate repeat length-dependent responses, some of which may be shared but occur at different times depending on the repeat length, while others might be unique to DPRs of a specific length. Ongoing proteomics analyses of brains expressing each GA DPR for various times should be informative to better understand the molecular underpinnings of GA-mediated toxicity. This information could be clinically relevant, as it could unravel novel mechanisms of GA toxicity.

5. List of figures

Figure 1. Selective neuronal vulnerability hypothesis. 8
Figure 2. Hypothesis of pathological spread of toxic proteins and associated mechanisms. . 11
Figure 3. The C9orf72 mutation and its molecular products. 23
Figure 4. Proposed toxicity mechanisms mediated by GA DPRs. 32
Figure 5. mCherry-tagged GA constructs can be detected by imaging their endogenous mCherry signal. 53
Figure 6. mCherry-tagged GR constructs can be detected upon pan-neuronal expression. .. 54
Figure 7. mCherry-tagged PR constructs can be detected upon pan-neuronal expression. ... 55
Figure 8. Eye-specific expression of mCherry-tagged, arginine-rich DPRs disrupts eye morphology. 56
Figure 9. GA36mCherry, GR36mCherry and PR36mCherry do not spread from ORNs. 58
Figure 10. GA100mCherry, but not GR100mCherry or PR100mCherry spread outside of ORNs. 60
Figure 11. GA100mCherry, but not GR100mCherry, spread outside of MNCs. 61
Figure 12. Validation of GA200 and GA200mCherry constructs. 62
Figure 13. GA200mCherry forms more abundant puncta and oligomeric bands than GA100mCherry upon pan-neuronal expression. 63
Figure 14. Spreading is higher for longer mCherry-tagged GA repeat proteins. 65
Figure 15. Spreading is increased in longer untagged GA repeat proteins. 67
Figure 16. GA spreads in a repeat length-dependent manner from an independent neuronal population. 69
Figure 17. Age-associated factors exacerbate GA spread 70
Figure 18. Propagated GA accumulates over time. 72
Figure 19. Long-term GA propagation from ORNs does not increase brain DNA damage. 73
Figure 20. Long-term GA propagation from ORNs lowers brain Ref(2)P puncta. 74
Figure 21. Genetic down-regulation of genes involved in the release of exosomes lowers GA200 spread from ORNs. 77
Figure 22. Genetic down-regulation of comatose lowers synaptic and propagated GA200 ... 78
Figure 23. Genetic modulation of genes involved in synaptic transmission affects GA200 spread from ORNs. 81
Figure 24. GA100 toxicity is not recapitulated by tagged GA100. 85
Figure 25. Large tags affect the aggregation pattern of GA100. 86
Figure 26. GA100GFP and GA100mCherry form puncta that are not recognized by an anti-GA antibody. 88
Figure 27. Brain cells do not up-regulate Ref(2)P upon long-term GA100GFP or GA100mCherry expression 89
Figure 28. Long-term GA100mCherry expression does not increase DNA damage. 90
Figure 29. Validation of the GA400 construct. 94
Figure 30. Eye-specific expression of GA400 disrupts eye morphology. 96
Figure 31. GA400 shortens lifespan more than GA100 and GA200. 97
Figure 32. Long term GA400 expression increases neurodegenerative vacuoles. 99
Figure 33. GA subcellular location and aggregation propensity are repeat length-dependent. 101
Figure 34. GA400 does not spread from ORNs. 103

Figure 35. Brain cells up-regulate Ref(2)P upon GA400 expression earlier than upon GA100 or GA200 expression induction. 105
Figure 36. GA expression does not trigger an early increase in poly-ubiquitinated proteins or a decrease in proteasome activity. 107
Figure 37. GA-mediated DNA damage does not match GA toxicity. 109

6. List of tables

Table 1. Summary of epidemiological, clinical and neuropathological traits of AD, PD, ALS and bvFTD. 4
Table 2. List of primers used for this thesis. 39
Table 3. List of plasmids generated for this thesis. 40
Table 4. List of fly lines generated for this thesis. 41

7. Contributions

Transgenic flies were generated by microinjection by Jacqueline Eßer or BestGene as indicated in the Materials and Methods section. Tessa Supèr helped me to generate the GA400 transgenic fly line, and conducted immunostainings parallel to the ones presented in my thesis to further confirm my results. However, all of the results presented in this thesis were obtained by Javier Morón Oset. Prof. Linda Partridge and Dr. Sebastian Grönke supervised my research.

8. Bibliography

Al-Chalabi, Ammar et al. 2016. “Amyotrophic Lateral Sclerosis: Moving towards a New Classification System.” *The Lancet Neurology*.

Al-Sarraj, Safa et al. 2011. “P62 Positive, TDP-43 Negative, Neuronal Cytoplasmic and Intranuclear Inclusions in the Cerebellum and Hippocampus Define the Pathology of C9orf72-Linked FTLD and MND/ALS.” *Acta Neuropathologica* 122(6): 691–702.

Aladesuyi Arogundade, Olubankole et al. 2019. “Antisense RNA Foci Are Associated with Nucleoli and TDP-43 Mislocalization in C9orf72-ALS/FTD: A Quantitative Study.” *Acta Neuropathologica*.

Albertson, Roger et al. 2013. “Mapping Wolbachia Distributions in the Adult Drosophila Brain.” *Cellular Microbiology* 15(9): 1527–44.

Alzheimer’sAssociation. 2020. “2020 Alzheimer’s Disease Facts and Figures.” *Alzheimer’s and Dementia* 16(3): 391–460.

Aparicio, Ricardo, Anil Rana, and David W. Walker. 2019. “Upregulation of the Autophagy Adaptor P62/SQSTM1 Prolongs Health and Lifespan in Middle-Aged Drosophila.” *Cell Reports*.

Arzberger, Thomas et al. 2013. “Bidirectional Transcripts of the Expanded C9orf72 Hexanucleotide Repeat Are Translated into Aggregating Dipeptide Repeat Proteins.” *Acta neuropathologica*.

Asai, Hirohide et al. 2015. “Depletion of Microglia and Inhibition of Exosome Synthesis Halt Tau Propagation.” *Nature Neuroscience* 18(11): 1584–93.

Atanasio, Amanda et al. 2016. “C9orf72 Ablation Causes Immune Dysregulation Characterized by Leukocyte Expansion, Autoantibody Production, and Glomerulonephropathy in Mice.” *Scientific Reports*.

Avila, Jesús. 2010. “Intracellular and Extracellular Tau.” *Frontiers in Neuroscience*.

Azpurua, Jorge, Rebekah E. Mahoney, and Benjamin A. Eaton. 2018. “Transcriptomics of Aged Drosophila Motor Neurons Reveals a Matrix Metalloproteinase That Impairs Motor Function.” *Aging Cell*.

Babcock, Daniel T., and Barry Ganetzky. 2015. “Transcellular Spreading of Huntingtin Aggregates in the Drosophila Brain.” *Proceedings of the National Academy of Sciences of the United States of America* 112(39): E5427–33.

Bae, Haneui et al. 2016. “Rab3-GEF Controls Active Zone Development at the Drosophila Neuromuscular Junction.” *eNeuro*.

Balendra, Rubika, and Adrian M. Isaacs. 2018. “C9orf72-Mediated ALS and FTD: Multiple Pathways to Disease.” *Nature Reviews Neurology* 14(9): 544–58. http://dx.doi.org/10.1038/s41582-018-0047-2.

Bang, Jee, Salvatore Spina, and Bruce L. Miller. 2015. “Frontotemporal Dementia.” *The Lancet*.

Bartlett, Bryan J. et al. 2011. “P62, Ref(2)P and Ubiquitinated Proteins Are Conserved Markers of Neuronal Aging, Aggregate Formation and Progressive Autophagic Defects.” *Autophagy*.

Bass, Timothy M. et al. 2007. “Optimization of Dietary Restriction Protocols in Drosophila.” *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* 62(10): 1071–81.

Bellingham, Mark C. 2011. “A Review of the Neural Mechanisms of Action and Clinical Efficiency of Riluzole in Treating Amyotrophic Lateral Sclerosis: What Have We Learned

in the Last Decade?" *CNS Neuroscience and Therapeutics*.
Belzil, Veronique V. et al. 2013. "Reduced C9orf72 Gene Expression in C9FTD/ALS Is Caused by Histone Trimethylation, an Epigenetic Event Detectable in Blood." *Acta Neuropathologica*.
Ben-Nissan, Gili, and Michal Sharon. 2014. "Regulating the 20S Proteasome Ubiquitin-Independent Degradation Pathway." *Biomolecules*.
Bengtson, J, and Toshihiro Kitamoto. 2001. "Conditional Modification of Behavior by Targeted Expression of a Temperature-Sensitive Shibire Allele in Subsets of Neurons." *CSH Neurobiology of Drosophila*.
Bensimon, G., L. Lacomblez, and V. Meininger. 1994. "A Controlled Trial of Riluzole in Amyotrophic Lateral Sclerosis." *New England Journal of Medicine*.
Bertram, Lars, and Rudolph E Tanzi. 2005. "Review Series The Genetic Epidemiology of Neurodegenerative Disease." 115(6): 1449–57.
Bischof, Johannes et al. 2007. "An Optimized Transgenesis System for Drosophila Using Germ-Line-Specific ΦC31 Integrases." *Proceedings of the National Academy of Sciences of the United States of America* 104(9): 3312–17.
van Blitterswijk, Marka et al. 2013. "Association between Repeat Sizes and Clinical and Pathological Characteristics in Carriers of C9ORF72 Repeat Expansions (Xpansize-72): A Cross-Sectional Cohort Study." *The Lancet Neurology* 12(10): 978–88. http://dx.doi.org/10.1016/S1474-4422(13)70210-2.
Borroni, Barbara et al. 2008. "Education Plays a Different Role in Frontotemporal Dementia and Alzheimer's Disease." *International Journal of Geriatric Psychiatry*.
Boxer, Adam L. et al. 2013. "Memantine in Patients with Frontotemporal Lobar Degeneration: A Multicentre, Randomised, Double-Blind, Placebo-Controlled Trial." *The Lancet Neurology*.
Braak, H., and E. Braak. 1991. "Neuropathological Stageing of Alzheimer-Related Changes." *Acta Neuropathologica*.
Braak, Heiko et al. 2003. "Staging of Brain Pathology Related to Sporadic Parkinson's Disease." *Neurobiology of Aging*.
Brasseur, Laurent et al. 2020. "Dipeptide Repeat Derived from C9orf72 Hexanucleotide Expansions Forms Amyloids or Natively Unfolded Structures in Vitro." *Biochemical and Biophysical Research Communications*.
Brettschneider, Johannes et al. 2013. "Stages of PTDP-43 Pathology in Amyotrophic Lateral Sclerosis." *Annals of Neurology* 74(1): 20–38.
———. 2014. "Sequential Distribution of PTDP-43 Pathology in Behavioral Variant Frontotemporal Dementia (BvFTD)." *Acta Neuropathologica*.
Brettschneider, Johannes, Kelly Del Tredici, Virginia M.Y. Lee, and John Q. Trojanowski. 2015. "Spreading of Pathology in Neurodegenerative Diseases: A Focus on Human Studies." *Nature Reviews Neuroscience* 16(2): 109–20. http://dx.doi.org/10.1038/nrn3887.
Brichta, Lars, and Paul Greengard. 2014. "Molecular Determinants of Selective Dopaminergic Vulnerability in Parkinson's Disease: An Update." *Frontiers in Neuroanatomy*.
Broughton, Susan J., Toshihiro Kitamoto, and Ralph J. Greenspan. 2004. "Excitatory and Inhibitory Switches for Courtship in the Brain of Drosophila Melanogaster." *Current Biology*.
Brown, Rebecca C, Alan H Lockwood, and Babasaheb R Sonawane. 2005. "Neurodegenerative Diseases : An Overview of Environmental Risk Factors." 1250(9): 1250–56.
Buch, Susanne et al. 2008. "Opposing Effects of Dietary Protein and Sugar Regulate a

Transcriptional Target of Drosophila Insulin-like Peptide Signaling." *Cell Metabolism*.
Burberry, Aaron et al. 2016. "Loss-of-Function Mutations in the C9ORF72 Mouse Ortholog Cause Fatal Autoimmune Disease." *Science Translational Medicine*.
———. 2020. "C9orf72 Suppresses Systemic and Neural Inflammation Induced by Gut Bacteria." *Nature*.
Calafate, Sara, William Flavin, Patrik Verstreken, and Diederik Moechars. 2016. "Loss of Bin1 Promotes the Propagation of Tau Pathology." *Cell Reports*.
De Calignon, Alix et al. 2012. "Propagation of Tau Pathology in a Model of Early Alzheimer's Disease." *Neuron*.
Callister, Janis Bennion et al. 2016. "Modelling C9orf72 Dipeptide Repeat Proteins of a Physiologically Relevant Size." *Human Molecular Genetics*.
Chang, Yu Jen et al. 2016. "The Glycine-Alanine Dipeptide Repeat from C9 or F72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties." *Journal of Biological Chemistry* 291(10): 4903–11.
Chen, Shirui et al. 2015. "Mutational Analysis of Rab3 Function for Controlling Active Zone Protein Composition at the Drosophila Neuromuscular Junction." *PLoS ONE*.
Chen, Xiaoying, Jennica L. Zaro, and Wei Chiang Shen. 2013. "Fusion Protein Linkers: Property, Design and Functionality." *Advanced Drug Delivery Reviews*.
Chen, Ying Da et al. 2017. "Exophagy of Annexin A2 via RAB11, RAB8A and RAB27A in IFN-γ-Stimulated Lung Epithelial Cells." *Scientific Reports*.
Cheng, Hsiao Chun, Christina M. Ulane, and Robert E. Burke. 2010. "Clinical Progression in Parkinson Disease and the Neurobiology of Axons." *Annals of Neurology*.
Chiò, Adriano et al. 2020. "ALS Phenotype Is Influenced by Age, Sex, and Genetics: A Population-Based Study." *Neurology*.
Chitiprolu, Maneka et al. 2018. "A Complex of C9ORF72 and P62 Uses Arginine Methylation to Eliminate Stress Granules by Autophagy." *Nature Communications*.
Chong, Samuel S. et al. 1995. "Gametic and Somatic Tissue–Specific Heterogeneity of the Expanded SCA1 CAG Repeat in Spinocerebellar Ataxia Type 1." *Nature Genetics*.
Chou, Ching Chieh et al. 2018. "TDP-43 Pathology Disrupts Nuclear Pore Complexes and Nucleocytoplasmic Transport in ALS/FTD." *Nature Neuroscience*.
Clavaguera, Florence et al. 2013. "Brain Homogenates from Human Tauopathies Induce Tau Inclusions in Mouse Brain." *Proceedings of the National Academy of Sciences of the United States of America*.
Coelho, Dina S. et al. 2018. "Culling Less Fit Neurons Protects against Amyloid-β-Induced Brain Damage and Cognitive and Motor Decline." *Cell Reports* 25(13): 3661-3673.e3. https://doi.org/10.1016/j.celrep.2018.11.098.
Conlon, Erin G. et al. 2016. "The C9ORF72 GGGGCC Expansion Forms RNA G-Quadruplex Inclusions and Sequesters HnRNP H to Disrupt Splicing in ALS Brains." *eLife* 5(September2016): 1–28.
Cooper-Knock, Johnathan et al. 2014. "Sequestration of Multiple RNA Recognition Motif-Containing Proteins by C9orf72 Repeat Expansions." *Brain* 137(7): 2040–51.
———. 2015. "Antisense RNA Foci in the Motor Neurons of C9ORF72-ALS Patients Are Associated with TDP-43 Proteinopathy." *Acta Neuropathologica*.
Cooper-Knock, Johnathan, Pamela J. Shaw, and Janine Kirby. 2014. "The Widening Spectrum of C9ORF72-Related Disease; Genotype/Phenotype Correlations and Potential Modifiers of Clinical Phenotype." *Acta Neuropathologica*.
Cragnaz, Lucia et al. 2014. "Aggregate Formation Prevents DTDP-43 Neurotoxicity in the Drosophila Melanogaster Eye." *Neurobiology of Disease*.

Curran, Jack A., Edgar Buhl, Krasimira Tsaneva-Atanasova, and James J.L. Hodge. 2019. "Age-Dependent Changes in Clock Neuron Structural Plasticity and Excitability Are Associated with a Decrease in Circadian Output Behavior and Sleep." *Neurobiology of Aging*.

Darling, April L. et al. 2019. "Repeated Repeat Problems: Combinatorial Effect of C9orf72-Derived Dipeptide Repeat Proteins." *International Journal of Biological Macromolecules* 127: 136–45. https://doi.org/10.1016/j.ijbiomac.2019.01.035.

Dean, Douglas C. et al. 2014. "Brain Differences in Infants at Differential Genetic Risk for Late-Onset Alzheimer Disease: A Cross-Sectional Imaging Study." *JAMA Neurology* 71(1): 11–22.

DeJesus-Hernandez, Mariely et al. 2011. "Expanded GGGGCC Hexanucleotide Repeat in Non-Coding Region of C9ORF72 Causes Chromosome 9p-Linked Frontotemporal Dementia and Amyotrophic Lateral Sclerosis." *Neuron* 72(2): 245–56. http://dx.doi.org/10.1016/j.neuron.2011.09.011%5Cnpapers3://publication/doi/10.1016/j.neuron.2011.09.011.

Deutsch, Mariel B., Mario F. Mendez, and Edmond Teng. 2015. "Interactions between Traumatic Brain Injury and Frontotemporal Degeneration." *Dementia and Geriatric Cognitive Disorders*.

