

<https://doi.org/10.1038/s41531-025-00899-z>

# Longitudinal multi-omics in alpha-synuclein *Drosophila* model discriminates disease- from age-associated pathologies in Parkinson's disease

Check for updates

Justin Moore<sup>1,2,3</sup>, Timothy Wu<sup>1,3,4</sup>, Justin Dhindsa<sup>1,3,4</sup>, Omar El Fadel<sup>1,3</sup>, Anh Le<sup>1,3</sup>, Alma Perez<sup>1,3</sup>, Bismark Amoh<sup>1,3</sup>, Akash Tarkunde<sup>1,3</sup>, Katy F. Zhu<sup>1,3</sup>, Matthew Avalos<sup>1,3</sup>, Eric B. Dammer<sup>5</sup>, Duc M. Duong<sup>5</sup>, Nicholas T. Seyfried<sup>5,6</sup>, Joshua M. Shulman<sup>1,3,7,8,9</sup>, Ismael Al-Ramahi<sup>1,3,7,9</sup> & Juan Botas<sup>1,2,3,9,10</sup> ✉

Parkinson's disease (PD) starts decades before symptoms appear, usually in the later decades of life, when age-related changes are occurring. To identify molecular changes early in the disease course and distinguish PD pathologies from aging, we generated *Drosophila* expressing alpha-synuclein ( $\alpha$ Syn) in neurons and performed longitudinal bulk transcriptomics and proteomics on brains at six time points across the lifespan and compared the data to healthy control flies as well as human post-mortem brain datasets. We found that translational and energy metabolism pathways were downregulated in  $\alpha$ Syn flies at the earliest timepoints; comparison with the aged control flies suggests that elevated  $\alpha$ Syn accelerates changes associated with normal aging. Unexpectedly, single-cell analysis at a mid-disease stage revealed that neurons upregulate protein synthesis and nonsense-mediated decay, while glia drive their overall downregulation. Longitudinal multi-omics approaches in animal models can thus help elucidate the molecular cascades underlying neurodegeneration vs. aging and co-pathologies.

Neurodegenerative diseases are, by definition, progressive: the dysfunction begins in one particular brain region or neuronal type and intensifies over time, ramifying outward to alter the functions of multiple brain regions. Such is the case with Parkinson's disease (PD). Loss of dopaminergic neurons in the substantia nigra pars compacta and serotonergic neurons in the dorsal raphe nuclei inhibits the thalamus and eventually affects the entire basal ganglia circuit and neocortex<sup>1</sup>, leading to the development of the classic PD motor triad of tremor, bradykinesia, and rigidity after a period of many years. Of course, aging is also progressive, and it is the chief risk factor for the development of PD<sup>2</sup>.

Developing a treatment that could alter the course of PD would require understanding the chain of events that leads from the earliest pathologic changes to later disease stages. A number of transcriptomic and proteomic

studies have analyzed human PD brain tissue<sup>3–9</sup>, but to our knowledge, only one has sought to map transcriptomic changes throughout the progression of PD<sup>6</sup>—and given the scarcity of samples from people who died before developing advanced PD, they had to combine Braak stages 1–4, which makes it difficult to trace the evolution of the disease from early to late stages. Another complication of human post-mortem studies is that the brains of many elderly PD patients show pathological signs of Alzheimer's disease (AD) as well, with amyloid plaques and tau tangles<sup>10–12</sup>. These co-pathologies can be as prominent as Lewy bodies, which are characteristic of PD<sup>13,14</sup> and may influence gene expression. Performing parallel studies on mouse models of AD and PD could help delineate the changes induced by a specific disease-related protein<sup>15–17</sup>, but because these diseases are usually idiopathic in humans it is challenging to develop accurate animal

<sup>1</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030, USA. <sup>2</sup>Quantitative and Computational Bioscience Graduate Program, Baylor College of Medicine, Houston, TX, 77030, USA. <sup>3</sup>Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, 77030, USA. <sup>4</sup>Medical Scientist Training Program, Baylor College of Medicine, Houston, TX, 77030, USA. <sup>5</sup>Department of Biochemistry, Emory University School of Medicine, Atlanta, GA, 30322, USA. <sup>6</sup>Department of Neurology, Emory University School of Medicine, Atlanta, GA, 30322, USA. <sup>7</sup>Department of Neuroscience, Baylor College of Medicine, Houston, TX, 77030, USA. <sup>8</sup>Department of Neurology, Baylor College of Medicine, Houston, TX, 77030, USA.

<sup>9</sup>Center for Alzheimer's and Neurodegenerative Diseases, Baylor College of Medicine, Houston, TX, 77030, USA. <sup>10</sup>Genetics and Genomics Graduate Program, Baylor College of Medicine, Houston, TX, 77030, USA. ✉e-mail: [jbotas@bcm.edu](mailto:jbotas@bcm.edu)



models<sup>18–23</sup>. Models of genetic forms of PD such as transgenic alpha-synuclein (SNCA) mice or *LRRK2*, *PRKN*, or *PINK1* mutant mice do not replicate the dopaminergic degeneration seen in PD or the classic motor triad but have nonetheless been useful for surveys of differential expression, although most studies have been cross-sectional and have considered transcripts or proteins, but not both<sup>18–20</sup>.

*Drosophila melanogaster* models, however, have provided a powerful means to complement human and mouse studies and capture the progression and mechanisms of neurodegeneration because of their conservation of disease-related pathways, their genetic tractability, and suitability for high-throughput experimentation<sup>24–26</sup>. A transgenic PD fly model expressing pan-neuronal human  $\alpha$ Syn recapitulates many PD pathologies such as  $\alpha$ Syn accumulation and aggregation, dopaminergic neuron degeneration, and progressive locomotor deficits<sup>26–29</sup>. We therefore generated  $\alpha$ Syn flies, similar to those previously described<sup>26</sup>, to conduct longitudinal RNA-seq and tandem mass tag-based mass spectrometry (TMT-MS) experiments to better interpret existing human post-mortem data.

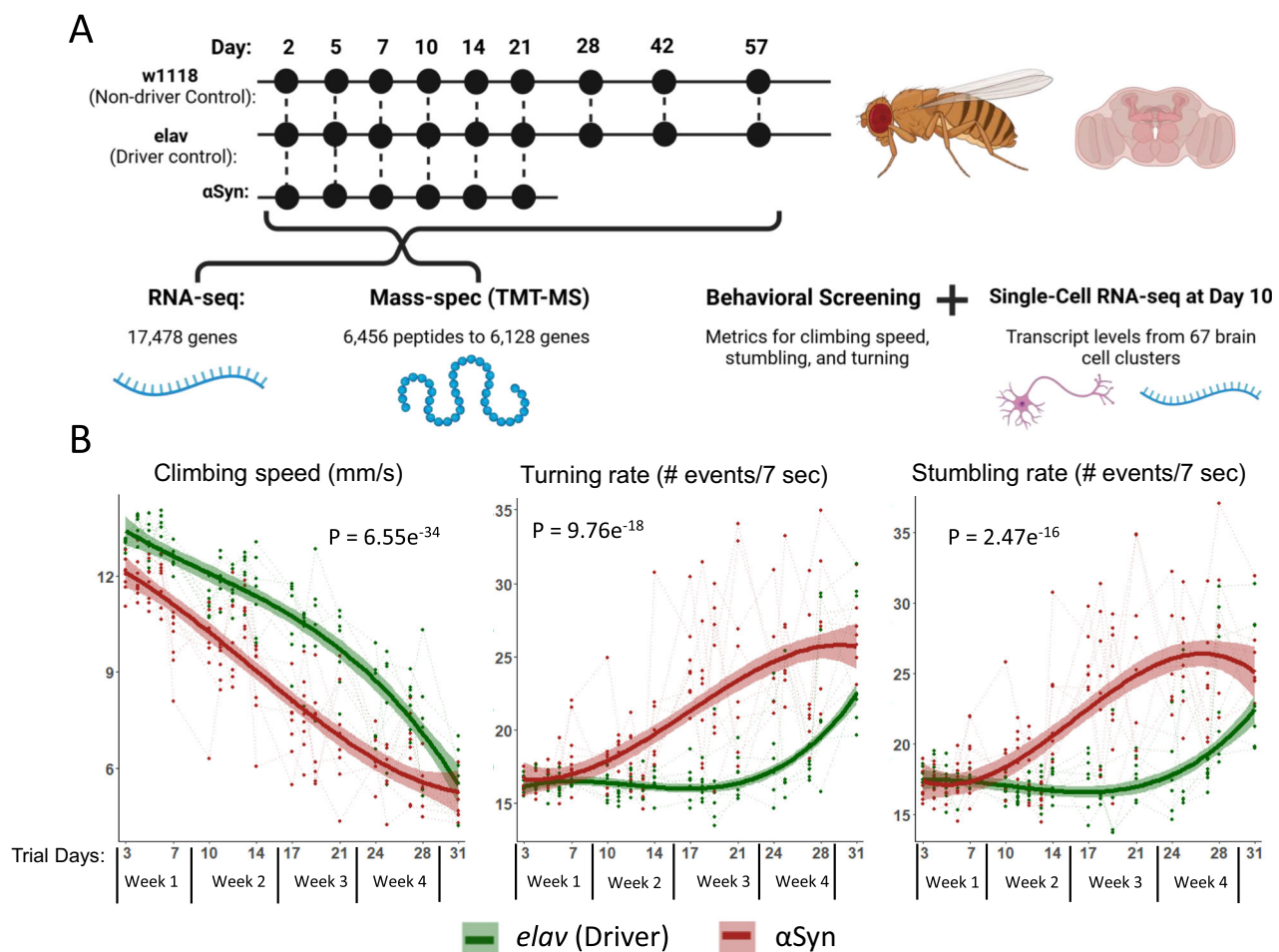
## Results

### Cross-sectional differential expression analysis identifies concordant signatures of $\alpha$ Syn neurotoxicity

To gain insight into mechanisms of pathogenesis we generated flies expressing  $\alpha$ Syn<sup>26</sup> in neurons under the pan-neuronal *elav*-Gal4 driver.