Devenney, Emma et al. 2014. " Frontotemporal Dementia Associated With the C9ORF72 Mutation ." *JAMA Neurology*.

DiLuca, Monica, and Jes Olesen. 2014. "The Cost of Brain Diseases: A Burden or a Challenge?" *Neuron* 82(6): 1205–8. http://dx.doi.org/10.1016/j.neuron.2014.05.044.

Dols-Icardo, Oriol et al. 2014. "Characterization of the Repeat Expansion Size in C9orf72 in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia." *Human Molecular Genetics*.

Donnelly, Christopher J. et al. 2013. "RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention." *Neuron*.

Duffy, Joseph B. 2002. "GAL4 System in Drosophila: A Fly Geneticist's Swiss Army Knife." *Genesis*.

Duyao, M. et al. 1993. "Trinucleotide Repeat Length Instability and Age of Onset in Huntington's Disease." *Nature Genetics*.

Edbauer, Dieter, and Christian Haass. 2016. "An Amyloid-like Cascade Hypothesis for C9orf72 ALS/FTD." *Current Opinion in Neurobiology* 36: 99–106. http://dx.doi.org/10.1016/j.conb.2015.10.009.

Elobeid, Adila et al. 2016. "Altered Proteins in the Aging Brain." *Journal of Neuropathology and Experimental Neurology* 75(4): 316–25.

Emmanouilidou, Evangelia et al. 2010. "Cell-Produced α-Synuclein Is Secreted in a Calcium-Dependent Manner by Exosomes and Impacts Neuronal Survival." *Journal of Neuroscience*.

Evans, Lewis D. et al. 2018. "Extracellular Monomeric and Aggregated Tau Efficiently Enter Human Neurons through Overlapping but Distinct Pathways." *Cell Reports*.

Flores, Brittany N. et al. 2016. "Distinct C9orf72-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity." *PLoS ONE*.

Fournier, Clémence et al. 2019. "Relations between C9orf72 Expansion Size in Blood, Age at Onset, Age at Collection and Transmission across Generations in Patients and Presymptomatic Carriers." *Neurobiology of Aging* 74: 234.e1-234.e8.

Franzmeier, Nicolai et al. 2020. "Functional Brain Architecture Is Associated with the Rate of Tau Accumulation in Alzheimer's Disease." *Nature Communications* 11(1): 1–17.

Freer, Rosie et al. 2016. "A Protein Homeostasis Signature in Healthy Brains Recapitulates

Tissue Vulnerability to Alzheimer's Disease." *Science Advances*.
Freibaum, Brian D. et al. 2015. "GGGGCC Repeat Expansion in C9orf72 Compromises Nucleocytoplasmic Transport." *Nature* 525(7567): 129–33.
Freibaum, Brian D., and J. Paul Taylor. 2017. "The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD." *Frontiers in Molecular Neuroscience*.
Fu, Hongjun et al. 2019. "A Tau Homeostasis Signature Is Linked with the Cellular and Regional Vulnerability of Excitatory Neurons to Tau Pathology." *Nature Neuroscience*.
Fu, Hongjun, John Hardy, and Karen E. Duff. 2018. "Selective Vulnerability in Neurodegenerative Diseases." *Nature Neuroscience*.
Gao, Ju et al. 2018. "Pathomechanisms of TDP-43 in Neurodegeneration." *Journal of Neurochemistry*.
Gendron, Tania F. et al. 2013. "Antisense Transcripts of the Expanded C9ORF72 Hexanucleotide Repeat Form Nuclear RNA Foci and Undergo Repeat-Associated Non-ATG Translation in C9FTD/ALS." *Acta Neuropathologica* 126(6): 829–44.
Gijselinck, I. et al. 2016. "The C9orf72 Repeat Size Correlates with Onset Age of Disease, DNA Methylation and Transcriptional Downregulation of the Promoter." *Molecular Psychiatry*.
Gittings, Lauren M. et al. 2020. "Symmetric Dimethylation of Poly-GR Correlates with Disease Duration in C9orf72 FTLD and ALS and Reduces Poly-GR Phase Separation and Toxicity." *Acta Neuropathologica*.
Glass, Christopher K. et al. 2010. "Mechanisms Underlying Inflammation in Neurodegeneration." *Cell*.
Gossye, Helena, Christine Van Broeckhoven, and Sebastiaan Engelborghs. 2019. "The Use of Biomarkers and Genetic Screening to Diagnose Frontotemporal Dementia: Evidence and Clinical Implications." *Frontiers in Genetics*.
Gottlieb, Roberta A., Allen M. Andres, Jon Sin, and David P.J. Taylor. 2015. "Untangling Autophagy Measurements All Fluxed Up." *Circulation Research*.
Graf, Ethan R. et al. 2009. "Rab3 Dynamically Controls Protein Composition at Active Zones." *Neuron*.
Graff-Radford, Neill R., and Bryan K. Woodruff. 2007. "Frontotemporal Dementia." *Seminars in Neurology*.
Green, Katelyn M. et al. 2017. "RAN Translation at C9orf72-Associated Repeat Expansions Is Selectively Enhanced by the Integrated Stress Response." *Nature Communications*.
Guo, Qiang et al. 2018. "In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment." *Cell* 172(4): 696-705.e12.
Guo, Yaxin et al. 2019. "Effects of Exosomes on Pre-Metastatic Niche Formation in Tumors." *Molecular Cancer*.
Hara, Taichi et al. 2006. "Suppression of Basal Autophagy in Neural Cells Causes Neurodegenerative Disease in Mice." *Nature*.
Harms, Matthew B. et al. 2013. "Lack of C9ORF72 Coding Mutations Supports a Gain of Function for Repeat Expansions in Amyotrophic Lateral Sclerosis." *Neurobiology of Aging*.
Hasegawa, Masato et al. 2008. "Phosphorylated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis." *Annals of Neurology*.
He, Fang et al. 2020. "The Carboxyl Termini of RAN Translated GGGGCC Nucleotide Repeat Expansions Modulate Toxicity in Models of ALS/FTD." *Acta Neuropathologica Communications*.
He, Zhuohao et al. 2018. "Amyloid-β Plaques Enhance Alzheimer's Brain Tau-Seeded

Pathologies by Facilitating Neuritic Plaque Tau Aggregation." *Nature Medicine*.
Higginbottom, Adrian et al. 2013. "C9ORF72 Transcription in a Frontotemporal Dementia Case with Two Expanded Alleles." *Neurology*.
Van Hoecke, Annelies et al. 2012. "EPHA4 Is a Disease Modifier of Amyotrophic Lateral Sclerosis in Animal Models and in Humans." *Nature Medicine*.
Höglinger, G. U., G. Respondek, and G. G. Kovacs. 2018. "New Classification of Tauopathies." *Revue Neurologique*.
Holmes, Brandon B. et al. 2013. "Heparan Sulfate Proteoglycans Mediate Internalization and Propagation of Specific Proteopathic Seeds." *Proceedings of the National Academy of Sciences of the United States of America*.
Hou, Yujun et al. 2019. "Ageing as a Risk Factor for Neurodegenerative Disease." *Nature Reviews Neurology* 15(10): 565–81. http://dx.doi.org/10.1038/s41582-019-0244-7.
Humphrey, Jack et al. 2017. "Quantitative Analysis of Cryptic Splicing Associated with TDP-43 Depletion." *BMC Medical Genomics*.
Hutton, M. et al. 1998. "Association of Missense and 5'-Splice-Site Mutations in Tau with the Inherited Dementia FTDP-17." *Nature*.
Igaz, Lionel M. et al. 2009. "Expression of TDP-43 C-Terminal Fragments in Vitro Recapitulates Pathological Features of TDP-43 Proteinopathies." *Journal of Biological Chemistry*.
Jadhav, Santosh et al. 2019. "A Walk through Tau Therapeutic Strategies." *Acta neuropathologica communications*.
Jaspert, Andrea, Raimund Fahsold, Holger Grehl, and Detlef Claus. 1995. "Myotonic Dystrophy: Correlation of Clinical Symptoms with the Size of the CTG Trinucleotide Repeat." *Journal of Neurology*.
Javed, Kinaan, and Daniel T. Daly. 2019. StatPearls *Neuroanatomy, Lower Motor Neuron Lesion*.
Jensen, Brigid K et al. 2020. " Synaptic Dysfunction Induced by Glycine-alanine Dipeptides in C9orf72- ALS / FTD Is Rescued by SV 2 Replenishment ." *EMBO Molecular Medicine*.
Jiang, Jie et al. 2016. "Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs." *Neuron*.
Johnson, Brian S. et al. 2009. "TDP-43 Is Intrinsically Aggregation-Prone, and Amyotrophic Lateral Sclerosis-Linked Mutations Accelerate Aggregation and Increase Toxicity." *Journal of Biological Chemistry*.
Jones, Lesley, Henry Houlden, and Sarah J. Tabrizi. 2017. "DNA Repair in the Trinucleotide Repeat Disorders." *The Lancet Neurology*.
Jovičič, Ana et al. 2015. "Modifiers of C9orf72 Dipeptide Repeat Toxicity Connect Nucleocytoplasmic Transport Defects to FTD/ALS." *Nature Neuroscience* 18(9): 1226–29.
Kanekura, Kohsuke et al. 2016. "Poly-Dipeptides Encoded by the C9ORF72 Repeats Block Global Protein Translation." *Human Molecular Genetics*.
Kaniyappan, Senthilvelrajan et al. 2020. "FRET-Based Tau Seeding Assay Does Not Represent Prion-like Templated Assembly of Tau Filaments." *Molecular Neurodegeneration*.
Kanning, Kevin C., Artem Kaplan, and Christopher E. Henderson. 2010. "Motor Neuron Diversity in Development and Disease." *Annual Review of Neuroscience*.
Kaplan, Artem et al. 2014. "Neuronal Matrix Metalloproteinase-9 Is a Determinant of Selective Neurodegeneration." *Neuron*.
Katsinelos, Taxiarchis et al. 2018. "Unconventional Secretion Mediates the Trans-Cellular Spreading of Tau." *Cell Reports*.

Kaufman, Sarah K. et al. 2016. “Tau Prion Strains Dictate Patterns of Cell Pathology, Progression Rate, and Regional Vulnerability In Vivo.” *Neuron*.
Kawasaki, Fumiko, and Richard W. Ordway. 2009. “Molecular Mechanisms Determining Conserved Properties of Short-Term Synaptic Depression Revealed in NSF and SNAP-25 Conditional Mutants.” *Proceedings of the National Academy of Sciences of the United States of America*.
Kfoury, Najla et al. 2012. “Trans-Cellular Propagation of Tau Aggregation by Fibrillar Species.” *Journal of Biological Chemistry*.
Khosravi, Bahram et al. 2017. “Cytoplasmic Poly-GA Aggregates Impair Nuclear Import of TDP-43 in C9orf72 ALS/FTLD.” *Human Molecular Genetics* 26(4): 790–800.
———. 2020. “ Cell-to-cell Transmission of C9orf72 Poly-(Gly-Ala) Triggers Key Features of ALS / FTD .” *The EMBO Journal*.
Kim, Eun Joo et al. 2012. “Selective Frontoinsular von Economo Neuron and Fork Cell Loss in Early Behavioral Variant Frontotemporal Dementia.” *Cerebral Cortex*.
King, Thomas D., Justine E. Johnson, and Jack R. Bateman. 2019. “Position Effects Influence Transvection in Drosophila Melanogaster.” *Genetics*.
Klein, Christine, and Ana Westenberger. 2012. “Genetics of Parkinson’s Disease.” *Cold Spring Harbor Perspectives in Medicine*.
Koles, Kate et al. 2012. “Mechanism of Evenness Interrupted (Evi)-Exosome Release at Synaptic Boutons.” *Journal of Biological Chemistry*.
Kreher, Scott A., Jae Young Kwon, and John R. Carlson. 2005. “The Molecular Basis of Odor Coding in the Drosophila Larva.” *Neuron* 46(3): 445–56.
Kumar, David R., Florence Aslinia, Steven H. Yale, and Joseph J. Mazza. 2011. “Jean-Martin Charcot: The Father of Neurology.” *Clinical Medicine and Research*.
Kwon, Ilmin et al. 2014. “Poly-Dipeptides Encoded by the C9orf72 Repeats Bind Nucleoli, Impede RNA Biogenesis, and Kill Cells.” *Science* 345(6201): 1139–45.
Laaksovirta, Hannu et al. 2010. “Chromosome 9p21 in Amyotrophic Lateral Sclerosis in Finland: A Genome-Wide Association Study.” *The Lancet Neurology*.
LaClair, Katherine D. et al. 2020. “Congenic Expression of Poly-GA but Not Poly-PR in Mice Triggers Selective Neuron Loss and Interferon Responses Found in C9orf72 ALS.” *Acta Neuropathologica*.
Lagier-Tourenne, Clotilde et al. 2013. “Targeted Degradation of Sense and Antisense C9orf72 RNA Foci as Therapy for ALS and Frontotemporal Degeneration.” *Proceedings of the National Academy of Sciences of the United States of America*.
Lamark, Trond, and Terje Johansen. 2012. “Aggrephagy: Selective Disposal of Protein Aggregates by Macroautophagy.” *International Journal of Cell Biology*.
Lasagna-Reeves, Cristian A. et al. 2012. “Alzheimer Brain-Derived Tau Oligomers Propagate Pathology from Endogenous Tau.” *Scientific Reports* 2.
Latouche, Morwena et al. 2007. “A Conditional Pan-Neuronal Drosophila Model of Spinocerebellar Ataxia 7 with a Reversible Adult Phenotype Suitable for Identifying Modifier Genes.” *Journal of Neuroscience*.
Lauwers, Elsa et al. 2018. “Hsp90 Mediates Membrane Deformation and Exosome Release.” *Molecular Cell*.
Lee, Jin Gu et al. 2016. “Unconventional Secretion of Misfolded Proteins Promotes Adaptation to Proteasome Dysfunction in Mammalian Cells.” *Nature Cell Biology* 18(7): 765–76.
Lee, Kyung Ha et al. 2016. “C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles.” *Cell*.

Lee, Sangseok, and Dong Kyu Lee. 2018. “What Is the Proper Way to Apply the Multiple Comparison Test?” *Korean Journal of Anesthesiology*.
Lee, Youn Bok et al. 2013. “Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic.” *Cell Reports* 5(5): 1178–86.
———. 2017. “C9orf72 Poly GA RAN-Translated Protein Plays a Key Role in Amyotrophic Lateral Sclerosis via Aggregation and Toxicity.” *Human Molecular Genetics* 26(24): 4765–77.
Levine, Timothy P. et al. 2013. “The Product of C9orf72, a Gene Strongly Implicated in Neurodegeneration, Is Structurally Related to DENN Rab-GEFs.” *Bioinformatics*.
Liu, Li et al. 2012. “Trans-Synaptic Spread of Tau Pathology in Vivo.” *PLoS ONE*.
López-Otín, Carlos et al. 2013. “The Hallmarks of Aging.” *Cell* 153(6): 1194.
Maass, Anne et al. 2017. “Comparison of Multiple Tau-PET Measures as Biomarkers in Aging and Alzheimer’s Disease.” *NeuroImage*.
MacKenzie, Ian R. et al. 2013. “Dipeptide Repeat Protein Pathology in C9ORF72 Mutation Cases: Clinico-Pathological Correlations.” *Acta Neuropathologica*.
Mackenzie, Ian R., and Manuela Neumann. 2017. “Reappraisal of TDP-43 Pathology in FTLD-U Subtypes.” *Acta Neuropathologica*.
Mackenzie, Ian R.A. et al. 2007. “Pathological TDP-43 Distinguishes Sporadic Amyotrophic Lateral Sclerosis from Amyotrophic Lateral Sclerosis with SOD1 Mutations.” *Annals of Neurology*.
———. 2015. “Quantitative Analysis and Clinico-Pathological Correlations of Different Dipeptide Repeat Protein Pathologies in C9ORF72 Mutation Carriers.” *Acta Neuropathologica* 130(6): 845–61.
Marambaud, Philippe, Ute Dreses-Werringloer, and Valérie Vingtdeux. 2009. “Calcium Signaling in Neurodegeneration.” *Molecular Neurodegeneration*.
Martini-Stoica, Heidi et al. 2018. “TFEB Enhances Astroglial Uptake of Extracellular Tau Species and Reduces Tau Spreading.” *Journal of Experimental Medicine*.
May, Stephanie et al. 2014. “C9orf72 FTLD/ALS-Associated Gly-Ala Dipeptide Repeat Proteins Cause Neuronal Toxicity and Unc119 Sequestration.” *Acta Neuropathologica* 128(4): 485–503.
McEachin, Zachary T. et al. 2020. “Chimeric Peptide Species Contribute to Divergent Dipeptide Repeat Pathology in C9ALS/FTD and SCA36.” *Neuron*.
McGuire, Sean E. et al. 2003. “Spatiotemporal Rescue of Memory Dysfunction in Drosophila.” *Science* 302(5651): 1765–68.
Mejzini, Rita et al. 2019. “ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now?” *Frontiers in Neuroscience*.
Melamed, Ze’ev et al. 2019. “Premature Polyadenylation-Mediated Loss of Stathmin-2 Is a Hallmark of TDP-43-Dependent Neurodegeneration.” *Nature Neuroscience*.
Menzies, Fiona M. et al. 2017. “Autophagy and Neurodegeneration: Pathogenic Mechanisms and Therapeutic Opportunities.” *Neuron* 93(5): 1015–34.
Millecamps, Stéphanie et al. 2012. “Phenotype Difference between ALS Patients with Expanded Repeats in C9ORF72 and Patients with Mutations in Other ALS-Related Genes.” *Journal of Medical Genetics*.
Miller, Timothy M. et al. 2013. “An Antisense Oligonucleotide against SOD1 Delivered Intrathecally for Patients with SOD1 Familial Amyotrophic Lateral Sclerosis: A Phase 1, Randomised, First-in-Man Study.” *The Lancet Neurology*.
Mirbaha, Hilda et al. 2015. “Tau Trimers Are the Minimal Propagation Unit Spontaneously Internalized to Seed Intracellular Aggregation.” *Journal of Biological Chemistry* 290(24):