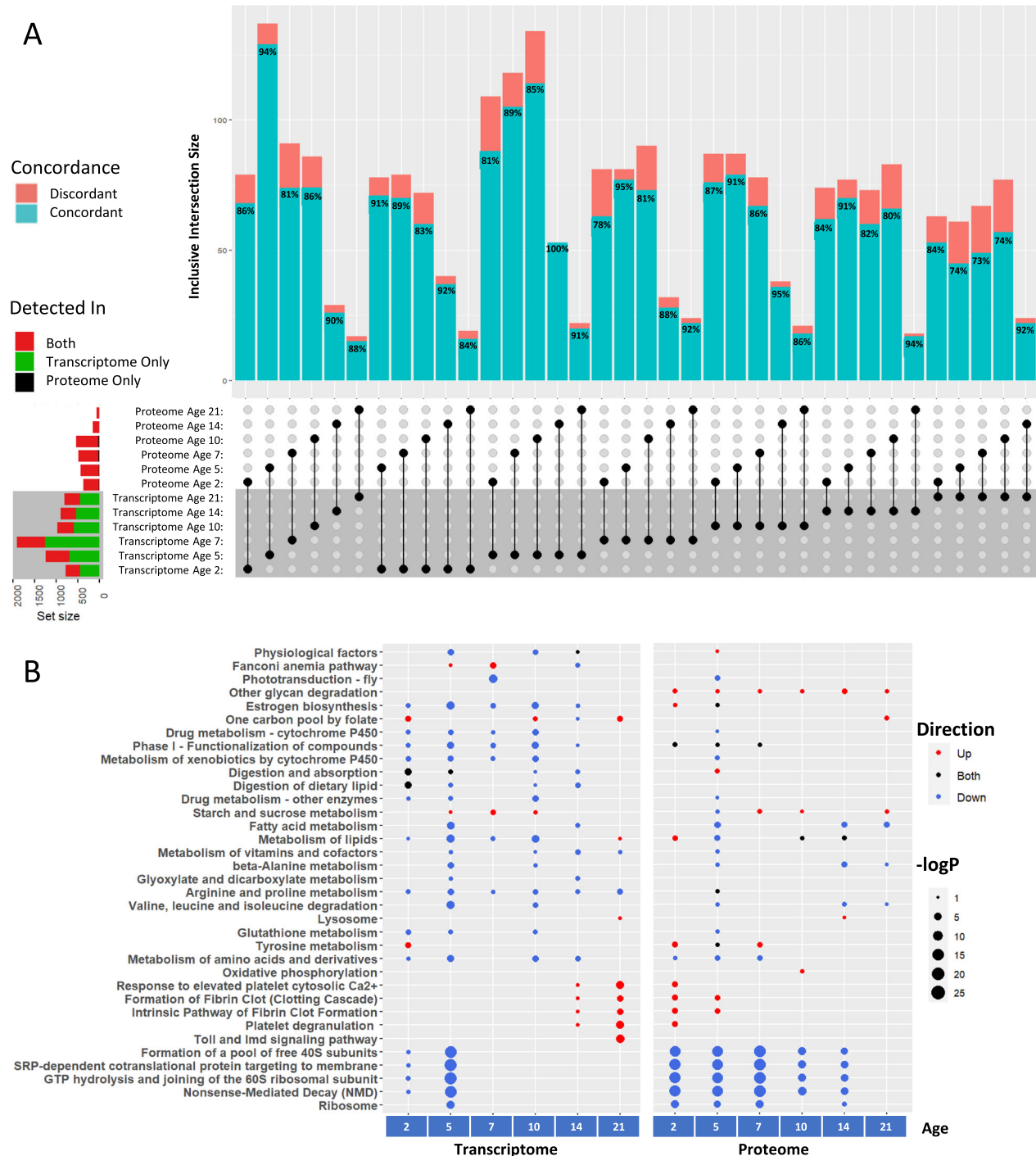
Immunoblotting analyses in these flies show the expected expression of the  $\alpha$ Syn transgene consistent with age-dependent *elav* activity<sup>30</sup>, and immunofluorescence studies reveal the aggregation of intracellular  $\alpha$ Syn positive inclusions that are reminiscent of Lewy-bodies found in humans<sup>31</sup> and disease models<sup>32,33</sup> (Supplementary Fig. 1). We collected brain samples for RNA-seq and TMT-MS studies at nine time points for the control genotypes (up to day 56 post-eclosion, PE) and six time points for the  $\alpha$ Syn genotype (up to day 21 PE, since they have a reduced lifespan; Fig. 1A), to directly compare the signatures of normal aging and  $\alpha$ Syn pathology over the entire lifespan. Longitudinal behavioral screening of  $\alpha$ Syn flies continues up to 31 days, while samples for omics analyses are collected only until day 21. This discrepancy arises from the high attrition rate in the  $\alpha$ Syn cohort, which is compatible with the smaller sample sizes needed for behavioral screening but insufficient for the larger sample sizes required for transcriptomic and proteomic analyses. The pan-neuronal expression of  $\alpha$ Syn significantly accelerated age-dependent behavioral impairments such as loss of climbing speed and increased rates of stumbling and turning (Fig. 1B). In total, RNA-seq captured transcripts from 17,478 genes, mass-spec captured 6456 unique peptides belonging to 6128 proteins/genes; 5817 genes had both transcript and protein quantifications. Quality control metrics such as variance partition and dimension reduction can be found in Supplementary Figs. 2 and 3.

To disentangle age-dependent and disease-associated alterations, we analyzed differential expression (DE) cross-sectionally, comparing the PD



**Fig. 1 | Approach to longitudinal multi-omic data collection from *Drosophila* disease model and controls. A** Samples were collected at multiple timepoints up to 57 days post-eclosion. RNA-seq (3 replicates each), mass-spec (TMT-MS, 3 replicates each), and behavioral screening were done at each time point. In addition, single-cell RNA-seq was performed on dissected fly brains at a single timepoint, day 10. **B** Longitudinal behavioral metrics performed on the same cohort of flies used for

multi-omic analyses reveal age-dependent decreases in climbing speed and increases in stumbling and turning rate. Note that longitudinal behavioral screening of the  $\alpha$ Syn flies extends to 31 days, while samples collected for omic-analysis end at day 21; this is a consequence of the high attrition of the  $\alpha$ Syn cohort that is compatible with the limited numbers required for behavioral screening, but not compatible with the numbers required for transcriptomic and proteomic analyses.



**Fig. 2 | Cross-sectional differential expression analysis identifies concordant signatures of  $\alpha$ Syn neurotoxicity.** **A** DEGs and DEPs overlap significantly, and concordance varies between 73% and 100%. UpSet plot shows the inclusive overlaps and genes are considered DE if they have an FDR < 0.05, and a  $|\log_2FC| > 0.5$  for the transcriptome and 0.25 for the proteome. Concordance percentages compare the concordance of effect direction for the overlapping DEGs/DEPs. Horizontal bars on

the left side of the graph represent the total number of DEGs/DEPs at that time point and the proportion of those genes that were detected in RNA-seq only (green), mass-spec only (black), or in both (red). **B** Functional enrichments (Reactome and KEGG) of the cross-sectional DEGs and DEPs. Circle color reflects the direction of dysregulation and circle sizes are scaled continuously from the  $-\log_{10}(p\text{-value})$  of the enrichment.

model flies with the healthy aging control lines at each time point (Fig. 2A, Supplementary Fig. 4, Supplementary Tables 1 and 2). Analysis of the transcriptome and proteome yielded similar proportions of differentially expressed genes (DEGs) and proteins (DEPs) at most time points, although significance cutoffs were calibrated to match the different technical limitations of RNA-seq and TMT-MS. To validate the results of TMT-MS, we analyzed an additional independent cohort of flies at a single time point

(Day 10, Supplementary Fig. 5). The results were largely concordant with the original cohort and show that our DE results are reliably reproducible. When comparing the DE results for the original RNA-seq and TMT-MS cohorts, there were more DEGs and DEPs in the middle of the disease course (e.g. Day 10 had 969 DEGs and 537 DEPs) than at the first (778 DEGs and 375 DEPs) and last (1007 DEGs and 63 DEPs) time points tested (Supplementary Fig. 4). This trend was not explained by technical artifacts such

as batch effects or missing data but reflects the fact that at the end stage of the disease, neuronal loss and consequent changes in the cellular composition of the brain tissues studied limit the ability to find DEGs. Similar trends have been found in other animal models of neurodegeneration<sup>34</sup>.

As expected, there were far more DEGs than DEPs. Given this imbalance, the amount of overlap between transcriptome and proteome (DEGs and DEPs) varied between 4% and 16% depending on the time point. The overlap between these DEGs and DEPs increased to up to 25% when limiting the analysis to only those genes detected in both RNA-seq and TMT-MS. To determine the concordance of these overlapping DEGs and DEPs, we compared the direction of expression change at each time point and found high (73–94%) concordance (Fig. 2A). We also correlated the genome-wide transcriptomic and proteomic log-fold changes between  $\alpha$ Syn and driver controls (Supplementary Fig. 6). For all comparisons, there was a positive correlation between transcript and protein. Interestingly, the lowest (but still positive) correlations between transcriptomic and proteomic datasets occurred at a late time point (day 21). The lower correlation in late-stage disease parallels the age-dependent incongruence in transcript and protein levels that has been shown in post-mortem human brain tissue from PD cases<sup>35</sup>.

Since ~66% of human genes are conserved in *Drosophila*, functional annotation is made possible through cross-species inferred interactions, which fill in an otherwise incomplete annotation database<sup>36</sup>. Therefore, we used *Drosophila* Reactome and KEGG over-representation analyses on the upregulated and downregulated genes from the cross-sectional -omics data (Fig. 2B, Supplementary Table 3). From the earliest stages, the most strongly downregulated pathways were translational (the formation of 40S subunits, SRP-dependent cotranslational protein targeting the membrane, and GTP hydrolysis during translation). The transcriptomic changes were most notable on days 2 and 5 while proteomic data showed changes from days 2 to 14. Dysregulation of the ribosomal machinery has been gaining increasing recognition in PD<sup>37</sup> and has previously been observed in post-mortem PD<sup>4,5,38,39</sup>; our results indicate that the change occurs well before the end stages of the disease.

Among the most upregulated pathways was the fibrin clotting cascade, an evolutionary byproduct of the innate immune system (and often categorized as such), which was upregulated late (days 14 and 21) in the transcriptome but early (days 2 and 5) in the proteomic data. This interesting temporal discrepancy seems to be caused by an enrichment of SERPIN (Serine Protease Inhibitors) gene dysregulation in both models, but the exact genes are different. Specifically, *Nec*, *Spn28Dc*, *Spn42De*, *Spn31A*, *Spn47C*, and *Spn77Bc* were upregulated late in the transcriptome, but *Spn27A*, *Spn42Da*, *Spn47C*, *Spn55B*, and *Spn43Ab* were upregulated early in the proteome. Interestingly, human homologs of these SERPINs (*SERPINH1*, *SERPINF2*, *SERPINC1*, *SERPINA3*, *SERPINA1*, and *SERPINE2*), which are related to inflammation and coagulation, are concordantly upregulated in post-mortem brain RNA-seq and proteomic data<sup>3,5,40–42</sup>. Upregulation of inflammation-related genes has been observed in human post-mortem data, but this data suggests that this upregulation begins early in the disease.

To investigate the significance of SERPIN dysregulation we knocked down a selected group of upregulated SERPINs in the *Drosophila*  $\alpha$ Syn model and ran longitudinal behavioral screening (Supplementary Fig. 7). This analysis revealed that knocking down *Spn27A*, *Spn31A*, *Spn42Da*, *Spn42De* and *Spn47C* exacerbated the behavioral deficits induced by  $\alpha$ Syn neurotoxicity suggesting that the upregulation of these SERPINs is a compensatory response to the disease. Control experiments knocking down the same SERPINs in healthy flies did not change their behavior (Supplementary Fig. 8).

The dysregulated signatures (DEGs and DEPs) observed in  $\alpha$ Syn flies are not caused by overexpression of any protein because they are absent in the controls overexpressing GAL4. To further investigate their specificity we compared the  $\alpha$ Syn signatures with previously reported tau DEGs and DEPs<sup>43</sup>. Briefly, we overlapped the longitudinal tau DEGs and DEPs and performed functional enrichment on the overlapping

signatures and the signatures specific to  $\alpha$ Syn (Supplementary Fig. 9). Our analysis suggests that dysregulation of metabolism-related pathways is specific to the  $\alpha$ Syn transcriptome. Furthermore, we found that the proteomic dysregulation of pathways related to protein synthesis and NMD have both  $\alpha$ Syn-specific and overlapping signatures with the tau proteome. Interestingly, a deeper look into the direction of change of the overlapping tau and  $\alpha$ Syn NMD-related genes reveals that tau and  $\alpha$ Syn alter the abundance of some of the same proteins but in different directions.

### **$\alpha$ Syn alters aging-related changes in translation, energy metabolism, and innate immunity**

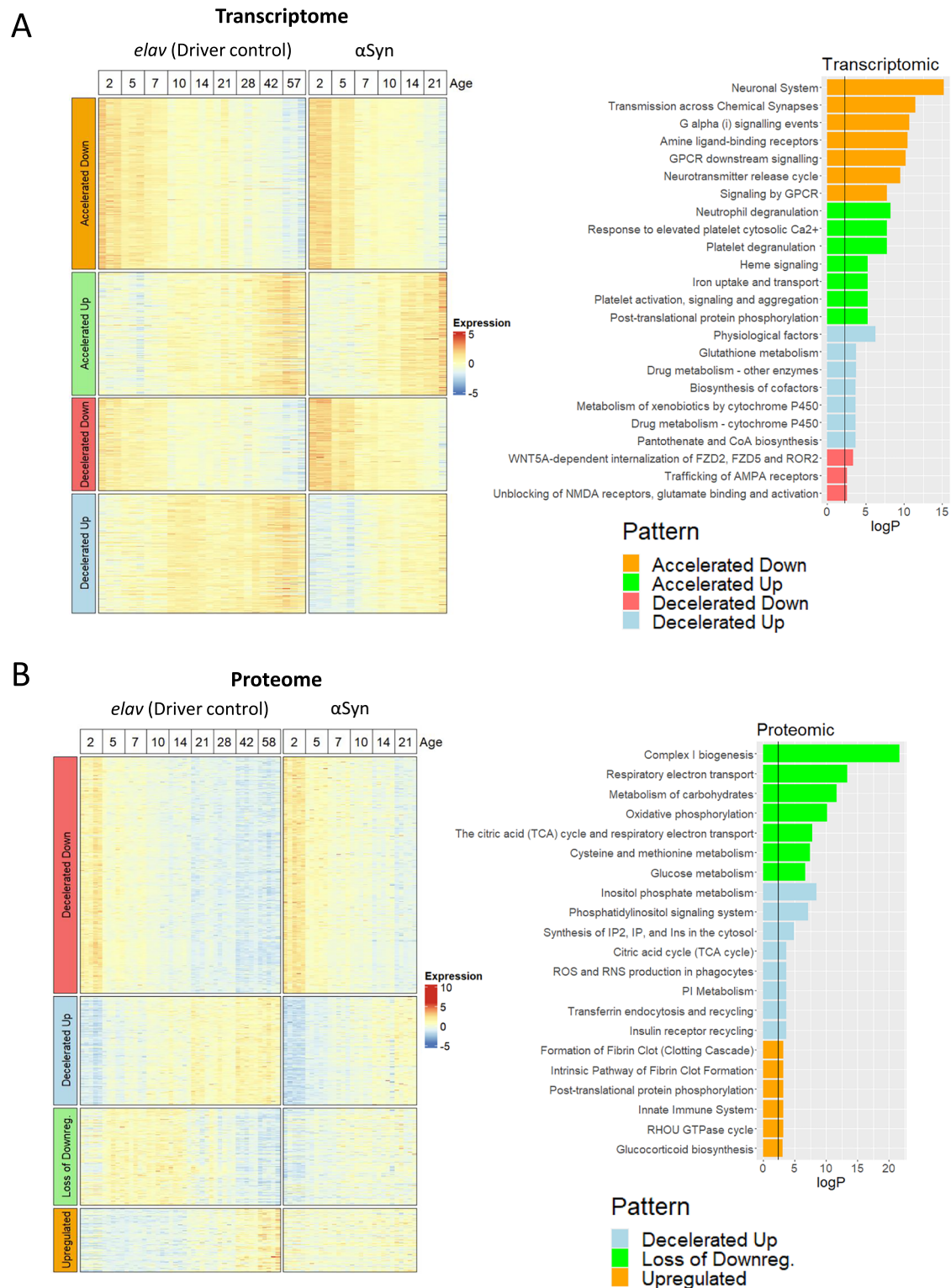
Several studies have noted that the molecular signatures of disease progression are comparable to what would typically be seen later in the course of normal aging leading to postulate an accelerated aging hypothesis of PD<sup>44–49</sup>. To explore this idea, we ran longitudinal DE analyses of both the transcriptome and proteome and identified genes that were significant relative to both the age and genotype\*age interaction term. These genes, termed age-associated DEGs, have normal expression changes with age in the control flies, but their up- or down-regulation is either accelerated or decelerated in the transgenic  $\alpha$ Syn fly. For the transcriptome, we used an age-adjusted DESeq2 model to identify 1140 age-associated DEGs; for the proteomic data, we used age-adjusted linear models to identify 1130 DEPs. Although only 629 (55%) of the transcriptomic age-associated DEGs had their respective protein detected in TMT-MS, 120 (19%) of these were also age-associated in the proteome analysis, suggesting these aging trends are reflected in both.