14893–903.
———. 2018. “Inert and Seed-Competent Tau Monomers Suggest Structural Origins of Aggregation.” *eLife*.
Mizielinska, Sarah et al. 2013. “C9orf72 Frontotemporal Lobar Degeneration Is Characterised by Frequent Neuronal Sense and Antisense RNA Foci.” *Acta Neuropathologica*.
———. 2014. “C9orf72 Repeat Expansions Cause Neurodegeneration in Drosophila through Arginine-Rich Proteins.” *Science* 345(6201): 1192–94.
Moens, Thomas G. et al. 2018. “Sense and Antisense RNA Are Not Toxic in Drosophila Models of C9orf72-Associated ALS/FTD.” *Acta Neuropathologica*.
———. 2019. “C9orf72 Arginine-Rich Dipeptide Proteins Interact with Ribosomal Proteins in Vivo to Induce a Toxic Translational Arrest That Is Rescued by EIF1A.” *Acta Neuropathologica* 137(3): 487–500.
Mori, Kohji, Thomas Arzberger, et al. 2013. “Bidirectional Transcripts of the Expanded C9orf72 Hexanucleotide Repeat Are Translated into Aggregating Dipeptide Repeat Proteins.” *Acta Neuropathologica* 126(6): 881–93.
Mori, Kohji, Shih-ming Weng, et al. 2013. “The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS.” *Science* 339(March): 1335–39. https://science.sciencemag.org/content/339/6125/1335/tab-pdf.
Morón-Oset, Javier et al. 2019. “Glycine-Alanine Dipeptide Repeats Spread Rapidly in a Repeat Length-and Age-Dependent Manner in the Fly Brain.” *Acta Neuropathologica Communications*.
Van Mossevelde, Sara, Julie Van Der Zee, et al. 2017. “Clinical Evidence of Disease Anticipation in Families Segregating a C9 Orf 72 Repeat Expansion.” *JAMA Neurology*.
Van Mossevelde, Sara, Julie van der Zee, Marc Cruts, and Christine Van Broeckhoven. 2017. “Relationship between C9orf72 Repeat Size and Clinical Phenotype.” *Current Opinion in Genetics and Development*.
Mure, Guest Editor D F, and Kurt A Jellinger. 2010. “Basic Mechanisms of Neurodegeneration : A Critical Update.” 14(3): 457–87.
Narasimhan, Sneha et al. 2017. “Pathological Tau Strains from Human Brains Recapitulate the Diversity of Tauopathies in Nontransgenic Mouse Brain.” *Journal of Neuroscience*.
Nassif, Melissa, Ute Woehlbier, and Patricio A. Manque. 2017. “The Enigmatic Role of C9ORF72 in Autophagy.” *Frontiers in Neuroscience*.
Naumann, Marcel et al. 2019. “Phenotypes and Malignancy Risk of Different FUS Mutations in Genetic Amyotrophic Lateral Sclerosis.” *Annals of Clinical and Translational Neurology*.
Nelson, Peter T. et al. 2012. “Correlation of Alzheimer Disease Neuropathologic Changes with Cognitive Status: A Review of the Literature.” *Journal of Neuropathology and Experimental Neurology*.
Neumann, Manuela et al. 2006. “Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis.” *Science*.
Nezis, Ioannis P. et al. 2008. “Ref(2)P, the Drosophila Melanogaster Homologue of Mammalian P62, Is Required for the Formation of Protein Aggregates in Adult Brain.” *Journal of Cell Biology* 180(6): 1065–71.
Ngolab, Jennifer et al. 2017. “Brain-Derived Exosomes from Dementia with Lewy Bodies Propagate α-Synuclein Pathology.” *Acta neuropathologica communications*.
Nguyen, Lien et al. 2019. “Antibody Therapy Targeting RAN Proteins Rescues C9 ALS/FTD Phenotypes in C9orf72 Mouse Model.” *Neuron*: 1–18. https://linkinghub.elsevier.com/retrieve/pii/S0896627319309705.

Nguyen, Lien, John Douglas Cleary, and Laura P.W. Ranum. 2019. “Repeat-Associated Non-ATG Translation: Molecular Mechanisms and Contribution to Neurological Disease.” *Annual Review of Neuroscience*.

Niccoli, Teresa, Linda Partridge, and Adrian M. Isaacs. 2017. “Ageing as a Risk Factor for ALS/FTD.” *Human Molecular Genetics* 26(R2): R105–13.

Nicholson, Alexandra M. et al. 2013. “TMEM106B p.T185S Regulates TMEM106B Protein Levels: Implications for Frontotemporal Dementia.” *Journal of Neurochemistry*.

Nihei, Yoshihiro et al. 2020. “Poly-Glycine–Alanine Exacerbates C9orf72 Repeat Expansion-Mediated DNA Damage via Sequestration of Phosphorylated ATM and Loss of Nuclear HnRNPA3.” *Acta Neuropathologica*.

Nilsson, Christer et al. 2014. “Age-Related Incidence and Family History in Frontotemporal Dementia: Data from the Swedish Dementia Registry.” *PLoS ONE*.

Nitabach, Michael N., Justin Blau, and Todd C. Holmes. 2002. “Electrical Silencing of Drosophila Pacemaker Neurons Stops the Free-Running Circadian Clock.” *Cell*.

Nizzardo, M. et al. 2020. “Synaptotagmin 13 Is Neuroprotective across Motor Neuron Diseases.” *Acta Neuropathologica*.

O’Rourke, J. G. et al. 2016. “C9orf72 Is Required for Proper Macrophage and Microglial Function in Mice.” *Science*.

Onyike, Chiadi U., and Janine Diehl-Schmid. 2013. “The Epidemiology of Frontotemporal Dementia.” *International Review of Psychiatry*.

Oskarsson, Björn, D. Kevin Horton, and Hiroshi Mitsumoto. 2015. “Potential Environmental Factors in Amyotrophic Lateral Sclerosis.” *Neurologic Clinics*.

Osterwalder, Thomas, Kenneth S. Yoon, Benjamin H. White, and Haig Keshishian. 2001. “A Conditional Tissue-Specific Transgene Expression System Using Inducible GAL4.” *Proceedings of the National Academy of Sciences of the United States of America* 98(22): 12596–601.

Palmqvist, Sebastian et al. 2020. “Discriminative Accuracy of Plasma Phospho-Tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders.” *JAMA*.

Pankiv, Serhiy et al. 2007. “P62/SQSTM1 Binds Directly to Atg8/LC3 to Facilitate Degradation of Ubiquitinated Protein Aggregates by Autophagy*[S].” *Journal of Biological Chemistry*.

Panza, Francesco et al. 2020. “Development of Disease-Modifying Drugs for Frontotemporal Dementia Spectrum Disorders.” *Nature Reviews Neurology*.

Panza, Francesco, Madia Lozupone, Giancarlo Logroscino, and Bruno P. Imbimbo. 2019. “A Critical Appraisal of Amyloid-β-Targeting Therapies for Alzheimer Disease.” *Nature Reviews Neurology*.

Partridge, Linda, Joris Deelen, and P. Eline Slagboom. 2018. “Facing up to the Global Challenges of Ageing.” *Nature* 561(7721): 45–56. http://dx.doi.org/10.1038/s41586-018-0457-8.

Pattamatta, Amrutha et al. 2020. “Repeat Length Increases Disease Penetrance and Severity In.” (352): 1–47.

Pattamatta, Amrutha, John D. Cleary, and Laura P.W. Ranum. 2018. “All in the Family: Repeats and ALS/FTD.” *Trends in Neurosciences*.

Peng, Chao, John Q. Trojanowski, and Virginia M.Y. Lee. 2020. “Protein Transmission in Neurodegenerative Disease.” *Nature Reviews Neurology*.

Peters, Owen M. et al. 2015. “Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice.” *Neuron*.

Petersen, Robert B. et al. 1996. “Effect of the D178N Mutation and the Codon 129

Polymorphism on the Metabolism of the Prion Protein." *Journal of Biological Chemistry*.
Pick, A. 1892. "Über Die Beziehungen Der Senilen Hirnatrophie Zur Aphasie." *Prager Medische Wochenschrift*.
Pooler, Amy M. et al. 2013. "Physiological Release of Endogenous Tau Is Stimulated by Neuronal Activity." *EMBO Reports*.
Prusiner, Stanley B. 2001. "Shattuck Lecture - Neurodegenerative Diseases and Prions." *New England Journal of Medicine*.
Purice, Maria D., Sean D. Speese, and Mary A. Logan. 2016. "Delayed Glial Clearance of Degenerating Axons in Aged Drosophila Is Due to Reduced PI3K/Draper Activity." *Nature Communications*.
Raggi, Alberto, and Matilde Leonardi. 2020. "Burden of Brain Disorders in Europe in 2017 and Comparison with Other Non-Communicable Disease Groups." *Journal of Neurology, Neurosurgery and Psychiatry* 91(1): 104–5.
Rauch, Jennifer N. et al. 2020. "LRP1 Is a Master Regulator of Tau Uptake and Spread." *Nature*.
Ren, D. et al. 2001. "A Prokaryotic Voltage-Gated Sodium Channel." *Science*.
Renton, Alan E. et al. 2011. "A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD." *Neuron* 72(2): 257–68.
Robinson, John L. et al. 2018. "Neurodegenerative Disease Concomitant Proteinopathies Are Prevalent, Age-Related and APOE4-Associated." *Brain*.
Roe, Catherine M., Chengjie Xiong, J. Phillip Miller, and John C. Morris. 2007. "Education and Alzheimer Disease without Dementia: Support for the Cognitive Reserve Hypothesis." *Neurology*.
Rohrer, Jonathan D. et al. 2011. "Clinical and Neuroanatomical Signatures of Tissue Pathology in Frontotemporal Lobar Degeneration." *Brain*.
Russ, Jenny et al. 2015. "Hypermethylation of Repeat Expanded C9orf72 Is a Clinical and Molecular Disease Modifier." *Acta Neuropathologica*.
Saberi, Shahram et al. 2018. "Sense-Encoded Poly-GR Dipeptide Repeat Proteins Correlate to Neurodegeneration and Uniquely Co-Localize with TDP-43 in Dendrites of Repeat-Expanded C9orf72 Amyotrophic Lateral Sclerosis." *Acta Neuropathologica* 135(3): 459–74. https://doi.org/10.1007/s00401-017-1793-8.
Saiz-Baggetto, Sara et al. 2017. "Chimeric Proteins Tagged with Specific 3xHA Cassettes May Present Instability and Functional Problems." *PLoS ONE* 12(8): 1–12.
Saman, Sudad et al. 2012. "Exosome-Associated Tau Is Secreted in Tauopathy Models and Is Selectively Phosphorylated in Cerebrospinal Fluid in Early Alzheimer Disease." *Journal of Biological Chemistry*.
Sanders, David W. et al. 2014. "Distinct Tau Prion Strains Propagate in Cells and Mice and Define Different Tauopathies." *Neuron*.
Savić, Dušanka et al. 2002. "250 CTG Repeats in DMPK Is a Threshold for Correlation of Expansion Size and Age at Onset of Juvenile-Adult DM1." *Human Mutation*.
Sawada, Hideyuki. 2017. "Clinical Efficacy of Edaravone for the Treatment of Amyotrophic Lateral Sclerosis." *Expert Opinion on Pharmacotherapy*.
Schludi, Martin H. et al. 2015. "Distribution of Dipeptide Repeat Proteins in Cellular Models and C9orf72 Mutation Cases Suggests Link to Transcriptional Silencing." *Acta Neuropathologica*.
———. 2017. "Spinal Poly-GA Inclusions in a C9orf72 Mouse Model Trigger Motor Deficits and Inflammation without Neuron Loss." *Acta Neuropathologica* 134(2): 241–54.
Schlüter, Oliver M., Jayeeta Basu, Thomas C. Südhof, and Christian Rosenmund. 2006. "Rab3

Superprimes Synaptic Vesicles for Release: Implications for Short-Term Synaptic Plasticity." *Journal of Neuroscience*.
Schwarz, Adam J. et al. 2016. "Regional Profiles of the Candidate Tau PET Ligand 18F-AV-1451 Recapitulate Key Features of Braak Histopathological Stages." *Brain*.
Sellier, Chantal et al. 2016. " Loss of C9 ORF 72 Impairs Autophagy and Synergizes with PolyQ Ataxin-2 to Induce Motor Neuron Dysfunction and Cell Death ." *The EMBO Journal*.
Shahheydari, Hamideh et al. 2017. "Protein Quality Control and the Amyotrophic Lateral Sclerosis/Frontotemporal Dementia Continuum." *Frontiers in Molecular Neuroscience* 10(May): 1–25.
Singh, N. N., M. D. Howell, E. J. Androphy, and R. N. Singh. 2017. "How the Discovery of ISS-N1 Led to the First Medical Therapy for Spinal Muscular Atrophy." *Gene Therapy*.
Slack, Cathy et al. 2011. "DFOXO-Independent Effects of Reduced Insulin-like Signaling in Drosophila." *Aging Cell*.
Smolek, Tomas et al. 2019. "First-in-Rat Study of Human Alzheimer's Disease Tau Propagation." *Molecular Neurobiology*.
Snell, Russell G. et al. 1993. "Relationship between Trinucleotide Repeat Expansion and Phenotypic Variation in Huntington's Disease." *Nature Genetics*.
Solomon, Daniel A. et al. 2018. "A Feedback Loop between Dipeptide-Repeat Protein, TDP-43 and Karyopherin-α Mediates C9orf72-Related Neurodegeneration." *Brain*.
Starr, Alexander, and Rita Sattler. 2018. "Synaptic Dysfunction and Altered Excitability in C9ORF72 ALS/FTD." *Brain Research*.
Sunderhaus, Elizabeth R., and Doris Kretzschmar. 2016. "Mass Histology to Quantify Neurodegeneration in Drosophila." *Journal of Visualized Experiments*.
Surmeier, D. James, José A. Obeso, and Glenda M. Halliday. 2017. "Selective Neuronal Vulnerability in Parkinson Disease." *Nature Reviews Neuroscience*.
Swinnen, Bart et al. 2018. "A Zebrafish Model for C9orf72 ALS Reveals RNA Toxicity as a Pathogenic Mechanism." *Acta Neuropathologica*.
Tabet, Ricardos et al. 2018. "CUG Initiation and Frameshifting Enable Production of Dipeptide Repeat Proteins from ALS/FTD C9ORF72 Transcripts." *Nature Communications*.
Tain, Luke S et al. 2017. "A Proteomic Atlas of Insulin Signalling Reveals Tissue-specific Mechanisms of Longevity Assurance." *Molecular Systems Biology*.
Takatori, Sho, Wenbo Wang, Akihiro Iguchi, and Taisuke Tomita. 2019. "Genetic Risk Factors for Alzheimer Disease: Emerging Roles of Microglia in Disease Pathomechanisms." In *Advances in Experimental Medicine and Biology*,.
Tao, Zhouteng et al. 2015. "Nucleolar Stress and Impaired Stress Granule Formation Contribute to C9orf72 RAN Translation-Induced Cytotoxicity." *Human Molecular Genetics* 24(9): 2426–41.
Tardivel, Meryem et al. 2016. "Tunneling Nanotube (TNT)-Mediated Neuron-to Neuron Transfer of Pathological Tau Protein Assemblies." *Acta neuropathologica communications*.
Taylor, J. Paul, Robert H. Brown, and Don W. Cleveland. 2016. "Decoding ALS: From Genes to Mechanism." *Nature*.
Thies, William, and Laura Bleiler. 2012. "2012 Alzheimer's Disease Facts and Figures." *Alzheimer's and Dementia*.
Tower, John et al. 2017. "Mifepristone/RU486 Acts in Drosophila Melanogaster Females to Counteract the Life Span-Shortening and pro-Inflammatory Effects of Male Sex Peptide." *Biogerontology*.