Next, we generated expression heatmaps to compare the age-dependent expression patterns in the control and the disease model flies (Fig. 3A, B, Supplementary Tables 4–9). Since the transcriptome trends are generally linear, we classified these categories as accelerated or decelerated up/down with age. In the proteome, the expression patterns are more complicated, so we categorized additional trends as a loss of age-related downregulation (i.e., downregulated at early stages but upregulated compared to late) or as an age-independent upregulation (Supplementary Fig. 10 shows representative expression plots for each of these trends.) It is important to note that these expression patterns are general trends with age and do not represent all possible patterns.

Next, we sought to identify the biological pathways affected by this accelerated aging. Functional annotation of the transcriptomic ‘accelerated aging’ signatures (Fig. 3A, right panel) identified an early upregulation of innate immunity pathways such as neutrophil degranulation and platelet activation. In other words, innate immunity genes are upregulated over the course of normal aging, and  $\alpha$ Syn neurotoxicity accelerates this trend. This supports the hypothesis of accelerated aging in PD and specifically points to age-dependent chronic inflammation (inflammaging) as a key player. Pathways whose downregulation was accelerated (see Day 21 at the top of the heatmap in Fig. 3A) were generally related to synaptic/neuronal biology, likely because of neuronal impairments or death.

The proteomic data show upregulation of similar innate immunity pathways (innate immune system, fibrin clot formation, glucocorticoid biosynthesis) (Fig. 3B). Innate immunity proteins are expressed at low levels in the young fly and high levels in the aged healthy fly but are consistently upregulated from early time points in the disease course. These findings confirm our previous results (Fig. 2B) indicating that these pathways are dysregulated from early in the disease course and become exacerbated with age. Only two other patterns in the proteome had significant functional enrichments: decelerated upregulation and loss of downregulation. In both cases, energy metabolism pathways (Complex I Biogenesis, Oxidative phosphorylation, TCA cycle, glucose metabolism) were downregulated, depending on the age. Other interesting enrichments include downregulation of phosphatidylinositol signaling (dme04070) and ROS/RNS production in phagocytes (R-DME-1222556). In summary,  $\alpha$ Syn neurotoxicity accelerates a number of aging-related expression changes.





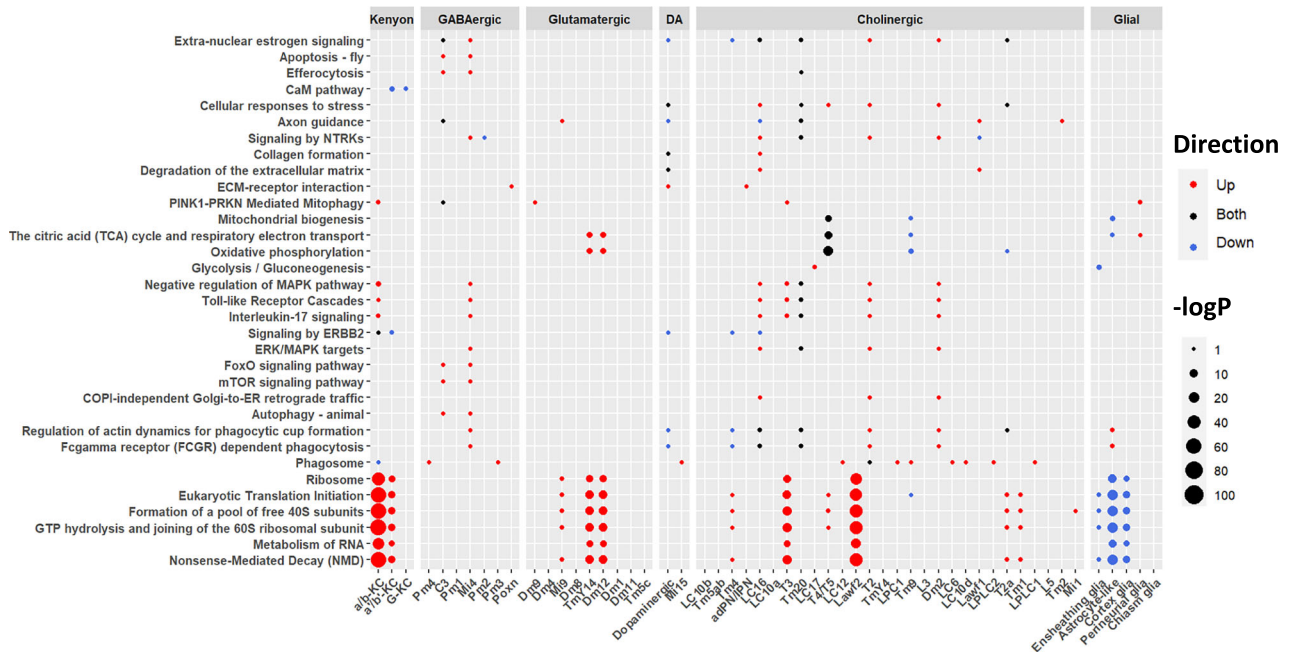
**Fig. 3 |  $\alpha$ Syn alters aging-related changes in translation, energy metabolism, and innate immunity.** **A** Expression heatmap of the 1140 age-adjusted DEGs in the transcriptome and the Reactome and KEGG functional enrichments for each pattern. **B** Heatmap of the 1130 age-adjusted DEPs in the proteome and their functional enrichments. Columns are annotated by genotype and age. The rows are split into

four categories based on the expression trend with age and the direction of the genotype\*age interaction term. The vertical black line in the functional enrichments is a 0.1 significance cutoff. The proteome “decelerated down” pattern had no significant enrichments.

### Single-cell RNA-seq data reveal divergent neuronal and glial responses to $\alpha$ Syn neurotoxicity

An obvious limitation of bulk RNA sequencing and mass-spectrometry is that they obscure differences between cell types. Therefore, we generated a

single-cell RNA-seq dataset at a single time point, 10 days post-eclosion, with the same *elav* (driver control) and  $\alpha$ Syn model used in the longitudinal experiments. These data captured ~65,000 cells and assigned cell identities for 67 of the 239 cell clusters (Supplementary Fig. 11).



**Fig. 4 | Single-cell RNA-seq data reveal divergent neuronal and glial responses to  $\alpha$ Syn neurotoxicity.** Single-cell differential expression was run using Seurat integrated DESeq2 and functional enrichment was run on the output using Reactome and KEGG ontologies. Both DEGs and functional enrichments were corrected for

FDR with a significance cutoff at  $FDR < 0.05$ . Circle color reflects the direction of dysregulation and circle sizes are scaled continuously from the  $-\log_{10}(p\text{-value})$  of the enrichment. Translation and NMD pathways (bottom of the figure) are upregulated in several neuronal cell types but downregulated in most glia.

To test whether pan-neuronal expression of human  $\alpha$ Syn changes the cell-type composition of the diseased brain, we used standard DESeq2 differential abundance protocol on each of the cell clusters. After FDR correction, none of the 67 cell clusters proved to be differentially abundant. This could indicate that day 10 is still sufficiently early in the disease course that there is not enough cell loss to produce a statistically significant difference by this method. For example, dopaminergic neurons were among the most underrepresented in the  $\alpha$ Syn model relative to the control at day 10, though this difference did not reach statistical significance (Supplementary Fig. 12). Loss of TH-positive, dopaminergic neurons, in the *elav* >  $\alpha$ Syn fly model has been documented at later stages (post day 20)<sup>29,50</sup>. Overall, these data lend confidence to the results of the longitudinal bulk data by showing that early time points are not confounded by changes in cell type composition.

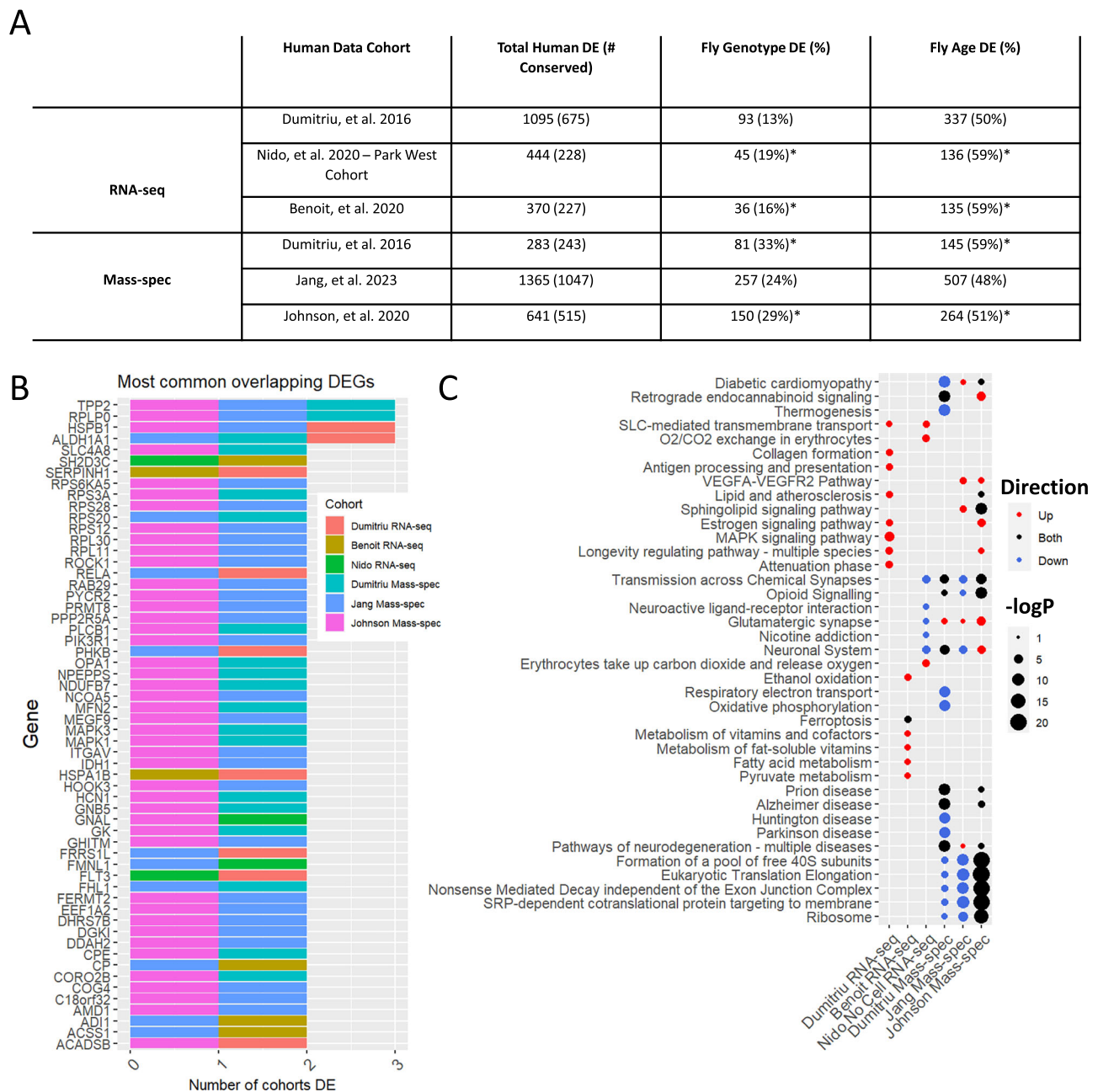
To learn which cell types contributed most to the bulk gene expression changes, we took the genes that had been differentially expressed in the bulk data and examined their expression in the scRNA-seq data. The two datasets accorded well with each other but with an interesting twist: transcriptomic upregulation at all time points was explained primarily by upregulation across neurons, including dopaminergic neurons, while both transcriptomic and proteomic downregulation were primarily glial (Supplementary Fig. 13). To further investigate these results, we ran expression-weighted cell-type enrichment on the bulk DEGs which revealed that the expression of the downregulated proteins is over-represented in healthy glial cells relative to neurons (Supplementary Fig. 14). Thus, the transcriptomic downregulation of pathways in Fig. 2B primarily reflects the comparative loss of expression in glia. We asked whether dysregulated transcripts and upregulated proteins also have high expression in glia depending on the age analyzed, but we found no consistent pattern (Supplementary Figs. 14, 15). This response of glial cells to  $\alpha$ Syn toxicity is cell nonautonomous and likely caused either by glia reacting to the neuronal-specific expression of  $\alpha$ Syn, or the propagation of oligomeric or fibrillar forms of  $\alpha$ Syn from neurons to glial cells<sup>51</sup>.

Finally, we sought to determine the pathways that are altered within each cell cluster using single-cell DE analyses such as DESeq2, as implemented in Seurat. Overall, functional enrichment analysis of the single-cell DEGs identified pathways similar to those seen in the bulk differential

expression (Figs. 2B and 4, Supplementary Table 10). For example, the most significant enrichment among several clusters was in pathways related to translation and nonsense-mediated decay. Interestingly, although the enrichment was significant in both neurons and glia, the direction of effect on translation was opposite in the two cell types: upregulated in neuronal cell clusters such as a/b-KC (Kenyon cells) and Lawf2 (cholinergic) cells, while downregulated in glial cells (specifically astrocyte-like, cortex, and ensheathing glia) (Fig. 4, Supplementary Table 10). Given that in the *Drosophila* brain neurons outnumber glia<sup>52</sup>, it was surprising that the bulk downregulation of protein synthesis pathways is explained by glial downregulation, but again, glia overrepresents the expression of these pathways (Supplementary Figs. 13, 14). Clusters such as a/b-KC, T4/T5 neurons, ensheathing glia, and astrocyte-like glia were among the most altered clusters with the highest DEG count (Supplementary Fig. 16). The pathways that are dysregulated in both neurons and glia, and in both post-mortem human scRNA-seq and our disease model flies, are related to energy metabolism, MAPK pathways, and Rho GTPase signaling<sup>7,8</sup>.

### Drosophila model data help distinguish disease and age-associated changes in PD patient data

As noted earlier, synuclein, tau, and amyloid deposits often coexist in the brains of individuals with PD. To distinguish the pathogenic effects of these three distinct proteins in patient data, we looked for genes that were differentially expressed in both humans and the  $\alpha$ Syn flies (Fig. 5A, Supplementary Table 11). First, we took genes that had been identified as differentially expressed by RNA-seq in human PD from three independent cohorts: Dumitriu et al.<sup>3</sup> (1095 genes, prefrontal cortex), Nido et al.<sup>4</sup> (444 genes, prefrontal cortex), and Benoit et al.<sup>40</sup> (370 genes, frontal lobe). Next, we identified the fly homologs of these DEGs by homology mapping, which yielded 675, 228, and 227 genes for the respective cohorts. We then overlapped these homologs with our longitudinally differentially expressed *Drosophila* genes with respect to genotype ( $\alpha$ Syn expression) and age. Only 13–19% of the human DEGs were differentially expressed in the flies with respect to genotype, whereas roughly half (50–59%) were associated with aging. The difference in the overlap between genotype and age suggests that age has a greater influence on these signatures than disease.



**Fig. 5 | *Drosophila* model data helps distinguish disease and age-associated changes in PD patient data.** **A** Table showing the degree of overlap between DEGs and DEPs in PD patient samples and those in a longitudinal aSyn *Drosophila* model. Homology mapping was done using DIOPT at a score of 3 or greater. Fly columns represent the proportion of human DEGs that overlap with the longitudinal fly model DEGs/DEPs relative to the genotype or age term. An asterisk indicates that

the overlap size is statistically significant ( $p < .05$ ), calculated from a permutation test of 1000 iterations. **B** Histogram of the most common overlapping human DEGs/DEPs and their respective cohorts. All genes shown overlap with longitudinal fly DEGs. **C** Functional enrichments of the overlapping human and fly DEGs/DEPs. Circle color reflects the direction of dysregulation relative to log2FC of human DE and circle sizes are scaled continuously from the  $-\log_{10}(p\text{-value})$  of the enrichment.

We then compared the DEPs identified in mass-spec data from Dumitriu et al.<sup>3</sup> (283 proteins, prefrontal cortex), Jang et al.<sup>5</sup> (1365 proteins, substantia nigra), and Johnson et al.<sup>41</sup> (641 proteins, dorsolateral prefrontal cortex) with our synucleinopathy fly mass-spec data. Homology mapping identified 243, 1047, and 515 *Drosophila* homologs for each respective cohort. Of these, 24–33% were also differentially expressed in the longitudinal mass-spec fly data with respect to genotype and 48–59% with respect to age. To assess the statistical significance of the overlaps between our fly data and the human datasets, we ran a permutation test (1000 iterations) by permuting gene labels and recalculating the overlap size. Of the six human datasets, four overlapped significantly (permutation test  $p < 0.05$ ) with our *Drosophila* data for both genotype and age. Furthermore, KEGG overrepresentation analysis

for all overlapping DEPs revealed enrichment of PD-related genes (hsa05012, Parkinson's disease) for both genotypes ( $p = 6.4e^{-4}$ ) and age ( $p = 1.98e^{-5}$ ). We also compared the human DEGs to cross-sectional fly DE data and found overlap among all sampled time points (Supplementary Fig. 17).

Age influences a vast number of physiological processes, and aSyn toxicity anticipates or accelerates many of these age-related changes. To identify specific pathways that were more prominently disrupted by aSyn expression than age, we ran functional enrichment on the human PD genes that overlapped with the aSyn fly DEGs but did not significantly change with age in the control flies. This analysis yielded pathways such as Primary active transmembrane transporter activity (GO:0015399, FDR = 0.00023), Endocrine and other factor-regulated calcium reabsorption (hsa04961,



FDR =  $1.09e^{-05}$ ), and Dopaminergic synapse (hsa04728, FDR = 0.0012). See Supplementary Table 12 for a complete view of enrichments. Although some genes in these pathways may change with age, our data suggest these pathways are more affected by  $\alpha$ Syn toxicity.