Tran, Helene et al. 2015. “Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS.” *Neuron* 87(6): 1207–14.
Tsakiri, Eleni N. et al. 2019. “Proteasome Dysfunction Induces Excessive Proteome Instability and Loss of Mitostasis That Can Be Mitigated by Enhancing Mitochondrial Fusion or Autophagy.” *Autophagy*.
Tsvetkov, Andrey S. et al. 2013. “Proteostasis of Polyglutamine Varies among Neurons and Predicts Neurodegeneration.” *Nature Chemical Biology*.
Ulusoy, Ayse et al. 2013. “Caudo-Rostral Brain Spreading of α-Synuclein through Vagal Connections.” *EMBO Molecular Medicine*.
United Nations, World Population Ageing, 2019. 2019. Economic and Social Affairs, Population Division *World Population Ageing 2019*. http://link.springer.com/chapter/10.1007/978-94-007-5204-7_6.
Vance, Caroline et al. 2006. “Familial Amyotrophic Lateral Sclerosis with Frontotemporal Dementia Is Linked to a Locus on Chromosome 9p13.2-21.3.” *Brain*.
———. 2009. “Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6.” *Science*.
Vatsavayai, Sarat C. et al. 2016. “Timing and Significance of Pathological Features in C9orf72 Expansion-Associated Frontotemporal Dementia.” *Brain* 139(12): 3202–16.
Velentzas, Panagiotis D. et al. 2013. “Detrimental Effects of Proteasome Inhibition Activity in Drosophila Melanogaster: Implication of ER Stress, Autophagy, and Apoptosis.” *Cell Biology and Toxicology*.
Verkerk, Annemiske J.M.H. et al. 1991. “Identification of a Gene (FMR-1) Containing a CGG Repeat Coincident with a Breakpoint Cluster Region Exhibiting Length Variation in Fragile X Syndrome.” *Cell*.
Vernace, Vita A., Lisette Arnaud, Thomas Schmidt-Glenewinkel, and Maria E. Figueiredo-Pereira. 2007. “ Aging Perturbs 26S Proteasome Assembly in Drosophila Melanogaster .” *The FASEB Journal*.
Viodé, Arthur et al. 2018. “New Antibody-Free Mass Spectrometry-Based Quantification Reveals That C9ORF72 Long Protein Isoform Is Reduced in the Frontal Cortex of Hexanucleotide-Repeat Expansion Carriers.” *Frontiers in Neuroscience*.
Vogel, Jacob W. et al. 2020. “Spread of Pathological Tau Proteins through Communicating Neurons in Human Alzheimer’s Disease.” *Nature communications* 11(1): 2612.
Vogels, Thomas et al. 2020. “Propagation of Tau Pathology: Integrating Insights From Postmortem and In Vivo Studies.” *Biological Psychiatry*.
Walsh, Dominic M., and Dennis J. Selkoe. 2016. “A Critical Appraisal of the Pathogenic Protein Spread Hypothesis of Neurodegeneration.” *Nature Reviews Neuroscience*.
Wang, Liang et al. 2016. “Evaluation of Tau Imaging in Staging Alzheimer Disease and Revealing Interactions between β-Amyloid and Tauopathy.” *JAMA Neurology* 73(9): 1070–77.
Wang, Yipeng et al. 2017. “The Release and Trans-Synaptic Transmission of Tau via Exosomes.” *Molecular Neurodegeneration*.
Webster, Christopher P et al. 2016. “ The C9orf72 Protein Interacts with Rab1a and the ULK 1 Complex to Regulate Initiation of Autophagy .” *The EMBO Journal*.
Wegmann, Susanne et al. 2015. “Removing Endogenous Tau Does Not Prevent Tau Propagation yet Reduces Its Neurotoxicity.” *The EMBO Journal*.
———. 2019. “Experimental Evidence for the Age Dependence of Tau Protein Spread in the Brain.” *Science Advances* 5(6).
Wen, Xinmei et al. 2014. “Antisense Proline-Arginine RAN Dipeptides Linked to C9ORF72-

ALS/FTD Form Toxic Nuclear Aggregates That Initiate Invitro and Invivo Neuronal Death." *Neuron* 84(6): 1213–25.
Westergard, Thomas et al. 2016. "Cell-to-Cell Transmission of Dipeptide Repeat Proteins Linked to C9orf72-ALS/FTD." *Cell Reports*.
Whitmer, R. A. et al. 2005. "Midlife Cardiovascular Risk Factors and Risk of Dementia in Late Life." *Neurology*.
Wong, L. J.C. et al. 1995. "Somatic Heterogeneity of the CTG Repeat in Myotonic Dystrophy Is Age and Size Dependent." *American Journal of Human Genetics*.
Wu, Jessica W. et al. 2013. "Small Misfolded Tau Species Are Internalized via Bulk Endocytosis and Anterogradely and Retrogradely Transported in Neurons." *Journal of Biological Chemistry*.
———. 2016. "Neuronal Activity Enhances Tau Propagation and Tau Pathology in Vivo." *Nature Neuroscience*.
Wu, Shih Cheng, Zih Syuan Cao, Kuo Ming Chang, and Jyh Lyh Juang. 2017. "Intestinal Microbial Dysbiosis Aggravates the Progression of Alzheimer's Disease in Drosophila." *Nature Communications* 8(1): 1–8. http://dx.doi.org/10.1038/s41467-017-00040-6.
Wyss-Coray, Tony. 2016. "Ageing, Neurodegeneration and Brain Rejuvenation." *Nature*.
Xi, Zhengrui et al. 2014. "Identical Twins with the C9ORF72 Repeat Expansion Are Discordant for ALS." *Neurology*.
Xiao, Shangxi et al. 2015. "Isoform-Specific Antibodies Reveal Distinct Subcellular Localizations of C9orf72 in Amyotrophic Lateral Sclerosis." *Annals of Neurology*.
Xu, Wangchao, and Jin Xu. 2018. "C9orf72 Dipeptide Repeats Cause Selective Neurodegeneration and Cell-Autonomous Excitotoxicity in Drosophila Glutamatergic Neurons." *Journal of Neuroscience*.
Xu, Zihui et al. 2013. "Expanded GGGGCC Repeat RNA Associated with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Causes Neurodegeneration." *Proceedings of the National Academy of Sciences of the United States of America*.
Yamada, Kaoru et al. 2014. "Neuronal Activity Regulates Extracellular Tau in Vivo." *Journal of Experimental Medicine*.
Yamada, Kaoru, and Takeshi Iwatsubo. 2018. "Extracellular α-Synuclein Levels Are Regulated by Neuronal Activity." *Molecular Neurodegeneration*.
Yang, Dejun et al. 2015. "FTD/ALS-Associated Poly(GR) Protein Impairs the Notch Pathway and Is Recruited by Poly(GA) into Cytoplasmic Inclusions." *Acta Neuropathologica*.
Zhang, Ke et al. 2015. "The C9orf72 Repeat Expansion Disrupts Nucleocytoplasmic Transport." *Nature* 525(7567): 56–61.
Zhang, Man et al. 2019. "Edaravone Attenuates Traumatic Brain Injury through Anti-inflammatory and Anti-oxidative Modulation." *Experimental and Therapeutic Medicine*.
Zhang, Yong Jie et al. 2014. "Aggregation-Prone C9FTD/ALS Poly(GA) RAN-Translated Proteins Cause Neurotoxicity by Inducing ER Stress." *Acta Neuropathologica* 128(4): 505–24.
———. 2016. "C9ORF72 Poly(GA) Aggregates Sequester and Impair HR23 and Nucleocytoplasmic Transport Proteins." *Nature Neuroscience*.
———. 2018. "Poly(GR) Impairs Protein Translation and Stress Granule Dynamics in C9orf72-Associated Frontotemporal Dementia and Amyotrophic Lateral Sclerosis." *Nature Medicine*.
———. 2019. "Heterochromatin Anomalies and Double-Stranded RNA Accumulation Underlie C9orf72 Poly(PR) Toxicity." *Science*.

Zhang, Yong Q., Christopher K. Rodesch, and Kendal Broadie. 2002. “Living Synaptic Vesicle Marker: Synaptotagmin-GFP.” *Genesis* 34(1–2): 142–45.

Zhou, Qihui et al. 2017. “ Antibodies Inhibit Transmission and Aggregation of C9orf72 Poly-GA Dipeptide Repeat Proteins .” *EMBO Molecular Medicine*.

———. 2020. “ Active Poly-GA Vaccination Prevents Microglia Activation and Motor Deficits in a C9orf72 Mouse Model .” *EMBO Molecular Medicine* 12(2): 1–13.

Zhu, Qiang et al. 2020. “Reduced C9ORF72 Function Exacerbates Gain of Toxicity from ALS/FTD-Causing Repeat Expansion in C9orf72.” *Nature Neuroscience*.

Zu, Tao et al. 2011. “Non-ATG-Initiated Translation Directed by Microsatellite Expansions.” *Proceedings of the National Academy of Sciences of the United States of America*.

———. 2013. “RAN Proteins and RNA Foci from Antisense Transcripts in C9ORF72 ALS and Frontotemporal Dementia.” *Proceedings of the National Academy of Sciences of the United States of America* 110(51).

9. Signed thesis declaration

Erklärung zur Dissertation
gemäß der Promotionsordnung vom 12. März 2020

Diese Erklärung muss in der Dissertation enthalten sein.
(This version must be included in the doctoral thesis)

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht."

Teilpublikationen:

- Morón-Oset et al, 2019, Acta Neuropathol Comm.

Datum, Name und Unterschrift

07.09.2020, Javier Morón Oset, Javier Morón Oset

10. Publications

As indicated in section 3.1, part of the results presented in that section were published in 2019 in Acta Neuropathologica Communications. Along with Prof. Dr. Linda Partridge and Dr. Sebastian Grönke, Javier Morón Oset (JMO) designed the study. In addition, JMO performed cloning and fly crossings. Moreover, along with Tessa Supèr, JMO performed immunostainings, imaging and analysis. Together with Prof. Dr. Linda Partridge, Dr. Sebastian Grönke and Prof. Dr. Adrian M Isaacs, JMO wrote the manuscript. All authors read and approved the final manuscript.

The published manuscript has been added to this thesis to facilitate consultation (next page).

Morón-Oset J, Supèr T, Esser J, Isaacs AI, Groenke S & Partridge L. (2019) Glycine-alanine dipeptide repeats spread rapidly in a repeat length- and age-dependent manner in the fly brain. Acta Neuropathol Commun., 7: 209.

Acta Neuropathologica Communications

RESEARCH Open Access

Glycine-alanine dipeptide repeats spread rapidly in a repeat length- and age-dependent manner in the fly brain

Javier Morón-Oset[1], Tessa Supèr[1], Jacqueline Esser[1], Adrian M. Isaacs[2], Sebastian Grönke[1] and Linda Partridge[1,3*]

Abstract

Hexanucleotide repeat expansions of variable size in *C9orf72* are the most prevalent genetic cause of amyotrophic lateral sclerosis and frontotemporal dementia. Sense and antisense transcripts of the expansions are translated by repeat-associated non-AUG translation into five dipeptide repeat proteins (DPRs). Of these, the polyGR, polyPR and, to a lesser extent, polyGA DPRs are neurotoxic, with polyGA the most abundantly detected DPR in patient tissue. Trans-cellular transmission of protein aggregates has recently emerged as a major driver of toxicity in various neurodegenerative diseases. In vitro evidence suggests that the C9 DPRs can spread. However, whether this phenomenon occurs under more complex in vivo conditions remains unexplored. Here, we used the adult fly brain to investigate whether the C9 DPRs can spread in vivo upon expression in a subset of neurons. We found that only polyGA can progressively spread throughout the brain, which accumulates in the shape of aggregate-like puncta inside recipient cells. Interestingly, GA transmission occurred as early as 3 days after expression induction. By comparing the spread of 36, 100 and 200 polyGA repeats, we found that polyGA spread is enhanced upon expression of longer GA DPRs. Transmission of polyGA is greater in older flies, indicating that age-associated factors exacerbate the spread. These data highlight a unique propensity of polyGA to spread throughout the brain, which could contribute to the greater abundance of polyGA in patient tissue. In addition, we present a model of early GA transmission that is suitable for genetic screens to identify mechanisms of spread and its consequences in vivo.

Keywords: *C9orf72*, Dipeptide repeat proteins, PolyGA, *Drosophila*, Spread, Repeat size, Ageing

Introduction

Frontotemporal Dementia (FTD) and Amyotrophic Lateral Sclerosis (ALS) are devastating and currently intractable neurodegenerative diseases, characterized histologically by the progressive loss of neurons in the frontal and temporal lobes, or upper and lower motor neurons, respectively [1]. Patients with either disease show a time-dependent progression of symptoms, yet the causes of this deterioration remain unknown.

An expansion of the hexanucleotide sequence GGGGCC in the *C9orf72* (C9) gene, ranging from 30 to several thousand repeats, is the most common familial cause for both FTD and ALS [2–4]. The hexanucleotide expansion is transcribed in both sense and antisense directions, and gives rise to hexanucleotide repeat RNA that accumulates in intranuclear and extranuclear RNA foci [2, 4–6]. In addition, the repeat RNAs can be translated in both directions in all reading frames, by repeat-associated non-AUG (RAN) translation, into 5 different dipeptide repeat (DPR) proteins: polyGA, polyGP, polyGR, polyPA and polyPR [7–9]. Numerous studies have addressed the differential toxicity of C9 RNA foci and DPRs, and have largely concluded that DPRs exert greater toxicity, especially the arginine-rich DPRs and, to a lesser extent, polyGA [10, 11]. However, the relative toxicity of the five DPRs has been mostly deduced from experimental production of proteins with much lower numbers of repeats than those seen in human patients, due to the difficulties in cloning repeat constructs. Importantly, the repeat length of proteins involved in other

* Correspondence: partridge@age.mpg.de
[1]Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Strasse 9b, 50931 Cologne, Germany
[3]Department of Genetics, Evolution and Environment, Institute of Healthy Ageing, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK
Full list of author information is available at the end of the article

neurodegenerative diseases, such as huntingtin and ataxin-3, greatly influences their toxicity [12].

Although the arginine containing DPRs have so far proved the most toxic in animal and cellular models, in patients the contribution of different DPRs to overall toxicity is likely to be affected by their abundance. GA aggregates are the most abundantly detected DPR in patient tissues [8, 13], and it is therefore important to understand the behaviour of this protein at different lengths.

An emerging theme in the field of neurodegenerative diseases is that specific toxic proteins can spread transcellularly, thus contributing to the clinical progression shown by patients [14]. For instance, in experimental models, TDP-43, which typically aggregates in ALS and FTD [1], can spread trans-neuronally in cells [15] and mice [16]. Similarly, three independent studies have reported transmission of the C9 DPRs in cell culture models [17–19]. However, whether this phenomenon occurs in vivo remains unexplored.

We have investigated whether the C9 DPRs spread in vivo. We used the powerful genetics of *Drosophila* and found that, out of the three toxicity-associated DPRs, only GA DPRs spread in vivo in the fly brain, which accumulate in recipient cells as intracellular aggregate-like puncta. Furthermore, spreading was dependent on the repeat length of the GA DPRs, and their transmission was greater in the brains of older flies.

Materials and methods

Drosophila stocks and maintenance

Fly stocks were kept at 65% humidity on a 12:12 h light: dark cycle and fed a standard sugar/yeast/agar (SYA) diet [20]. For experiments using the pan-neuronal elav-GS driver, experimental flies developed and were allowed to mate for 2 days at 25 °C, after which female flies were sorted to SYA food with 200 μM RU486 (Mifepristone) at a density of 20 flies/vial and maintained at 25 °C for 3 days. Flies used for propagation experiments expressed the temperature-sensitive Gal4 inhibitor $Gal80^{ts}$ to minimize the expression of the UAS transgenes during development. This inhibitor is active at 18 °C and can be inhibited to derepress Gal4 activity by shifting flies to 29 °C [21]. Therefore, flies used for propagation experiments developed and were allowed to mate for 2 days at 18 °C, after which female flies were sorted into SYA food at a fly density of 20 flies/vial and maintained at 18 °C or 29 °C as indicated for each experiment.

The following transgenic fly lines were obtained from the Bloomington *Drosophila* Stock Center: tubulin-$Gal80^{ts}$ (BDSC_7019), orco-Gal4 (BDSC_23292), R9D03-Gal4 (BDSC_40726; hereafter referred to as OL-Gal4), UAS-eGFP.NLS (BDSC_4776) and UAS-syt.eGFP (BDSC_6926). The elav-GS driver line was obtained as a generous gift from Dr. Hervé Tricoire (CNRS, France) [22]. The rest of the fly lines used were generated for this study.

Generation of transgenic fly lines and genetics

To generate the mCherry-tagged DPR constructs, we first PCR amplified mCherry using the Phusion polymerase (NEB) and the primers JOL13 and JOL14, which allowed for the addition of an N-terminal NotI restriction site (RS) followed by the linker GGTAGTGGAAGTGGTAGT, as well as a C-terminal KpnI RS after the stop codon. This amplicon was then ligated into the pUAST attB *Drosophila* transgenesis vector, thus forming the hereafter referred to as pUAST-mCherry-C plasmid. In parallel, we PCR amplified the sequences for GA36, GR36, PR36, GA100, GR100 and PR100 [10] using the TaKaRa LA Taq polymerase (Takara Bio Inc.) and the following primers: GA36fwd: JOL26; GA36rev: JOL33; GR36fwd: JOL26; GR36rev: JOL34; PR36fwd: JOL26; PR36rev: JOL35; GA100fwd: JOL26; GA100rev: JOL28; GR100fwd: JOL26; GR100rev: JOL30; PR100fwd: JOL26 and PR100rev: JOL30. This allowed for the addition of an N-terminal EcoRI RS followed by the ATG initiation site, as well as a C-terminal NotI RS. These amplicons were first ligated into the pBlueScript SK(+) plasmid for amplification and subsequently subcloned into the pUAST-mCherry-C plasmid. As a control, we also PCR amplified mCherry using the primers JOL9 and JOL14, which allowed for the addition of an N-terminal EcoRI RS followed by the ATG initiation site, as well as a C-terminal NotI RS after the stop codon. This amplicon was then directly ligated into the pUAST attB plasmid.

To clone the GA200 and GA200-mCherry constructs, we PCR amplified the GA100 sequence [10] in two independent reactions using the TaKaRa LA Taq polymerase and then ligated them together. First, we used primers JOL26 and JOL69 to add an N-terminal EcoRI RS followed by the ATG initiation site, as well as a SmaI RS and an XbaI RS at the C terminus. This amplicon was ligated into the pBlueScript SK(+) plasmid to obtain an EcoRI-ATG-GA100-SmaI-XbaI pBlueScript SK(+) plasmid. Second, we used JOL43 and JOL44 to add an N-terminal XbaI RS and a C-terminal stop codon followed by a NotI RS to GA100. Alternatively, we used JOL43 and JOL28 to add an N-terminal XbaI RS and a C-terminal NotI RS without a stop codon to GA100. The former was ligated into the EcoRI-ATG-GA100-SmaI-XbaI pBlueScript SK(+) plasmid to generate an EcoRI-ATG-GA100-SmaI-XbaI-GA100-Stop-NotI pBlueScript SK(+) plasmid, which was then subcloned into the pUAST attB plasmid (hereafter referred to as GA200). The latter was ligated into the EcoRI-ATG-GA100-SmaI-XbaI pBlueScript SK(+) plasmid to generate an EcoRI-ATG-GA100-SmaI-

XbaI-GA100-NotI pBlueScript SK(+) plasmid, which was subcloned into the pUAST-mCherry-C plasmid, thus forming the hereafter referred to GA200-mCherry plasmid. To achieve high expression levels, all constructs contained the CACC Kozak sequence before their ATG initiation site. Finally, the sequence of all plasmids was verified by Sanger sequencing (Eurofins Genomics).

The sequences of all primers used for this study are included in Table 1.

Constructs were inserted into the fly genome using the phiC31 and attP/attB integration system [23]. For comparisons across the different DPRs, the landing site attP40 was used (i.e., Fig. 1 & Additional file 1: Figure S1, Additional file 2: Figure S2 and Additional file 3: Figures **S3**), whereas for comparisons across the different repeat lengths of GA the landing site attP2 was used (i.e., Figs. 2, 3, 4).

For all experiments, female Gal4 driver flies were crossed with UAS or wild-type (WT) male flies. To generate the final genotypes of the driver flies used for the propagation experiments, the orco-Gal4 and R9D03-Gal4 genes were recombined with UAS-syt.eGFP and UAS-eGFP.NLS, respectively. These flies were then crossed with tub-Gal80^{ts} flies and stable stocks were generated carrying the following genotypes: w^{-}; w, tub-Gal80^{ts}; w, orco-Gal4, w, UAS-syt.eGFP and w^{-}; w, tub-Gal80^{ts}; w, R9D03-Gal4, w, UAS-eGFP.NLS.

Staining and imaging of adult *Drosophila* brains

Brains of adult female flies were dissected in PBS and immediately fixed in 4% paraformaldehyde at 4 °C for 2 h. Tissues were then washed 4–6 × 30 min in PBT (PBS with 0.5% Triton X-100) at room temperature (RT). For experiments where the mCherry and eGFP signals were imaged, brains were subsequently incubated in 50% glycerol in PBS for 1 h at RT after washing and mounted in VectaShield Antifade Mounting Medium with DAPI (Vectorlabs). For experiments where GA or GR were immunostained, brains were blocked in PBT with 5% fetal bovine serum and 0.01% sodium azide for 1 h at RT after initial washing and incubated with a mouse monoclonal anti-GA antibody (1:3000, Merck Millipore) or the 5H9 antibody against polyGR (1:50, [24]) overnight at 4 °C. Following 4–6 × 30 min washes in PBT at RT, brains were incubated with a suitable Alexa Fluor Secondary Antibody (Molecular Probes) overnight at 4 °C. Finally, brains were washed 4–6 × 30 min in PBT, incubated in glycerol-PBS and mounted.

To label cell membranes, brains were incubated in a rhodamine-conjugated phalloidin solution (Life Technologies) diluted in PBT at 0.2 U/ml for 15 min at RT. Brains were subsequently washed 3 × 30 min in PBT, incubated in glycerol-PBS and mounted.

Series of 2-μm z-stacks across the whole fly brain were taken for each image using a Leica SP8-DLS confocal microscope and the same settings were used across genotypes and ages, unless otherwise stated. In experiments where DPR propagation was investigated, brains were imaged with settings where propagated puncta were overexposed, both in the case of the GA100 and the GA200 constructs, and where the signal in the negative control, devoid of any DPR construct, was minimal. This was done in an attempt to maximize the detectability of signal. To further maximize the detectability of specific signal, HyD detectors, gating and the excitation wavelength that

Table 1 List of primers

Primer name	Primer sequence	Purpose
JOL13	ATATGCGGCCGCCGGTAGTGGAAGTGGTAG TGTGAGCAAGGGCGAGGAG	Generation of pUAS T-mCherry-C
JOL14	CCCCGGTACCTCACTTGTACAGCTCGTCCATG	Generation of pUAS T-mCherry-C and mCherry-only pUAST plasmids
JOL26	ATATGAATTCGGATCCCACCATG	Generation of GA36-mCherry, GR36-mCherry, PR36-mCherry, GA100-mCherry, GR100-mCherry and PR100-mCherry, GA200 and GA200-mCherry plasmids
JOL33	AAGCGGCCGCTGAAGCG	Generation of GA36-mCherry plasmid
JOL34	AAGCGGCCGCTGATCTGC	Generation of GR36-mCherry plasmid
JOL35	AAGCGGCCGCTGATCTGG	Generation of PR36-mCherry plasmid
JOL28	AAAAGCGGCCGCTGATGCTC	Generation of GA100-mCherry plasmid
JOL30	AAAAGCGGCCGCTGAACGTC	Generation of GR100-mCherry plasmid
JOL34	AAAAGCGGCCGCTGATCGAG	Generation of PR100-mCherry plasmid
JOL9	AAAAGAATTCCAACATGGTGAGCAAGGGCGAG	Generation of the mCherry-only pUAST plasmids
JOL69	CCGCGGCCGCTCTAGACCCGGGTGATGCTC CTGCTCC	Generation of the GA200 and GA200-mCherry plasmids
JOL43	GAATTCGGATCCCACCATGTCTAGAGGAGCT	Generation of the GA200 and GA200-mCherry plasmids
JOL44	CTTGCGGCCGCTTATGCTCC	Generation of the GA200 and GA200-mCherry plasmids
JOL28	AAAAGCGGCCGCTGATGCTC	Generation of the GA200 and GA200-mCherry plasmids

maximized the fluorescence emission of all fluorophores were used in all cases during imaging.

Quantification of confocal images and statistics

All confocal images acquired for experiments where DPR spread was tested were first processed using ImageJ before subjecting them to quantification analysis. First, maximum z-stack projections were obtained to identify the lamina surrounding the optic lobes, as well as distinct artifacts, which were cropped from the stacks. In addition, areas of initial expression induction were also removed. For the latter, brain regions positive for eGFP were identified and cropped in experiments where eGFP was co-expressed along with the relevant DPR (i.e., Figs. 1, 3 & Additional file 3: Figure S3). This included the ORN axons and synaptic terminals for Fig. 1 and Additional file 3: Figure S3, or the medulla of the optic lobes, as well as a distinct region in the antennal lobes, when the OL driver was used. Alternatively, a rectangle spanning the visually detectable antennal lobes and the rest of the lower part of the central brain was drawn in experiments where eGFP was not co-expressed (i.e., Figs. 2 & 4), and its content was also cropped to ensure that no puncta within the axons or the terminals of ORNs were included in the quantification of propagated puncta. Puncta in the remaining brain areas were quantified from the cropped z-stacks in 3D using the image analysis software Imaris 9.2.0 (Oxford Instruments). After background correction, the built-in spot detection algorithm was used to identify spots with a minimum size of 500 nm. Detection settings were adjusted based on the maximum intensity of the spots, which proved the most accurate filter to distinguish between strongly labelled spots (considered as real GA puncta) and weak/low quality spots from trachea or background. The same parameters were used for all of the conditions compared in the same experiment.

For the quantification of eGFP or GA levels in ORNS, we used their fluorescent signal in the ORN terminals in the antennal lobes as a proxy for their overall levels. Briefly, whole-brain stacks were taken with non-saturating settings for the ORN eGFP or GA signals, maximum intensity projections were generated from each z-stack, and the mean intensity of eGFP or GA in the synaptic terminals of ORNs was measured using ImageJ. The same settings were used for all of the conditions compared in the same experiment.

Statistical analysis was performed using GraphPad Prism. Individual statistical tests are indicated in the figure legends. Both One-way and Two-way ANOVA were always followed by Bonferroni post hoc test. P values < 0.05 were considered significant: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$.

Results

GA DPRs, but not GR or PR DPRs, spread rapidly in the fly brain

To address whether toxicity-associated DPRs can spread in vivo, we generated novel fly lines that expressed mCherry-tagged GA, GR or PR with 36 or 100 repeats (hereafter GA36, GR36, PR36, GA100, GR100 and PR100) in a UAS-transgene. All transgenes were integrated into the same genomic locus, the attP40 landing site, which we previously confirmed to produce equal transcript levels of untagged DPRs [10]. mCherry-tagged DPRs were used in an effort to avoid differences in sensitivity of the different DPR-specific antibodies. To confirm expression of the DPR constructs, we generated flies with pan-neuronal induction of each of the mCherry-tagged DPRs, using the inducible elav-GS system, and imaged the mCherry signal in adult fly brains after induction for 3 days. Expression of all the DPR36-mCherry fusion proteins could be detected by imaging their mCherry signal (Additional file 1: Figure S1). The signal from the PR36-mCherry flies was stronger than that of GA36-mCherry and GR36-mCherry, probably due to the previously reported nuclear location of PR [25, 26]. However, while we detected the mCherry signal from mCherry-tagged GA100 and PR100 pan-neuronally (Additional file 2: Figure S2A-C), the presence of GR100-mCherry could only be verified when a GR-specific antibody was used (Additional file 2: Figure S2D-E). Interestingly, the expression of GR100-mCherry was almost exclusively detected in the median neurosecretory cells (MNCs) in the pars intercerebralis, where the expression levels of the majority of the DPRs tested, but not mCherry only, was also particularly high (Additional file 1: Figure S1 and Additional file 2: Figure S2). This suggests that MNCs may be particularly vulnerable to the accumulation of the C9 DPRs.

We next addressed whether the toxic DPRs have the ability to spread in vivo. Given that a previous study reported propagation of mutant huntingtin from Olfactory Receptor Neurons (ORNs) to other brain regions in *D. melanogaster* [27], we also initiated expression in this brain area. We imaged the brains of flies where ORN-specific expression of GA36-mCherry, GR36-mCherry or PR36-mCherry had been induced for 3 days in the adult fly using a temperature-inducible Gal80 and the ORN-specific orco-Gal4 driver [21, 28]. Since the cell bodies of ORNs are outside the central brain, and therefore only the axonal projections and synaptic terminals of ORNs can be detected in the adult central brain of *Drosophila* after dissection, we co-expressed eGFP-tagged synaptotagmin to label ORNs and control for driver specificity [29]. No specific mCherry signal was found outside of ORNs (Additional file 3: Figure S3), suggesting that the short isoforms of the toxic DPRs cannot spread, at least after short-term expression from this location. Interestingly, when flies

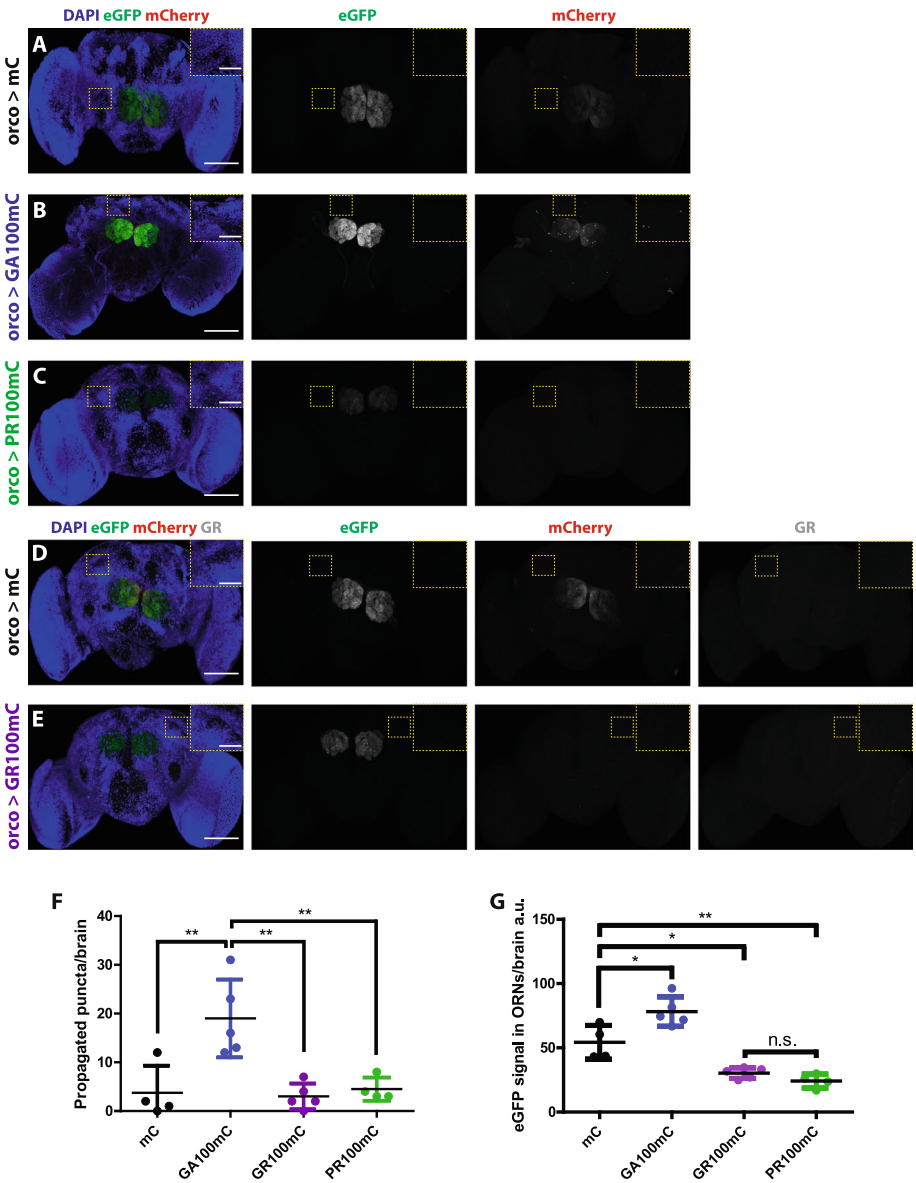

Fig. 1 GA100-mCherry, but not GR100-mCherry or PR100-mCherry, can spread outside of ORNs. **a**-**e** Representative images of 5-days-old fly brains expressing mCherry (**a** & **d**), GA100-mCherry (**b**), PR100-mCherry (**c**) or GR100-mCherry (**e**) in Olfactory Receptor Neurons (ORNs) for 3 days. The same settings were applied to all genotypes while imaging their eGFP and mCherry signals. Spreading was only observed in flies expressing GA100-mCherry. EGFP and mCherry were detected using fluorescence as read-out. A GR-specific antibody was used to detect GR100-mCherry (**e**) and mCherry (**d**). Fly brains expressing only mCherry (**a** & **d**) were used as control to show that mCherry cannot spread by itself and to verify antibody specificity (**d**). Insets of the highlighted brain regions are shown. **f** Quantification of the number of mCherry puncta detected outside of ORNs across genotypes per brain after induction for 3 days. **g** Quantification of the eGFP signal detected within ORNs per brain (**$P < 0.01$ and *$P < 0.05$; One-way ANOVA, $n = 4$–6). Scale bars in images and insets are 100 um and 10 um, respectively

were induced to express GA100-mCherry, GR100-mCherry or PR100-mCherry in ORNs for 3 days, mCherry-positive puncta were exclusively detected outside of the ORNs of GA100-mCherry-expressing flies (Fig. 1a-f), suggesting that longer GA DPRs are particularly prone to spread. Moreover, we detected a strong reduction in the eGFP fluorescence within the ORNs of flies expressing GR100-mCherry and PR100-mCherry (Fig. 1a-e, g), which may result from the well-known inhibitory effect of the arginine-rich DPRs on translation [30]. Also of note, unlike for GA100-mCherry, no specific mCherry signal was detected in the axons or synaptic terminals of ORNs in GR36-mCherry-, PR36-mCherry-, GR100-mCherry- or PR100-mCherry-expressing flies