### **Drosophila model data distinguish changes caused by $\alpha$ Syn from those caused by co-pathologies in patient data**

The overlapping human and fly DEGs are most likely to represent the genes and pathways whose dysfunction is downstream of  $\alpha$ Syn pathogenesis rather than from amyloid, tau, or other co-pathologies. We, therefore, plotted the most common overlapping DEGs (Fig. 5B) and identified the functional enrichments for each cohort (Fig. 5C, Supplementary Table 13). Overall, 4 DEGs in our fly data were identified in 3 out of 6 human cohorts, 54 were identified in 2 human cohorts, and 476 in only one human cohort. Interestingly, ALDH1A1, a gene that was downregulated in 3 human cohorts and flies, has been found to harbor variants that modulate PD risk<sup>33</sup> and has been previously studied in PD<sup>54,55</sup>. Functional enrichment of each of these overlaps yielded commonly dysregulated pathways related to ribosomes, translation machinery, neuronal biology, energy metabolism, and longevity-regulating pathways such as MAPK signaling. Interestingly, translation and nonsense-mediated decay pathways were downregulated in two of the human mass-spec datasets, as they were in our bulk mass-spec data (Fig. 2B), whereas the Johnson et al. dataset showed both down- and up-regulation of these pathways. These results support the usefulness of the *Drosophila*  $\alpha$ Syn model to reproduce the gene expression signatures of human PD. More importantly, they distinguish processes that are downstream  $\alpha$ Syn neurotoxicity and aging and are independent of comorbidities.

### **Discussion**

The results of this cross-species transcriptomic and proteomic analysis support the premature aging and “inflammaging” hypotheses for the changes occurring in PD. Inflammation has been proposed to contribute strongly to the aging process and may influence neurodegenerative disease onset<sup>56</sup>. Indeed, centenarians who have evaded age-related neurodegeneration do not bear classic signs of age-related inflammation<sup>57–59</sup>. It is conceivable that the increased risk of PD with age may be due in part to the compounding effects of  $\alpha$ Syn-induced inflammation and age-associated inflammation. It may, therefore, be possible to counter the progression of PD by modifying the feedback loops associated with accelerated aging, inflammation, and PD progression. Our data identified specific innate immune targets that contribute to these mechanisms, such as *Spn27A*, *Spn31A*, *Spn42Da*, *Spn42De* and *Spn47C*, which belong to a family of serine protease inhibitors (SERPINs). Human SERPINs such as *SERPINA1* and *SERPINE1*, have established roles in regulating inflammation<sup>60–62</sup>, but their role in neurodegeneration is just beginning to be explored. Our data suggests that the upregulation of at least some SERPINs is a compensatory mechanism since their experimental knockdown exacerbates the  $\alpha$ Syn fly behavioral phenotype. Overall, our results implicate SERPIN dysregulation as an age-dependent downstream effect of  $\alpha$ Syn neurotoxicity, although further research is required to investigate their therapeutic potential.

Previous studies in neurodegenerative diseases have found that the correlation between RNA-seq and mass spec data ranges from 0.08 to 0.27<sup>43,63,64</sup>. The coupling of the transcriptome and proteome may decline with age and is exacerbated in PD due to aberrant proteostasis<sup>35</sup>, but in this study, the correlation between proteomic and transcriptional data ranged from 0.07 to 0.3. Furthermore, we observed DEG/DEP concordance greater than ~80%. We note that a prior, similar study<sup>43</sup> using tauopathy flies found a rather lower concordance, due to technical differences in significance cutoffs for DEG identification.

The most significantly downregulated pathways we identified in the bulk transcriptome, bulk proteome, and single-cell data involved translational machinery and nonsense-mediated decay, which are also strongly downregulated in two post-mortem human mass-spec datasets<sup>3,5</sup> and both upregulated and downregulated in another<sup>41</sup>. Interestingly, our results suggest that the transcriptomic downregulation is specific to early time

points while proteomic downregulation may extend into later stages. This may explain why protein synthesis pathways were detected only in human post-mortem mass-spec and not in RNA-seq datasets. (We were unable to compare human post-mortem scRNA-seq because the relevant studies had removed ribosomal mRNAs from the analysis<sup>7,8</sup>.) Although the connection between protein translation and PD has been largely overlooked in research, a recent review<sup>37</sup> notes that human studies repeatedly find dysregulation of these pathways<sup>5,9,38,39,65,66</sup>. Based on this recurrent observation, researchers have hypothesized that downregulated protein synthesis could be an attempt either to conserve energy or to modulate the levels of aggregated  $\alpha$ Syn<sup>37,39</sup>. Our results do not fully support either of these hypotheses, since they reveal that glia downregulated these pathways while some neurons upregulated the translational machinery. This suggests that neurons increase protein synthesis, which is not consistent with either conservation of energy or reducing  $\alpha$ Syn aggregation. A third hypothesis<sup>39</sup> states that under stress conditions, the cell will reprogram the proteome by increasing the translation of stress-related proteins while decreasing the translation of non-essential proteins. A fourth possibility is that  $\alpha$ Syn is upregulated inadvertently while neurons increase protein synthesis of compensatory pathways such as phagocytosis or autophagy to manage toxicity. Meanwhile, glial cells downregulate protein synthesis and glycolysis to conserve energy for the neurons. Support for this fourth model comes from the fact that deficits in energy metabolism pathways, specifically glycolysis<sup>67,68</sup> and oxidative phosphorylation<sup>69,70</sup>, are commonly observed in PD. Our data are consistent with this last model and reveal that proteomic downregulation of glycolysis, insulin signaling, oxidative phosphorylation, and complex I biogenesis are all downstream of  $\alpha$ Syn aggregation.

In this context, the neuronal upregulation and glial downregulation of nonsense-mediated decay (NMD) seems particularly noteworthy: although the role of NMD has not been studied in synucleinopathies, there is mounting evidence for its pathogenic role in other neurodegenerative diseases such as Amyotrophic Lateral Sclerosis (ALS), Frontotemporal Dementia, and Alzheimer’s disease<sup>71–75</sup>. NMD is an RNA surveillance pathway that selectively degrades transcripts with a premature termination codon<sup>76</sup>. The efficient clearance of error-containing transcripts by NMD is essential for neuronal health as it plays key roles in learning, development, synaptic plasticity, and axonal guidance<sup>77</sup>. Zuniga et al.<sup>72</sup> showed that pan-neuronal expression of wild-type and mutant tau in *Drosophila* hinders NMD, whereas genetic or pharmacological activation of NMD suppressed tau-induced neurodegeneration. This suggested that NMD impairment is a major mechanism through which tau exerts its neurotoxicity. Similar studies in cellular and *Drosophila* ALS models showed the expression of a neurotoxic C9orf72 dipeptide repeat impaired NMD pathway activity, but overexpression of core NMD genes (*UPF1* & *UPF2*) mitigated neurotoxicity<sup>71,73</sup>. Our results are consistent with the hypothesis of neurotoxic  $\alpha$ Syn altering NMD and further refine our understanding by suggesting NMD impairment may be cell-type specific. These results suggest the possibility of testing whether experimental reactivation of NMD exerts a therapeutic effect in PD as it does in models of ALS and tauopathy.

In sum, we show that longitudinal ‘omics data from a disease model can complement similar datasets from human postmortem studies, mitigating the limitations of terminal samples. Not only can a combined analysis distinguish disease-related from aging-related changes, but it can disentangle the changes wrought by co-pathologies—in this case, the tendency of PD patients to also accumulate other disease-triggering proteins such as tau or amyloid. We defined the molecular alterations that are downstream of  $\alpha$ Syn; provided support for the accelerated aging model of PD; and find that glia mount a non-cell autonomous response to  $\alpha$ Syn toxicity that runs counter to that of neurons. Further analysis of these temporal and cell-specific alterations could promote the discovery of early biomarkers and potential therapeutic targets.

### **Methods**

#### **Drosophila models and stocks**

Flies were raised at 25 °C on *Drosophila* molasses-based media. To generate a Parkinson’s disease model fly line, we crossed a female virgin elav-Gal4/+



with a male UAS- $\alpha$ Syn line containing wild-type human  $\alpha$ Syn. We generated two sets of controls: we crossed the female virgin elav-Gal4/+ driver control line with a non-targeting shRNA (VDR: #2691) male fly, and we generated a non-driver control without elav-Gal4 (w1118/+) by crossing the same non-targeting shRNA line with a wild-type fly of the same genetic background (VDR: #60000).

Samples for bulk RNA-seq ( $n = 3$  biological replicates) and mass-spec ( $n = 5$  biological replicates) were collected at 6 timepoints up to day 21 post-eclosion (days 2, 5, 7, 10, 14, and 21) for the  $\alpha$ Syn genotype. To capture age-related changes in the wild-type context, we collected samples from the control genotypes at 9 time points up to day 57 post-eclosion (days 2, 5, 7, 10, 14, 21, 28, 42, and 57). For single-cell RNA-seq ( $n = 3$  biological replicates) we used only flies that were 10 days post-eclosion.

### Drosophila longitudinal behavior assays

We assessed neurological function by automatically quantifying climbing speed, turning rate, and stumbling rate in a high-throughput behavioral assay that takes advantage of the flies' instinctive negative geotaxis<sup>78</sup> as described before Onur et al.<sup>79</sup>. We assessed statistical significance using longitudinal mixed effects (ANOVA) models with smoothing splines to capture non-linear trends over time; we report  $p$ -values for additive and interaction effects in the figures.

### Immunoblotting

Female *Drosophila* were raised at 25 °C, then frozen at −80 °C at 5, 10, 15, 20, and 25 days post eclosion. For each replicate, 8 fly heads were collected, lysed, and homogenized in 30  $\mu$ L of lysis buffer (25% 4 $\times$  NuPage LDS Sample Buffer, 10% beta-mercaptoethanol, and 65% DI water). Samples were centrifuged for 10 min and then boiled at 100 °C for another 10 min. The sample was loaded into a Bis-Tri NuPage (BioRad) gel and run at 80 V for one hour. Proteins were then transferred to a Nitrocellulose membrane using 20% methanol transfer buffer at 0.3 amps for 1.5 h. After transfer, membranes were washed with TBST and blocked using 5% non-fat blocking milk in 1 $\times$  TBST. Antibodies were diluted in blocking solutions at concentrations of 1:1000 for  $\alpha$ Syn antibody (BD Biosciences #610787), and 1:10,000 for  $\beta$ -Tubulin antibody (Proteintech #66240-1-Ig). Membranes were treated with primary antibodies at 4 °C overnight and treated with 1:10,000 secondary antibodies the following morning. After treatment with a secondary antibody, the membranes were imaged using a LI-COR Odyssey CLx fluorescence imager. Each genotype was run with 3 replicates and  $\alpha$ Syn band intensity was normalized to loading control  $\beta$ -Tubulin intensity.

### Whole-mount staining and confocal microscopy

Using aged female flies (30 days post-eclosion), brains were dissected in ice-cold 1 $\times$  PBS and immediately fixed for 35 min using a 4% paraformaldehyde solution (PFA). Brains were washed three times for 5 min each using 1 $\times$  PBST. Samples were placed in a blocking buffer (1% bovine serum albumen in PBST) for 1 h at room temperature and then incubated with anti- $\alpha$ Syn (1:200, BD Biosciences #610787) for 3 days at 4 °C. After the primary was removed, samples were washed 3 times with 1 $\times$  PBST for 5 min each, then incubated in secondary (1:500 anti-Mouse-555, and 1:1000 DAPI) for 1 h at room temperature. Samples were removed from the secondary and washed 4 times using 1 $\times$  PBST, one time in 1 $\times$  PBS, and then mounted onto slides. Whole-mount stained slides were imaged using a confocal microscope (Leica SP8X) with a  $\times 63$  objective and 4 $\times$  zoom taking z-stack images of the midbrain. Images were visualized using the LasX Office software.

### RNA-seq data generation and processing

We used three biological replicates per genotype per time point for RNA-seq. Each replicate consisted of 100 heads collected from flash-frozen, age-matched virgin female flies. mRNA was extracted using TRIzol (#15596026, Invitrogen) followed by DNase treatment. A minimum of

500 nanograms of total DNase-treated extract RNA were used per replicate. Samples were prepared using the Illumina TruSeq Stranded mRNA Library Prep Kit. Samples were then sequenced on an Illumina NovaSeq 6000 with 350 bp paired-end reads. Raw reads were aligned to *Drosophila* reference genome dm6 r6.06. Genes with an average read count <50 across all samples were excluded. Preparation of sequencing libraries, RNA-sequencing, and alignment was performed by the New York Genome Center. Initially, these datasets captured read counts for 17,478 genes. RNA-seq read counts were filtered by removing the bottom 5% of expressed genes and further analyzed using DESeq2 adjusted for the preparation batch. After filtering, these datasets provide the counts of 16,604 genes from RNA-seq for further analysis.

### Mass-spectrometry data generation and processing

Five biological replicates per genotype per time point were generated. Tandem mass tag mass-spectrometry proteomics was performed following previously published protocols<sup>80</sup>. Homogenates were prepared from dissected adult female fly brains (50 heads per biological replicate). We removed protein isoforms and samples with >30% missingness and imputed the remaining missing values using the minimum value divided by 2. Peptides mapping to multiple isoforms were collapsed by sum. After initial processing, we captured information for 6456 peptides of 6128 unique genes. Mass-spec protein abundances were processed using the tunable median polish of ratio (TAMPOR) approach<sup>81</sup>. After filtering, these datasets provided the abundance of 6456 peptides collapsed to 6044 proteins for further analyses.

### Mass-spectrometry validation dataset generation

Using the same models and TMT-MS processing methods, we generated an independent sample at day 10 with 8 replicates each, run in a single batch. This validation set was analyzed for differential expression using a two-sided  $t$ -test and then compared to the original cohort.

### Dimension reduction for bulk RNA-Seq and mass spec data

Dimension reduction for the RNA-seq data was done on the DESeq2 normalized read counts and the built-in PCA capabilities of the DESeq2 R package. For the proteomic data, we used principal component analysis (PCA) as implemented in the stats R package version 4.2.2.  $t$ -distributed stochastic neighbor embedding for both models was done using the Rtsne R package version 0.16 with default parameters.

### Differential expression analysis for bulk RNA-Seq and TMT-MS data

For the bulk RNA-seq data, we used the standard DESeq2<sup>82</sup> protocol on the filtered read count data to compare and identify genes whose expression differed between the  $\alpha$ Syn and driver control genotypes.  $P$ -values were adjusted for FDR using Benjamini–Hochberg (BH) correction, and genes were considered differentially expressed if they had an FDR < 0.05 and log2 fold change (FC) > 0.5.

For the TMT-MS data, we ran a  $t$ -test on the TAMPOR-corrected protein abundances and adjusted  $p$ -values for FDR using the BH method. Proteins were considered differentially expressed if they had an FDR < 0.05 and a log2FC > 0.25.

For longitudinal differential expression, an age-adjusted DESeq2 model (Expr ~ Genotype + Age + Genotype\*Age) was used for RNA-seq, and an age-adjusted linear model was used for mass-spec (Expr ~ Genotype + Age + Genotype\*Age). Age was encoded as a numerical variable. Significance was assessed relative to the genotype term, and cutoffs for the age-adjusted models were set to  $<1 \times 10^{-4}$  for transcriptomic differences and  $<1 \times 10^{-3}$  for proteomic differences. Cutoffs for the accelerated aging trends were set for the age term (adjusted  $p \leq 0.05$ ) and the interaction term (adjusted  $p \leq 0.05$ ).

### Functional enrichment

Genes that were dysregulated were used as input for functional enrichment using the ReactomePA R package (Version 1.42) for Reactome and the

clusterProfiler package (Version 4.6.2) for KEGG (Kyoto Encyclopedia of Genes and Genomes) Pathway Database. For the background, we used all detected genes from RNA-seq or all detected proteins from mass-spec. *P*-values were adjusted for FDR using the Benjamini–Hochberg method. Given the hierarchical structure of the Reactome pathways, many significant enrichments are redundant because of overlapping signals. Therefore, we manually removed some redundancies and displayed only the most significant pathways within each hierarchy. Reactome pathway relationships were identified using Reactome website and the rbiapi R package.

### Human RNA-seq and mass-spectrometry data

We obtained human RNA-seq data from three published sources, Nido et al.<sup>4</sup> performed RNA-sequencing with ribosomal depletion from the prefrontal cortex of two cohorts, Park-west (PW) and the Netherlands Brain Bank (NBB). Because the authors' analysis of the NBB cohort did not identify any differentially expressed genes, we focused our analyses on the PW cohort ( $n = 18$  idiopathic PD and  $n = 11$  healthy controls). The authors performed differential expression analyses under two models (no cell-type correction and cell-type corrected), and we utilized the non-cell-type corrected in our results. We also removed all lncRNAs from the analysis due to their low conservation between humans and *Drosophila*. A second source of human RNA-seq, published by Benoit et al.<sup>40</sup> was gathered from living tissue from the frontal lobe. Dumitriu et al.<sup>3</sup> provided our third source of human RNA-seq data, from the substantia nigra of PD patients.

Mass-spec data from the Alzheimer's Disease Brain Proteome was recently made available by Johnson et al.<sup>41</sup> For our purposes in the current study, we selected samples from patients with mild to severe Lewy body pathology ( $n = 62$ ) and compared them to healthy controls ( $n = 69$ ). TAMPOR-normalized protein abundances were corrected for age, sex, and post-mortem interval. We analyzed differential expression using a two-sided *t*-test with FDR adjustment and DE cutoffs set at FDR < 0.05.

Taking the differentially expressed genes from each resource, we used DIOPT to identify *Drosophila* homologs for each human gene. In brief, a homolog was assigned to a human gene if it had a DIOPT score >3. A human gene overlaps as differentially expressed with *Drosophila* if it has a homolog that is differentially expressed in the longitudinal *Drosophila* analysis relative to genotype or age, respectively. Human RNA-seq data was overlapped *Drosophila* model RNA-seq data only, and human mass-spec data was overlapped with *Drosophila* model mass-spec only. The significance of these overlaps was calculated using a permutation test where *p*-values from human DE were randomly permuted, and then the overlap was recalculated; 1000 iterations were used to calculate a *p*-value. Benoit et al.<sup>40</sup> did not release complete DE results, so instead of permutation we substituted a hypergeometric test with parameters calculated from our data: population size = 13,676, number of successes in population = 1587 for genotype or 5568 for age, sample size = 227, and sample successes are 36 for genotype or 135 for age.