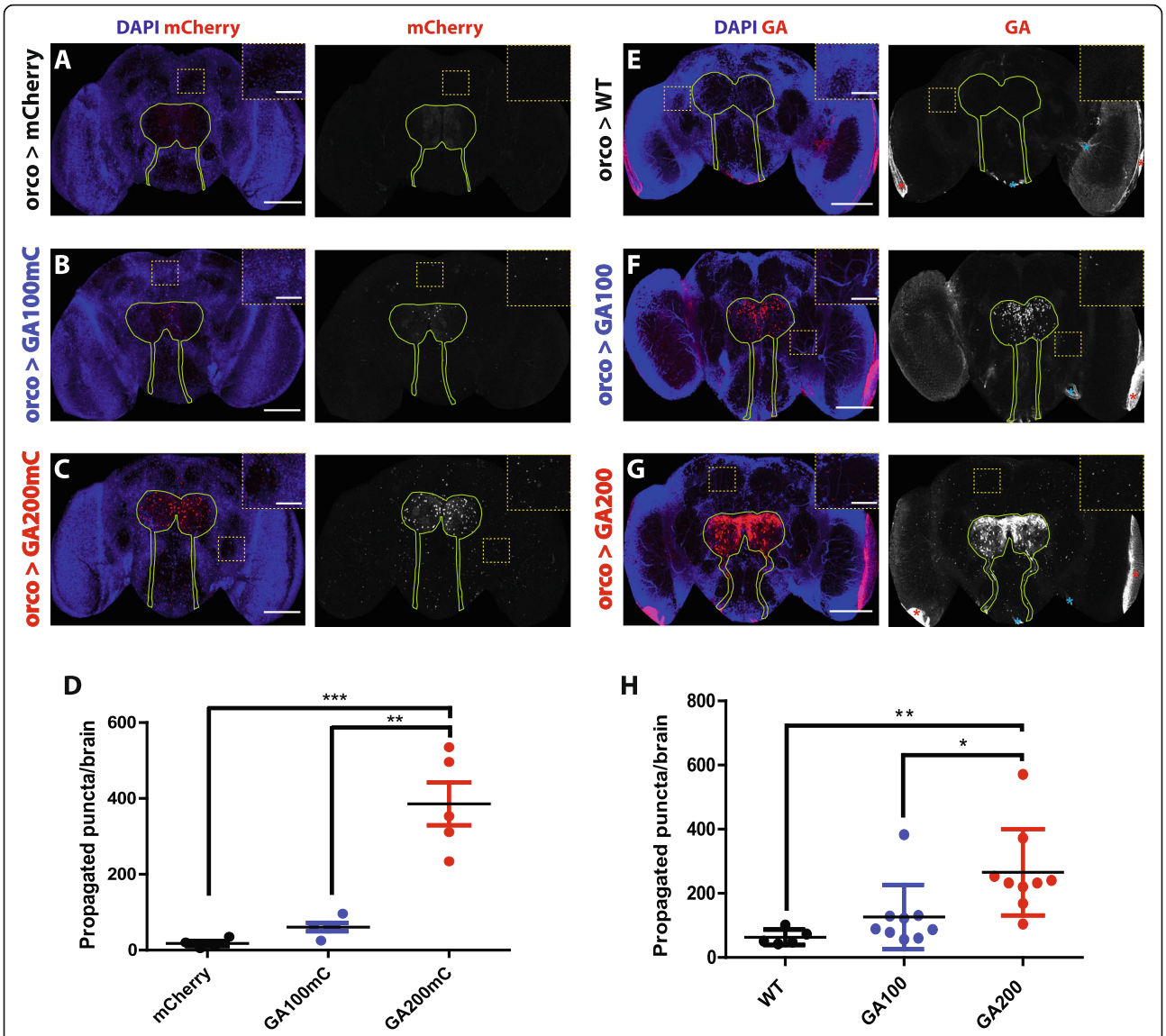

Fig. 2 Spreading is increased in longer GA repeat proteins. **a**-**c** Representative images of 5-days-old fly brains expressing mCherry (**a**), GA100-mCherry (**b**) and GA200-mCherry (**c**) from Olfactory Receptor Neurons (ORNs) for 3 days. The same settings were applied to all genotypes while imaging their mCherry signals. Spreading is greater for GA200-mCherry than for GA100-mCherry. The mCherry signal was detected using its fluorescence as read-out. Flies expressing only mCherry (**a**) served as negative control to ensure that mCherry does not spread by itself. **d** Quantification of mCherry puncta detected in the central brain outside of ORNs across genotypes per brain after induction for 3 days (***$P < 0.001$ and **$P < 0.01$; One-way ANOVA, $n = 4$–5). **e**-**g** Representative images of 5-days-old fly brains expressing GA100 (**f**) or GA200 (**g**) in ORNs for 3 days and probed with an anti-GA antibody. GA200 spreads more than GA100. The same settings were used while imaging the GA signal across genotypes. Flies expressing only the driver (**e**) served to control for unspecific binding of the anti-GA antibody. Unspecific binding to trachea (blue asterisks) and the lamina of the optic lobes (red asterisks) was observed. **h** Quantification of GA puncta detected outside of the ORN boundaries across genotypes per brain after induction for 3 days. (**$P < 0.01$ and *$P < 0.05$; One-way ANOVA, $n = 5$–9). The boundaries of the ORN axons and synaptic terminals are highlighted with a solid green line. Insets of the indicated areas are shown to facilitate visualization. Scale bars in images and insets are 100 um and 10 um, respectively

(Additional file 3: Figure S3C, D), indicating that these DPRs are not transported along axons, a potential requirement for DPR spread. Altogether, our data indicate that, out of the three toxic DPRs, at least for DPRs up to 100 repeats in length, only GA can spread from ORNs to the rest of the central brain.

GA repeat length modulates the spread of GA DPRs

Unlike GA100-mCherry, GA36-mCherry did not spread, suggesting that spread is repeat length-dependent. To test this hypothesis further, we generated novel fly lines expressing mCherry-tagged GA100 and GA200 from the same genomic locus (attP2 landing site) to ensure equal

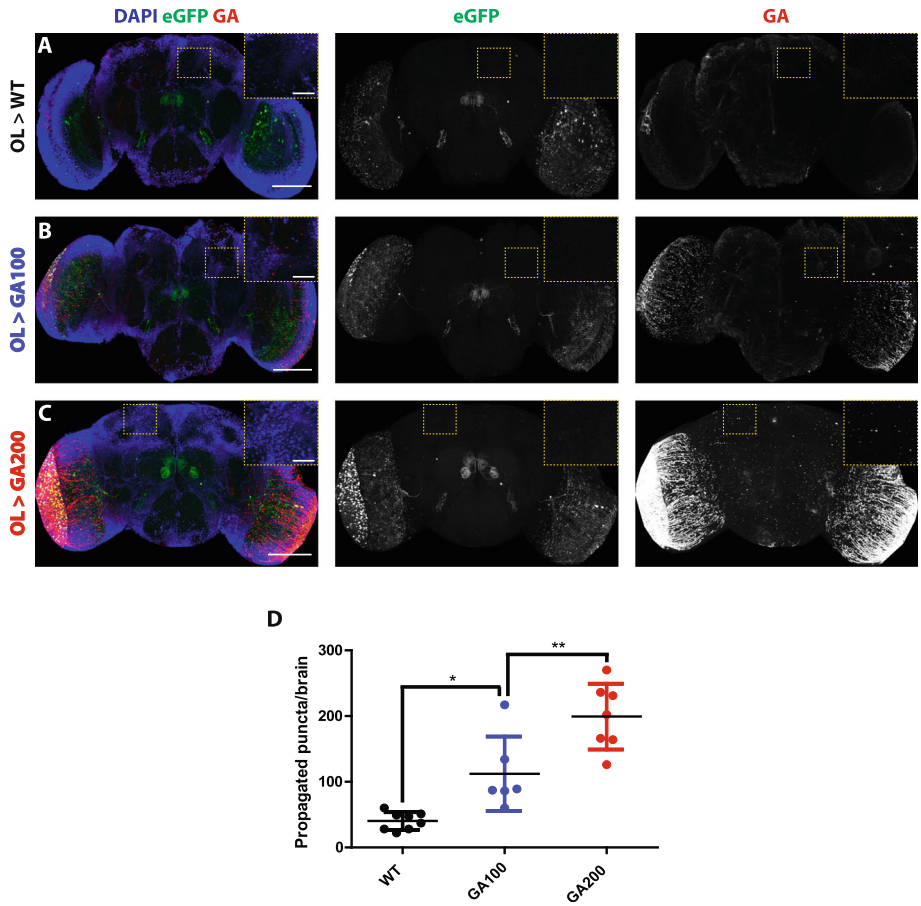

Fig. 3 GA spreads in a repeat length-dependent manner from an independent neuronal population. **a**-**c** Representative images of 5-days-old fly brains from (**a**) control flies, expressing only the Optic Lobe (OL)-Gal4 driver, (**b**) flies expressing GA100 or (**c**) GA200 in the OLs for 3 days and probed with an anti-GA antibody. GA200 also spreads more than GA100 from this brain region. The same settings were used while imaging the GA signal across genotypes. EGFP with a nuclear localization signal was co-expressed to identify the cells targeted by the OL-Gal4 driver. **d** Quantification of GA puncta detected in the central brain outside of the targeted cells after expression of the indicated constructs for 3 days. Flies expressing only the driver (**a**) were used to control for unspecific binding of the anti-GA antibody (**$P < 0.001$ and *$P < 0.05$; One-way ANOVA, $n = 6–8$). Insets of the indicated areas are shown to facilitate visualization. Scale bars in images and insets are 100 um and 10 um, respectively

transcript levels. In addition, to exclude the possibility that co-expression of eGFP-tagged synaptotagmin could influence transmission, we only expressed polyGA-mCherry. We measured the spread of the two mCherry-tagged constructs from ORNs, and found accumulation of mCherry puncta of both DPRs outside of this neuronal population after 3 days of expression induction, with substantially greater spread of the 200 than the 100 GA DPR (Fig. 2a-d). Spreading of GA was therefore more pronounced with longer repeats and was independent of eGFP-tagged synaptotagmin co-expression.

Since tags can interfere with protein function [31], we next tested the spread of untagged GA DPRs using GA100 and GA200 expressed in ORNs and a GA-specific antibody [32]. In agreement with our results using mCherry-tagged GA constructs, we found that the number of GA puncta detected outside of ORNs greatly increased with repeat length (Fig. 2e-h), further supporting the notion that the propensity of GA to spread is greater with longer GA repeats. To determine if GA could spread from different types of neurons, we tested whether untagged GA was transmitted from the optic lobes (OL). We expressed the GA constructs along with nuclear eGFP for 3 days in the OLs using the R9D03-Gal4 driver [33] and, consistent with our finding in ORNs, GA also spread in a repeat-length dependent manner from the OLs (Fig. 3a-d).

To determine if the propagated GA enters recipient cells, we co-stained brains from flies expressing GA200 in ORNs with fluorescently labelled phalloidin, a dye that strongly binds to actin F and can therefore be used to identify the boundaries of single cells in tissue [34]. Using this approach, we detected GA positive puncta in the cytoplasm of recipient cells, thus indicating that propagated GA puncta are intracellular (Additional file 4: Figure S4).

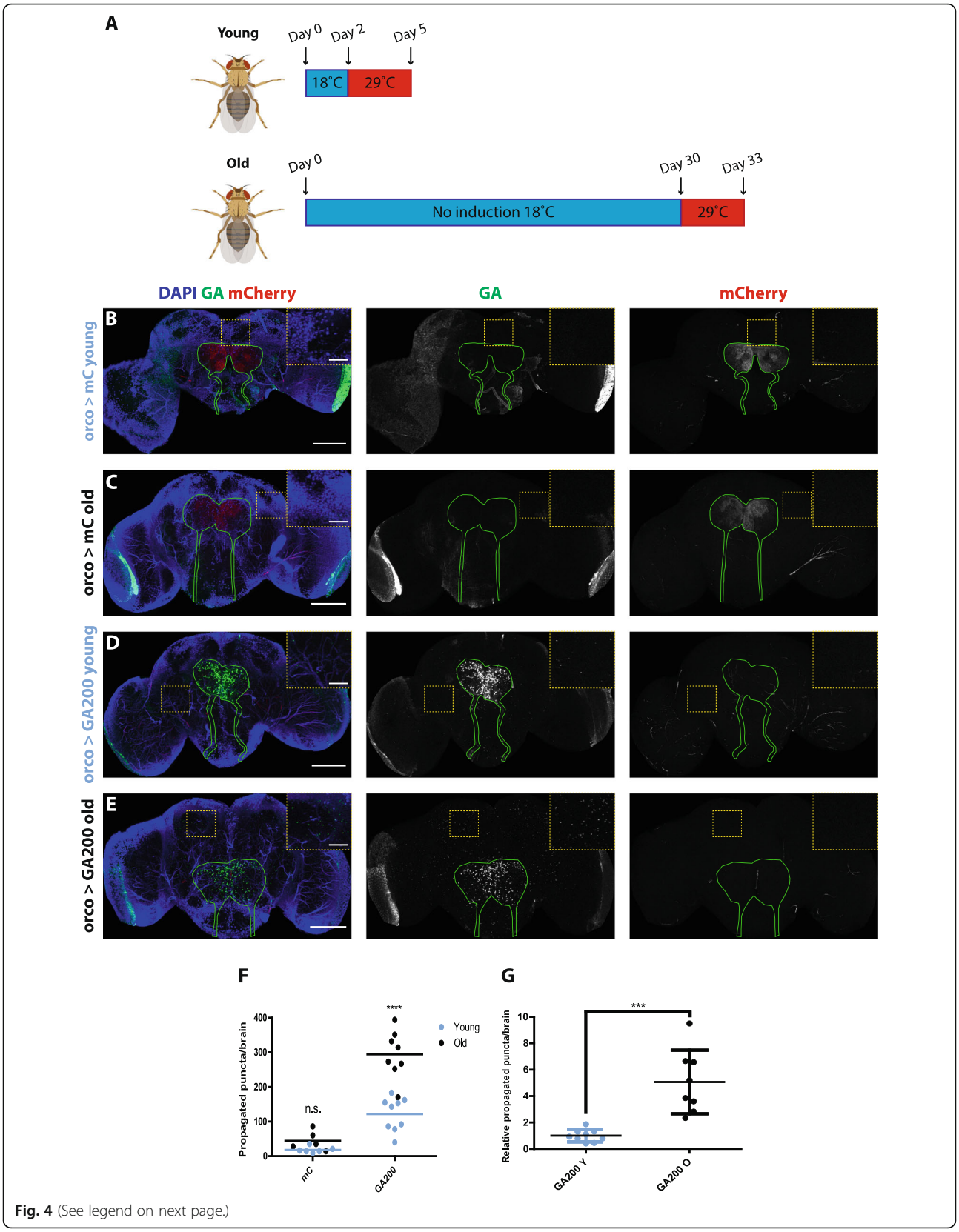

Fig. 4 (See legend on next page.)

(See figure on previous page.)
Fig. 4 Age-associated factors exacerbate GA spread. **a**The expression of GA200 was induced for 3 days in Olfactory Receptor Neurons (ORNs) in young (2 days old) and old flies (30 days old), after which GA spread was measured. Fly cartoons were created with BioRender. **b** & **c** Representative images of control fly brains expressing mCherry in ORNs for 3 days in young (**a**) or old (**b**) flies. **d** & **e** Representative images of fly brains expressing GA200 in ORNs for 3 days in young (**d**) and old (**e**) animals. Brains were probed with an anti-GA antibody. The outline of ORN axons and synaptic terminals is shown in green. Insets of indicated areas highlight differences in the number of propagated dots across conditions. **f** Quantification of the total number of GA-positive dots detected outside of ORNs after 3 days of expression in young and old flies (age: ****$P < 0.0001$; genotype: ****$P < 0.0001$; interaction: ***$P < 0.001$; Two-way ANOVA, $n = 5–9$). **g** Quantification of the number of propagated GA-positive dots relative to the GA signal in ORNs after 3 days of expression in young and old flies (***$P < 0.001$, t-test). Scale bars in images and insets are 100 um and 10 um, respectively

Altogether, our results show that, in two independent neuronal subsets, longer GA repeats spread in a length-dependent manner.

GA DPRs exhibit an age-related increase in spreading

Given that ageing is a major risk factor for ALS and FTD [35], we investigated whether GA spread was affected by the age at which we induced GA expression. We induced ORN-specific expression of untagged GA200 starting in 2-day-old or 30-day-old adult flies for 3 days, and measured the spread outside of ORNs (Fig. 4a). There was a 3-fold increase in the total number of propagated GA puncta when GA expression was induced at the older age (Fig. 4b-f). Given that the accumulation of the peptides could change after expression induction at different ages, we also quantified the cumulative number of propagated GA puncta relative to GA expression in ORNs, as an indicator of whether the proportion of propagated GA compared to the total amount of GA in ORNs would change at different ages. Indeed, we found a larger proportion of propagated GA compared to ORN GA after expression induction in older brains (Fig. 4g), thus showing that the increased spread in older brains is not simply due to changes in the accumulation of GA in ORNs upon expression in older brains. Collectively, these results suggest that age-associated factors strongly affect the propagation propensity of GA DPRs.

Discussion

In this study, we have explored whether the toxic C9 DPRs, namely GA, GR and PR, can spread, and the contributions of DPR repeat length and age to propensity to spread in vivo. Using the brain of adult *Drosophila* as a model, we show that GA, but not GR and PR, DPRs spread out of ORNs, at least as soon as 3 days after expression induction, and could therefore be a relevant early event in the pathogenesis initiation of C9 ALS and C9 FTD. In addition, while 36-repeat GA did not show evidence of spread, transmission occurred with expression of 100 and, to a greater extent, with 200 GA repeats. Finally, GA spread was more marked when its expression was induced at an older age.

Recent studies have shown that some proteins that typically aggregate in the brains of patients with various neurodegenerative diseases can be transmitted in model organisms. These findings have led to the hypothesis that protein transmission could underpin the clinical progression of such patients [14]. For instance, the clinical progression of ALS and FTD may be explained by the progressive spreading of TDP-43 pathology across conserved neuronal circuits relevant to these diseases [14, 16, 36, 37]. To investigate whether the toxicity-associated DPRs derived from a mutation in the *C9orf72* gene could also spread under in vivo conditions, we first used flies expressing, exclusively in ORNs, mCherry-tagged constructs of GA, GR and PR, to avoid an antibody bias, as antibodies specific for different DPRs could have different sensitivities. Interestingly, unlike mutant huntingtin, which strongly spread to a pair of large posterior neurons in the posterior protocerebrum upon expression in ORNs [27], the spreading pattern of GA DPRs was not specific to a single neuronal subset, either upon expression from ORNs or OLs, which might indicate that GA transmission does not only occur through synapses in vivo. By performing co-stainings with the cell membrane dye phalloidin, we found that the propagated GA signal accumulates in the shape of aggregate-like puncta inside recipient cells, which adds GA onto the growing list of proteins shown to spread from cell to cell in in vivo models of neurodegenerative diseases. In fact, GA accumulates in intracellular aggregates and is the most widely detected DPR in patient tissue [8, 13], which may be at least partially attributable to this ability to spread. Our result is also in line with a previous cell culture study, where GA, but not GR or PR, was found to spread from cell to cell [19]. Future studies should address what mechanisms are activated by GA, and not by GR or PR, that could be associated with its release from DPR-expressing neurons and/or its uptake by recipient neurons. While expression of the arginine-rich DPRs has been mostly associated with translation inhibition [38, 39], impairment of nucleocytoplasmatic transport [40–42], RNA processing [25, 43–45] and dysfunction of stress granule dynamics [39, 46], GA expression has been strongly correlated with impairment

of proteostasis [47, 48]. Indeed, we observed that the co-expression of eGFP-tagged synaptotagmin with mCherry-tagged GR and PR led to decreased eGFP signal, which is likely due to their well-known inhibitory effect on translation [30]. In contrast, when we co-expressed mCherry-tagged GA, we found increased eGFP signal, suggesting that GA may be impairing eGFP degradation by damaging the proteostasis machinery. Therefore, we hypothesize that the preferential spread of GA could be associated with GA-induced proteostasis impairment, which has been shown to exacerbate the release of toxic proteins [49]. Another non-exclusive reason why GA, but not GR or PR, may spread could be related to the differential ability of neurons to transport each DPR along axons. In fact, unlike GA, we did not detect GR or PR in the axons or synaptic terminals of ORNs. Furthermore, GA, unlike GR and PR, has been shown to form oligomeric amyloids [17], which have been associated with greater propensity to spread [50]. This feature could also account for the greater propensity of GA to spread compared to GR and PR.

Unlike the distribution of phosphorylated TDP-43, which has been proposed to occur in patients in a progressive pattern indicative of pathology spread [36, 37], DPR aggregates have not yet been shown to occur in a staged manner. Thus, an argument against the relevance of DPR spreading is that they are produced by all cells expressing C9orf72, so do not necessarily need to receive an aggregate from a neighboring/connected cell to allow aggregate formation. However, it is possible that the addition of exogenous GA seeds could initiate aggregation of non-aggregated GA molecules, thus speeding up the process in receiver cells. How this might lead to TDP-43 changes is an open question but several possibilities exist. First, several studies have shown that GA itself can trigger cytoplasmic accumulation and hyperphoshorylation of TDP-43 [32, 51, 52]. Therefore, given that GA transmission would exacerbate the accumulation of GA, it could play a relevant role in eliciting the cytoplasmic accumulation and aggregation of TDP-43, which would then spread across the neuronal circuits relevant to ALS/FTD symptoms. This idea is supported by reports that the accumulation of GA aggregates precedes TDP-43 pathology [8, 53–55]. Second, one in vitro study showed that GA can spread within exosomes [18], which are known to comprise different kinds of biological material, including TDP-43 [56]. Therefore, GA accumulation could favour the formation of secretory vesicles, like exosomes, which would comprise, apart from GA aggregates, TDP-43 seeds or other material that would contribute to the degeneration of the ALS/FTD-relevant neuronal circuits. Indeed, the relevance of exosomes in the transmission of aggregation-prone proteins has already been shown in vivo [57]. Finally, while GA aggregates do not strongly correlate with neurodegeneration, or TDP-43 pathology [58], it cannot be excluded that specific forms of GA, e.g., oligomers or aggregates of a specific size, potentially not detected by inclusion staining, may correlate better with ALS/FTD neurodegeneration and TDP-43 pathology than the overall burden of GA aggregates. However, this point remains completely unexplored.

Both the morphology and spreading propensity of GA DPRs is dependent on repeat lengths. In the adult fly brain, GA36-mCherry shows a diffuse staining and does not spread. In contrast, GA100-mCherry and GA200-mCherry look aggregated and spread in a repeat length-dependent manner. These differences highlight the relevance of testing DPRs of different lengths to fully understand the behaviour of these peptides, and support future in-depth studies to understand the effect that repeat size has in the *C9orf72* mutation context in humans, which remains controversial [59–61] . In the case of the untagged GA DPRs, differences in GA transmission could be confounded by the GA antibody having more epitopes to bind to in the longer constructs, which would lead to stronger propagation signal being detected for GA200 than for GA100. However, this is unlikely because we also see greater spread for mCherry-tagged GA200 compared to mCherry-tagged GA100 when directly recording the mCherry fluorescence signal, which is only influenced by the overall abundance of each DPR as both GA100-mCherry and GA200-mCherry carry one mCherry tag per polyGA molecule. Given that the amino acid composition of GA100 and GA200 DPRs is the same, we hypothesize that the longer GA DPRs spread more either because they are more difficult to degrade, presumably due to their greater propensity to aggregate or their stronger effect on proteostasis, or because they earlier acquire a conformation that triggers their transmission. For instance, tau and α-synuclein, which typically aggregate and spread in Alzheimer's disease and Parkinson's disease models respectively, need to be at least partially aggregated to be able to propagate [14].

Ageing is a major risk factor for ALS and FTD [35]. We found that GA spread is greater when its expression is induced at a later age, suggesting that ageing-associated factors, such as impaired proteostasis, promote DPR spread, and therefore that GA transmission is likely to occur at a greater rate at old ages. These findings are in agreement with a recent study that reported a 2-fold increase in the spreading rate of tau upon injection of human tau-expressing viruses in aged murine brains compared to young ones [62]. For instance, GA aggregates may be degraded when they are formed at a young age, when the proteostasis machinery works efficiently, but accumulate during the progressive decline of

proteostasis activity during ageing. The excessive accumulation of GA aggregates could trigger their mislocalization to exocytotic vesicles, thus mediating their extracellular release.

Taken together, we show that GA DPRs can rapidly spread in vivo, which is strongly influenced by their repeat length and the age of the GA-expressing neurons. The wide range of available tools to genetically manipulate flies makes this *Drosophila* model of early in vivo GA spread an attractive system that could be used in genetic or pharmacological screens to gain further insights into the mechanism(s) underlying GA transmission, its consequences for the recipient tissue and the search for interventions that can abolish this event.

Conclusions

In conclusion, we have provided the first evidence for transmission of GA DPRs in an in vivo setting using the adult fly brain as a complex model. The extent of spread was magnified for longer GA DPRs and upon expression in old flies, suggesting that this mechanism could be of relevance for both the initiation and the progression of C9 ALS and C9 FTD.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s40478-019-0860-x.

Additional file 1: Figure S1. mCherry-tagged DPR36 constructs can be detected by imaging their endogenous mCherry signal. A-E show representative images of 5-days-old adult fly brains from flies induced to pan-neuronally express each of the indicated mCherry-tagged DPR36 construct for 3 days. 10 times lower settings were used to image mCherry (B) and PR36mCherry (E), as the signal was much stronger in those genotypes. For the rest of the genotypes, the settings were the same. No antibodies were used. Insets highlight the brain area where Median Neurosecretory Cells (MNCs) are located. Scale bars in images and insets are 100 um and 10 um, respectively.

Additional file 2: Figure S2. Detection of mCherry-tagged DPR100 proteins upon pan-neuronal expression. A-E Representative images of 5-days-old adult fly brains that pan-neuronally express the indicated mCherry-tagged DPR100 constructs for 3 days. 10 times lower settings were used to image GA100mCherry (B) and mCherry (D) as the signal was much stronger in those genotypes. No antibodies were used for A-C. D-E Fly brains were stained with an anti-GR antibody. GR100mC can be most clearly detected in the brain area where Median Neurosecretory Cells (MNCs) are located. Insets of the indicated areas are shown to facilitate visualization. Scale bars in images and insets are 100 um and 10 um, respectively.

Additional file 3: Figure S3. GA36-mCherry, GR36-mCherry and PR36-mCherry cannot spread from ORNs. A-D Representative images of 5-days-old fly brains expressing GA36-mCherry (B), GR36-mCherry (C) or PR36-mCherry (D) in Olfactory Receptor Neurons (ORNs) for 3 days. Synaptotagmin-eGFP was co-expressed in all genotypes to identify ORNs. Flies expressing mCherry (A) were used as a negative control to ensure that mCherry cannot spread by itself. No antibodies were used. Insets of the indicated areas are also shown to facilitate visualization. Scale bars in images and insets are 100 um and 10 um, respectively.

Additional file 4: Figure S4. GA propagated puncta are intracellular. A Representative image of a 5-days-old fly brain expressing GA200 in Olfactory Receptor Neurons (ORNs) for 3 days, and stained with an anti-GA antibody (green) and the rhodamine-conjugated fluorophore phalloidin (red). Scale bar = 25 um. B Inset of the area highlighted in a yellow dotted square in A outside of the ORN synaptic terminals where GA has propagated. Five cells positive for GA intracellular puncta can be observed. Scale bar = 3 um.

Abbreviations

ALS: Amyotrophic lateral sclerosis; C9orf72: Chromosome 9 open reading frame 72; DPR: Dipeptide repeat; eGFP: Enhanced green fluorescent protein; FTD: Frontotemporal dementia; GA: Glycine-alanine; GP: Glycine-proline; GR: Glycine-arginine; MNCs: Median neurosecretory cells; OL: Optic lobe; ORN: Olfactory receptor neurons; PA: Proline-alanine; PR: Proline-arginine; RAN: Repeat-associated non-AUG; RS: Restriction site; TDP-43: TAR DNA-binding protein of 43 kDa; WT: Wild-type

Acknowledgements

The 5H9 anti GR antibody was a generous gift from Friederich A. Grässer [24]. Imaging analyses were performed in the FACS & Imaging Core Facility at the Max Planck Institute for Biology of Ageing. Stocks obtained from the Bloomington *Drosophila* Stock Center (NIH P40OD018537) were used in this study. We are also grateful to Natalia Kononenko (CECAD), David Vilchez (CECAD) and Teresa Niccoli (UCL) for helpful comments on the manuscript. Javier Morón-Oset received support by the Cologne Graduate School of Ageing Research, which is funded by the Deutsche Forschungsgemeinschaft (DFG), German Research Foundation under Germany's Excellence Strategy EXC 2030/1, Project-ID 390661388.

Authors' contributions

JMO, SG and LP designed the study; JMO performed cloning; JE injected plasmids for the generation of transgenic flies; JMO performed fly crossings; JMO and TS performed immunostainings, imaging and analysis; JMO, AMI, SG and LP wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was funded by the Max Planck Society and the Wellcome Trust (WT098565/Z/12/Z).

Availability of data and materials

All data generated or analysed during this study are included in this published article (and its supplementary information files).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Strasse 9b, 50931 Cologne, Germany. [2]Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK. [3]Department of Genetics, Evolution and Environment, Institute of Healthy Ageing, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK.

Received: 18 October 2019 Accepted: 2 December 2019
Published online: 16 December 2019

References

1. Ling SC, Polymenidou M, Cleveland DW (2013) Converging mechanisms in als and FTD: disrupted RNA and protein homeostasis. Neuron 79:416–438. https://doi.org/10.1016/j.neuron.2013.07.033
2. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Gilmer F, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung G-YR, Karydas A, Seeley WW, Josephs KA,

Geschwind DH, Wszolek ZK, Feldman H, Petersen R, Miller BL, Dickson D, Boylan K (2011) Expanded GGGGCC hexanucleotide repeat in non-coding region of C9ORF72 causes chromosome 9p-linked frontotemporal dementia and amyotrophic lateral sclerosis. Neuron 72:245–256. https://doi.org/10.1016/j.neuron.2011.09.011.Expanded

3. Majounie E, Renton AE, Mok K, Dopper EGP, Waite A, Rollinson S, Chiò A, Restagno G, Nicolaou N, Simon-Sanchez J, van Swieten JC, Abramzon Y, Johnson JO, Sendtner M, Pamphlett R, Orrell RW, Mead S, Sidle KC, Houlden H, Rohrer JD, Morrison KE, Pall H, Talbot K, Ansorge O, Hernandez DG, Arepalli S, Sabatelli M, Mora G, Corbo M, Giannini F, Calvo A, Englund E, Borghero G, Floris GL, Remes AM, Laaksovirta H, McCluskey L, Trojanowski JQ, Van Deerlin VM, Schellenberg GD, Nalls MA, Drory VE, Lu CS, Yeh TH, Ishiura H, Takahashi Y, Tsuji S, Le Ber I, Brice A, Drepper C, Williams N, Kirby J, Shaw P, Hardy J, Tienari PJ, Heutink P, Morris HR, Pickering-Brown S, Traynor BJ, Adamson G, Bayer AJ, Beck J, Callister JB, Blake DJ, Blumen SC, Collinge J, Dunckley T, Ealing J, East S, Elman L, Gerhard A, Guerreiro RJ, Gwinn K, Halliwell N, Hamdalla HH, Hewitt C, Ince P, Jablonka S, James C, Kent L, Knock JC, Lynch T, Mahoney C, Mann D, Neal J, Norris D, O'Dowd S, Richardson A, Rossor M, Rothstein J, Scholz SW, Snowden J, Stephan DA, Toulson G, Turner MR, Warren JD, Young K, Weng YH, Kuo HC, Lai SC, Huang CL, Camuzat A, Entraingues L, Guillot-Noël VP, Clerget-Darpoux F, Corcia P, Couratier P, Didic M, Dubois B, Duyckaerts C, Guedj E, Golfier V, Habert MO, Hannequin D, Lacomblez L, Meininger V, Salachas F, Levy R, Michel BF, Pasquier F, Puel M, Thomas-Anterion C, Sellal F, Vercelletto M, Moglia C, Cammarosano S, Canosa A, Gallo S, Brunetti M, Ossola I, Marinou K, Papetti L, Pisano F, Pinter GL, Conte A, Luigetti M, Zollino M, Lattante S, Marangi G, la Bella V, Spataro R, Colletti T, Battistini S, Ricci C, Caponnetto C, Mancardi G, Mandich P, Salvi F, Bartolomei I, Mandrioli J, Sola P, Lunetta C, Penco S, Monsurrò MR, Tedeschi G, Conforti FL, Gambardella A, Quattrone A, Volanti P, Floris G, Cannas A, Piras V, Marrosu F, Marrosu MG, Murru MR, Pugliatti M, Parish LD, Sotgiu A, Solinas G, Ulgheri L, Ticca A, Simone I, Logroscino G (2012) 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. https://doi.org/10.1016/S1474-4422(12)70043-1

4. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita VM, Kaivorinne AL, Hölttä-Vuori M, Ikonen E, Sulkava R, Benatar M, Wuu J, Chiò A, Restagno G, Borghero G, Sabatelli M, Heckerman D, Rogaeva E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72:257–268. https://doi.org/10.1016/j.neuron.2011.09.010

5. Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PEA, Caulfield T, Daughrity L, Dunmore JH, Castanedes-Casey M, Chew J, Cosio DM, Van Blitterswijk M, Lee WC, Rademakers R, Boylan KB, Dickson DW, Petrucelli L (2013) Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol 126:829–844. https://doi.org/10.1007/s00401-013-1192-8

6. Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P, Isaacs AM (2013) C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. Acta Neuropathol 126:845–857. https://doi.org/10.1007/s00401-013-1200-z

7. Ash PEA, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, Van Blitterswijk MM, Jansen-West K, Paul JW, Rademakers R, Boylan KB, Dickson DW, Petrucelli L (2013) Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 77:639–646. https://doi.org/10.1016/j.neuron.2013.02.004

8. Mori K, Weng S, Arzberger T, May S, Rentzsch K, Broeckhoven C, Van HC, Edbauer D (2013) The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. Science 339:1335–1339

9. Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, Ostrow LW, Rothstein JD, Troncoso JC, Ranum LPW (2013) RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A 110. https://doi.org/10.1073/pnas.1315438110

10. Mizielinska S, Grönke S, Niccoli T, Ridler CE, Clayton EL, Devoy A, Moens T, Norona FE, Woollacott IOC, Pietrzyk J, Cleverley K, Nicoll AJ, Pickering-brown S, Dols J, Cabecinha M, Hendrich O, Fratta P, Fisher EMC, Partridge L, Isaacs AM (2014) C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science 345:1192–1194

11. Tran H, Almeida S, Moore J, Gendron TF, Chalasani UD, Lu Y, Du X, Nickerson JA, Petrucelli L, Weng Z, Gao FB (2015) Differential toxicity of nuclear RNA foci versus dipeptide repeat proteins in a Drosophila model of C9ORF72 FTD/ALS. Neuron 87:1207–1214. https://doi.org/10.1016/j.neuron.2015.09.015

12. Everett CM, Wood NW (2004) Trinucleotide repeats and neurodegenerative disease. Brain 127:2385–2405. https://doi.org/10.1093/brain/awh278

13. Saberi S, Stauffer JE, Jiang J, Garcia SD, Taylor AE, Schulte D, Ohkubo T, Schloffman CL, Maldonado M, Baughn M, Rodriguez MJ, Pizzo D, Cleveland D, Ravits J (2018) Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. Acta Neuropathol 135:459–474. https://doi.org/10.1007/s00401-017-1793-8

14. Brettschneider J, Del Tredici K, Lee VMY, Trojanowski JQ (2015) Spreading of pathology in neurodegenerative diseases: a focus on human studies. Nat Rev Neurosci 16:109–120. https://doi.org/10.1038/nrn3887

15. Feiler MS, Strobel B, Freischmidt A, Helferich AM, Kappel J, Brewer BM, Li D, Thal DR, Walther P, Ludolph AC, Danzer KM, Weishaupt JH (2015) TDP-43 is intercellularly transmitted across axon terminals. J Cell Biol 211:897–911. https://doi.org/10.1083/jcb.201504057

16. Porta S, Xu Y, Restrepo CR, Kwong LK, Zhang B, Brown HJ, Lee EB, Trojanowski JQ, Lee VMY (2018) Patient-derived frontotemporal lobar degeneration brain extracts induce formation and spreading of TDP-43 pathology in vivo. Nat Commun 9. https://doi.org/10.1038/s41467-018-06548-9

17. Chang YJ, Jeng US, Chiang YL, Hwang IS, Chen YR (2016) The glycine-alanine dipeptide repeat from C9 or f72 hexanucleotide expansions forms toxic amyloids possessing cell-to-cell transmission properties. J Biol Chem 291:4903–4911. https://doi.org/10.1074/jbc.M115.694273

18. Westergard T, Jensen BK, Wen X, Cai J, Kropf E, Iacovitti L, Pasinelli P, Trotti D (2016) Cell-to-cell transmission of dipeptide repeat proteins linked to C9orf72-ALS/FTD. Cell Rep 17:645–652. https://doi.org/10.1016/j.celrep.2016.09.032

19. Zhou Q, Lehmer C, Michaelsen M, Mori K, Alterauge D, Baumjohann D, Schludi MH, Greiling J, Farny D, Flatley A, Feederle R, May S, Schreiber F, Arzberger T, Kuhm C, Klopstock T, Hermann A, Haass C, Edbauer D (2017) Antibodies inhibit transmission and aggregation of C9orf72 poly- GA dipeptide repeat proteins. EMBO Mol Med 9:687–702. https://doi.org/10.15252/emmm.201607054

20. Bass TM, Grandison RC, Wong R, Martinez P, Partridge L, Piper MDW (2007) Optimization of dietary restriction protocols in Drosophila. J Geron - Ser A Biol Sci Med Sci 62:1071–1081. https://doi.org/10.1093/gerona/62.10.1071

21. McGuire SE, Le PT, Osborn AJ, Matsumoto K, Davis RL (2003) Spatiotemporal Rescue of Memory Dysfunction in Drosophila. Science 302:1765–1768. https://doi.org/10.1126/science.1089035

22. Osterwalder T, Yoon KS, White BH, Keshishian H (2001) A conditional tissue-specific transgene expression system using inducible GAL4. Proc Natl Acad Sci U S A 98:12596–12601. https://doi.org/10.1073/pnas.221303298

23. Bischof J, Maeda RK, Hediger M, Karch F, Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific φC31 integrases. Proc Natl Acad Sci U S A 104:3312–3317. https://doi.org/10.1073/pnas.0611511104

24. Mori K, Arzberger T, Grässer FA, Gijselinck I, May S, Rentzsch K, Weng SM, Schludi MH, Van Der Zee J, Cruts M, Van Broeckhoven C, Kremmer E, Kretzschmar HA, Haass C, Edbauer D (2013) Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol 126:881–893. https://doi.org/10.1007/s00401-013-1189-3

25. Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, Wang T, Kim J, Yun J, Xie Y, McKnight SL (2014) Poly-dipeptides encoded by the C9orf72 repeats bind

nucleoli, impede RNA biogenesis, and kill cells. Science 345:1139–1145. https://doi.org/10.1126/science.1254917
26. Zhang YJ, Guo L, Gonzales PK, Gendron TF, Wu Y, Jansen-West K, O'Raw AD, Pickles SR, Prudencio M, Carlomagno Y, Gachechiladze MA, Ludwig C, Tian R, Chew J, DeTure M, Lin WL, Tong J, Daughrity LM, Yue M, Song Y, Andersen JW, Castanedes-Casey M, Kurti A, Datta A, Antognetti G, McCampbell A, Rademakers R, Oskarsson B, Dickson DW, Kampmann M, Ward ME, Fryer JD, Link CD, Shorter J, Petrucelli L (2019) Heterochromatin anomalies and double-stranded RNA accumulation underlie C9orf72 poly(PR) toxicity. Science 363. https://doi.org/10.1126/science.aav2606
27. Babcock DT, Ganetzky B (2015) Transcellular spreading of huntingtin aggregates in the Drosophila brain. Proc Natl Acad Sci U S A 112: E5427–E5433. https://doi.org/10.1073/pnas.1516217112
28. Kreher SA, Kwon JY, Carlson JR (2005) The molecular basis of odor coding in the Drosophila larva. Neuron 46:445–456. https://doi.org/10.1016/j.neuron.2005.04.007
29. Zhang YQ, Rodesch CK, Broadie K (2002) Living synaptic vesicle marker: Synaptotagmin-GFP. Genesis 34:142–145. https://doi.org/10.1002/gene.10144
30. Kanekura K, Yagi T, Cammack AJ, Mahadevan J, Kuroda M, Harms MB, Miller TM, Urano F (2016) Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. Hum Mol Genet 25:1803–1813. https://doi.org/10.1093/hmg/ddw052
31. Saiz-Baggetto S, Méndez E, Quilis I, Igual JC, Bañó MC (2017) Chimeric proteins tagged with specific 3xHA cassettes may present instability and functional problems. PLoS One 12:1–12. https://doi.org/10.1371/journal.pone.0183067
32. Solomon DA, Stepto A, Au WH, Adachi Y, Diaper DC, Hall R, Rekhi A, Boudi A, Tziortzouda P, Lee YB, Smith B, Bridi JC, Spinelli G, Dearlove J, Humphrey DM, Gallo JM, Troakes C, Fanto M, Soller M, Rogelj B, Parsons RB, Shaw CE, Hortobágyi T, Hirth F (2018) A feedback loop between dipeptide-repeat protein, TDP-43 and karyopherin-α mediates C9orf72-related neurodegeneration. Brain 141:2908–2924. https://doi.org/10.1093/brain/awy241
33. Sowade RF, Jahn TR (2017) Seed-induced acceleration of amyloid -β mediated neurotoxicity in vivo. Nat Commun 8:1–12. https://doi.org/10.1038/s41467-017-00579-4
34. Albertson R, Tan V, Leads RR, Reyes M, Sullivan W, Casper-Lindley C (2013) Mapping Wolbachia distributions in the adult Drosophila brain. Cell Microbiol 15:1527–1544. https://doi.org/10.1111/cmi.12136
35. Niccoli T, Partridge L, Isaacs AM (2017) Ageing as a risk factor for ALS/FTD. Hum Mol Genet 26:R105–R113. https://doi.org/10.1093/hmg/ddx247
36. Brettschneider J, Del Tredici K, Irwin DJ, Grossman M, Robinson JL, Toledo JB, Fang L, Van Deerlin VM, Ludolph AC, Lee VMY, Braak H, Trojanowski JQ (2014) Sequential distribution of pTDP-43 pathology in behavioral variant frontotemporal dementia (bvFTD). Acta Neuropathol 127:423–439. https://doi.org/10.1007/s00401-013-1238-y
37. Brettschneider J, Del Tredici K, Toledo JB, Robinson JL, Irwin DJ, Grossman M, Suh E, Van Deerlin VM, Wood EM, Baek Y, Kwong L, Lee EB, Elman L, McCluskey L, Fang L, Feldengut S, Ludolph AC, Lee VMY, Braak H, Trojanowski JQ (2013) Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. Ann Neurol 74:20–38. https://doi.org/10.1002/ana.23937
38. Moens TG, Niccoli T, Wilson KM, Atilano ML, Birsa N, Gittings LM, Holbling BV, Dyson MC, Thoeng A, Neeves J, Glaria I, Yu L, Bussmann J, Storkebaum E, Pardo M, Choudhary JS, Fratta P, Partridge L, Isaacs AM (2019) C9orf72 arginine-rich dipeptide proteins interact with ribosomal proteins in vivo to induce a toxic translational arrest that is rescued by eIF1A. Acta Neuropathol 137:487–500. https://doi.org/10.1007/s00401-018-1946-4
39. Zhang YJ, Gendron TF, Ebbert MTW, O'Raw AD, Yue M, Jansen-West K, Zhang X, Prudencio M, Chew J, Cook CN, Daughrity LM, Tong J, Song Y, Pickles SR, Castanedes-Casey M, Kurti A, Rademakers R, Oskarsson B, Dickson DW, Hu W, Gitler AD, Fryer JD, Petrucelli L (2018) Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. Nat Med 24: 1136–1142. https://doi.org/10.1038/s41591-018-0071-1
40. Freibaum BD, Lu Y, Lopez-Gonzalez R, Kim NC, Almeida S, Lee KH, Badders N, Valentine M, Miller BL, Wong PC, Petrucelli L, Kim HJ, Gao FB, Taylor JP (2015) GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature 525:129–133. https://doi.org/10.1038/nature14974
41. Jovičič A, Mertens J, Boeynaems S, Bogaert E, Chai N, Yamada SB, Paul JW, Sun S, Herdy JR, Bieri G, Kramer NJ, Gage FH, Van Den Bosch L, Robberecht W, Gitler AD (2015) Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat Neurosci 18:1226–1229. https://doi.org/10.1038/nn.4085
42. Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, Daley EL, Miller SJ, Cunningham KM, Vidensky S, Gupta S, Thomas MA, Hong I, Chiu SL, Huganir RL, Ostrow LW, Matunis MJ, Wang J, Sattler R, Lloyd TE, Rothstein JD (2015) The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature 525:56–61. https://doi.org/10.1038/nature14973
43. Conlon EG, Lu L, Sharma A, Yamazaki T, Tang T, Shneider NA, Manley JL (2016) The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. Elife 5: 1–28. https://doi.org/10.7554/eLife.17820
44. Cooper-Knock J, Walsh MJ, Higginbottom A, Highley JR, Dickman MJ, Edbauer D, Ince PG, Wharton SB, Wilson SA, Kirby J, Hautbergue GM, Shaw PJ (2014) Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. Brain 137:2040–2051. https://doi.org/10.1093/brain/awu120
45. Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Štalekar M, Troakes C, Nishimura AL, Scotter EL, Vance C, Adachi Y, Sardone V, Miller JW, Smith BN, Gallo JM, Ule J, Hirth F, Rogelj B, Houart C, Shaw CE (2013) Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. Cell Rep 5:1178–1186. https://doi.org/10.1016/j.celrep.2013.10.049
46. Tao Z, Wang H, Xia Q, Li K, Li K, Jiang X, Xu G, Wang G, Ying Z (2015) Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. Hum Mol Genet 24:2426–2441. https://doi.org/10.1093/hmg/ddv005
47. Guo Q, Lehmer C, Martínez-Sánchez A, Rudack T, Beck F, Hartmann H, Pérez-Berlanga M, Frottin F, Hipp MS, Hartl FU, Edbauer D, Baumeister W, Fernández-Busnadiego R (2018) In situ structure of neuronal C9orf72 poly-GA aggregates reveals proteasome recruitment. Cell 172:696–705.e12. https://doi.org/10.1016/j.cell.2017.12.030
48. Zhang YJ, Gendron TF, Grima JC, Sasaguri H, Jansen-West K, Xu YF, Katzman RB, Gass J, Murray ME, Shinohara M, Lin WL, Garrett A, Stankowski JN, Daughrity L, Tong J, Perkerson EA, Yue M, Chew J, Castanedes-Casey M, Kurti A, Wang ZS, Liesinger AM, Baker JD, Jiang J, Lagier-Tourenne C, Edbauer D, Cleveland DW, Rademakers R, Boylan KB, Bu G, Link CD, Dickey CA, Rothstein JD, Dickson DW, Fryer JD, Petrucelli L (2016) C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. Nat Neurosci 19:668–677. https://doi.org/10.1038/nn.4272
49. Lee JG, Takahama S, Zhang G, Tomarev SI, Ye Y (2016) Unconventional secretion of misfolded proteins promotes adaptation to proteasome dysfunction in mammalian cells. Nat Cell Biol 18:765–776. https://doi.org/10.1038/ncb3372
50. Lasagna-Reeves CA, Castillo-Carranza DL, Sengupta U, Guerrero-Munoz MJ, Kiritoshi T, Neugebauer V, Jackson GR, Kayed R (2012) Alzheimer brain-derived tau oligomers propagate pathology from endogenous tau. Sci Rep 2. https://doi.org/10.1038/srep00700
51. Khosravi B, Hartmann H, May S, Möhl C, Ederle H, Michaelsen M, Schludi MH, Dormann D, Edbauer D (2017) Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in C9orf72 ALS/FTLD. Hum Mol Genet 26:790–800. https://doi.org/10.1093/hmg/ddw432
52. Nonaka T, Masuda-Suzukake M, Hosokawa M, Shimozawa A, Hirai S, Okado H, Hasegawa M (2018) C9ORF72 dipeptide repeat poly-GA inclusions promote intracellular aggregation of phosphorylated TDP-43. Hum Mol Genet 27:2658–2670. https://doi.org/10.1093/hmg/ddy174
53. Baborie A, Griffiths TD, Jaros E, Perry R, McKeith IG, Burn DJ, Masuda-Suzukake M, Hasegawa M, Rollinson S, Pickering-Brown S, Robinson AC, Davidson YS, Mann DMA (2015) Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. Neuropathol Appl Neurobiol 41:601–612. https://doi.org/10.1111/nan.12178
54. Proudfoot M, Gutowski NJ, Edbauer D, Hilton DA, Stephens M, Rankin J, Mackenzie IRA (2014) Early dipeptide repeat pathology in a frontotemporal dementia kindred with C9ORF72 mutation and intellectual disability. Acta Neuropathol 127:451–458. https://doi.org/10.1007/s00401-014-1245-7
55. Vatsavayai SC, Yoon SJ, Gardner RC, Gendron TF, Vargas JNS, Trujillo A, Pribadi M, Phillips JJ, Gaus SE, Hixson JD, Garcia PA, Rabinovici GD, Coppola G, Geschwind DH, Petrucelli L, Miller BL, Seeley WW (2016) Timing and

significance of pathological features in C9orf72 expansion-associated frontotemporal dementia. Brain 139:3202–3216. https://doi.org/10.1093/brain/aww250
56. Ding X, Ma M, Teng J, Teng RKF, Zhou S, Yin J, Fonkem E, Huang JH, Wu E, Wang X (2015) Exposure to ALS-FTD-CSF generates TDP-43 aggregates in glioblastoma cells through exosomes and TNTs-like structure. Oncotarget 6: 24178–24191. https://doi.org/10.18632/oncotarget.4680
57. Asai H, Ikezu S, Tsunoda S, Medalla M, Luebke J, Haydar T, Wolozin B, Butovsky O, Kügler S, Ikezu T (2015) Depletion of microglia and inhibition of exosome synthesis halt tau propagation. Nat Neurosci 18:1584–1593. https://doi.org/10.1038/nn.4132
58. Mackenzie IRA, Frick P, Grässer FA, Gendron TF, Petrucelli L, Cashman NR, Edbauer D, Kremmer E, Prudlo J, Troost D, Neumann M (2015) Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. Acta Neuropathol 130: 845–861. https://doi.org/10.1007/s00401-015-1476-2
59. van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E, Murray ME, Heckman MG, Diehl NN, Brown PH, Baker MC, Finch NCA, Bauer PO, Serrano G, Beach TG, Josephs KA, Knopman DS, Petersen RC, Boeve BF, Graff-Radford NR, Boylan KB, Petrucelli L, Dickson DW, Rademakers R (2013) 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. https://doi.org/10.1016/S1474-4422(13)70210-2
60. Fournier C, Barbier M, Camuzat A, Anquetil V, Lattante S, Clot F, Cazeneuve C, Rinaldi D, Couratier P, Deramecourt V, Sabatelli M, Belliard S, Vercelletto M, Forlani S, Jornea L, Brice A, Auriacombe S, Blanc F, Bouteleau-Bretonnière C, Ceccaldi M, Didic M, Dubois B, Duyckaerts C, Etcharry-Bouix F, Golfier V, Hannequin D, Lacomblez L, Le Ber I, Levy R, Michel BF, Pasquier F, Thomas-Anterion C, Pariente J, Sellal F, Benchetrit E, Bertin H, Bertrand A, Bissery A, Bombois S, Boncoeur MP, Cassagnaud P, Chastan M, Chen Y, Chupin M, Colliot O, Delbeucq X, Delmaire C, Gerardin E, Hossein-Foucher C, Habert MO, Lautrette G, Lebouvier T, Lehéricy S, Le Toullec B, Martineau K, Mackowiak MA, Monteil J, Petyt G, Pradat PF, Oya AH, Rollin-Sillaire A, Salachas F, Sayah S, Wallon D, Leguern E (2019) 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. https://doi.org/10.1016/j.neurobiolaging.2018.09.010
61. Gijselinck I, Van Mossevelde S, Van Der Zee J, Sieben A, Engelborghs S, De Bleecker J, Ivanoiu A, Deryck O, Edbauer D, Zhang M, Heeman B, Bäumer V, Van Den Broeck M, Mattheijssens M, Peeters K, Rogaeva E, De Jonghe P, Cras P, Martin JJ, De Deyn PP, Cruts M, Van Broeckhoven C (2016) The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. Mol Psychiatry 21: 1112–1124. https://doi.org/10.1038/mp.2015.159
62. Wegmann S, Bennett RE, Delorme L, Robbins AB, Hu M, McKenzie D, Kirk MJ, Schiantarelli J, Tunio N, Amaral AC, Fan Z, Nicholls S, Hudry E, Hyman BT (2019) Experimental evidence for the age dependence of tau protein spread in the brain. Sci Adv 5. https://doi.org/10.1126/sciadv.aaw6404

Publisher's Note