### Single-cell RNA-seq

Three biological replicates of females at 10 days post-eclosion were generated per genotype (21 samples total). For each replicate, dissected brains (18 per replicate) were enzymatically dissociated into a single-cell suspension as previously described<sup>83</sup>. Single-cell RNA-seq libraries were prepared by the Baylor College of Medicine Single Cell Genomics Core using the Chromium Single Cell Gene Expression 3' v3.1 kit (10x Genomics) and then were sequenced by the Baylor Genomic and RNA Profiling Core. Sequencing was performed using the Illumina NovaSeq 6000 platform with a minimum depth of 300,000,000 reads per sample. Illumina BCL files were demultiplexed using the Cell Ranger 4.0.0 mkfastq function. The resultant FASTQ files were aligned to the *Drosophila* reference genome (BDGP6.22.98). Gene counts were quantified using the Cell Ranger 4.0.0 count pipeline and filtered count matrices were loaded into Seurat<sup>84</sup> in R for additional quality control and downstream analyses. Cells were filtered from the data if the proportion of mitochondrial reads per cell was >20% or the total number of unique genes per cell was fewer than 200 or >3000.

To predict and remove heterotypic doublets we used the DoubletFinder function within the Seurat package. Briefly, for each library, artificial doublets were generated from the existing data, and PCA was performed after merging the real and artificial data. Using the first 40 PCs, we generated a distance matrix to compute the proportion of artificial *K*-nearest-neighbors (pANN) for each cell. We estimated the PC neighborhood size for each library as previously described<sup>85</sup>, estimated the number of doublets per library, ranked the cells using pANN, and removed the suspected doublets.

We used SCTransform<sup>86</sup> to normalize gene expression independently per library using regularized negative binomial regression. All libraries normalized via SCTransform were integrated using the reciprocal principal component analysis (rPCA) pipeline in Seurat to correct for batch effects and identify similar cell identities across conditions. 3 *elav*-Gal4 controls were merged to create an integration reference, and anchors across libraries were computed over the first 100 rPCA dimensions in the combined dataset. We used the NormalizeData function in Seurat for the normalization of gene counts for computing cell cluster marker genes, cell identity annotation, and other applications directly comparing gene expression levels between cell clusters. Unique molecular identifiers (UMI) for each gene in each cell were divided by the sum of UMIs in that cell, multiplied by a scalar (10,000), and log-transformed to create the final integrated and normalized dataset.

### Single-cell RNA-Seq clustering and cell-type identification

We first used the FindNeighbors function in the Seurat package to construct a nearest-neighbor (SNN) graph of the integrated dataset. Next, we used the FindClusters function to identify clusters of cells using a shared nearest-neighbor modularity optimization-based clustering algorithm. Because cluster count varies with function parameters, we performed a grid search of 64 different combinations of clustering parameters. We varied the number of principal components from 80 to 150 (in step sizes of 10) in creating the SNN graph and varied the clustering resolution from 1 to 8 (step size of 1). This yielded between 179 and 368 clusters, and each cluster's biological relevance was compared to published bulk transcriptomes from 52 purified *Drosophila* neuronal types. We determined the similarity of expression profiles by calculating the Pearson correlation coefficient for each purified cell type and the average log-normalized non-integrated gene expression of each cluster. For correlation calculation, we used the most variable genes found during clustering, and for each pair of clustering parameters, we counted the number of purified neuronal cell types that matched 1–5 clusters. Specifically, for the top 6 most highly correlated clusters for each neuronal type, we tested whether the difference in correlation value between a cluster and the subsequent cluster was >0.05, as has been described by a prior publication Ozel et al.<sup>87</sup>. We annotated the dataset after it was clustered with 120 principal components, which yielded 239 clusters and 67 unique annotated cell types.

To assign these clusters to specific cell types, we utilized three methods of annotation. First, we compared the average log-normalized non-integrated gene expression of each cluster to published bulk transcriptomes of purified neuronal populations. We calculated the Pearson correlation coefficient using the top 10 cluster markers—the genes ranked among the top 10 most differentially expressed genes on the basis logFC for at least one cluster when compared to all other clusters<sup>87</sup>. Second, we used four published single-cell datasets—Wu et al. (2023), Ozel et al. (2020), Davie et al. (2018), and Li et al.<sup>83,87–89</sup>—as references and performed cell-label transfer using Seurat's FindTransferAnchors and TransferData functions. In brief, we identified pairs of similar cells between the reference and query datasets using a mutual nearest neighbor approach after projecting the new dataset onto the reference dataset in PCA-reduced space. Each cell was assigned a predicted label based on the score from the comparison to the reference dataset. Lastly, we utilized a trained two-layer neural network classifier of adult optic lobe cell types Ozel et al.<sup>87</sup>. Non-integrated log-normalized expression values of ~500 marker genes were used as the input to train the classifier. Cluster identity was assigned based on the most abundant population of cells in that cluster. To assign final identities to clusters, we

examined the concordance between our various annotating methods; at least two of the methods were required to predict the cluster as concordant to be annotated.

### Single-cell RNA-seq differential cell type abundance

In the same fashion that DESeq2 uses to test the differential abundance of a transcript using a read count matrix, we used DESeq2 to determine whether a cell type is differentially abundant between the disease model and control using a cell count matrix. We assessed three biological replicates for each model. P-values were calculated using a Wald test statistic and adjusted for FDR using the Benjamini-Hochberg method.

### Single-cell RNA-seq differential expression

We used the FindMarkers function as implemented in the Seurat package with the test set to DESeq2. We required that genes be expressed in >10% of cells and considered genes to be differentially expressed if they had a BH-corrected *p*-value < 0.05. When comparing bulk data to single-cell data in Supplementary Fig. 13, we ran a *t*-test comparing the mean normalized transcript abundances for each group of bulk cross-sectional DEGs between genotypes within each cluster.

### Expression-weighted cell-type enrichment (EWCE)

We used the R implementation of EWCE as described in Skene et al.<sup>90</sup>. The input and output species were set to *Drosophila*. We used a bootstrap test for significance calculation set to 1000 iterations and utilized our scRNA-seq read counts and collapsed cell type for the analysis.

### Data availability

The complete bulk RNA-seq, mass-spec, and single-cell RNA-seq data has been uploaded to Synapse and can be found at <https://www.synapse.org/Synapse:syn34767207>.

### Code availability

All codes used in this project are in publicly available R packages defined in the “Methods” section.

### Abbreviations

αSyn	alpha-synuclein
PD	Parkinson's disease
DE	differential expression
DEGs	differentially expressed genes
DEPs	differentially expressed proteins
NMD	nonsense-mediated decay
AD	Alzheimer's disease
ALS	amyotrophic lateral sclerosis

Received: 15 August 2024; Accepted: 17 February 2025;

Published online: 11 March 2025

### References

- McGregor, M. M. & Nelson, A. B. Circuit mechanisms of Parkinson's disease. *Neuron* **101**, 1042–1056 (2019).
- Reeve, A., Simcox, E. & Turnbull, D. Ageing and Parkinson's disease: why is advancing age the biggest risk factor? *Ageing Res. Rev.* **14**, 19–30 (2014).
- Dumitriu, A. et al. Integrative analyses of proteomics and RNA transcriptomics implicate mitochondrial processes, protein folding pathways and GWAS loci in Parkinson disease. *BMC Med. Genom.* **9**, 5 (2015).
- Nido, G. S. et al. Common gene expression signatures in Parkinson's disease are driven by changes in cell composition. *Acta Neuropathol. Commun.* **8**, 55 (2020).
- Jang, Y. et al. Mass spectrometry-based proteomics analysis of human substantia nigra from Parkinson's disease patients identifies multiple pathways potentially involved in the disease. *Mol. Cell. Proteom.* **22**, 100452 (2023).
- Cappelletti, C. et al. Transcriptomic profiling of Parkinson's disease brains reveals disease stage specific gene expression changes. *Acta Neuropathol. (Berl.)* **146**, 227–244 (2023).
- Martirosyan, A. et al. Unravelling cell type-specific responses to Parkinson's disease at single cell resolution. *Mol. Neurodegener.* **19**, 7 (2024).
- Smajić, S. et al. Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state. *Brain* **145**, 964–978 (2022).
- Zhao, Y. et al. Transcriptomics and weighted protein network analyses of the LRRK2 protein interactome reveal distinct molecular signatures for sporadic and LRRK2 Parkinson's Disease. *Npj Park. Dis.* **10**, 144 (2024).
- Sengupta, U. & Kayed, R. Amyloid β, Tau, and α-Synuclein aggregates in the pathogenesis, prognosis, and therapeutics for neurodegenerative diseases. *Prog. Neurobiol.* **214**, 102270 (2022).
- Shim, K. H., Kang, M. J., Youn, Y. C., An, S. S. A. & Kim, S. Alpha-synuclein: a pathological factor with Aβ and tau and biomarker in Alzheimer's disease. *Alzheimers Res. Ther.* **14**, 201 (2022).
- Moussaud, S. et al. Alpha-synuclein and tau: teammates in neurodegeneration? *Mol. Neurodegener.* **9**, 43 (2014).
- Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M. & Goedert, M. α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl Acad. Sci. USA* **95**, 6469–6473 (1998).
- Shults, C. W. Lewy bodies. *Proc. Natl Acad. Sci. USA* **103**, 1661–1668 (2006).
- Sanchez-Varo, R. et al. Transgenic mouse models of Alzheimer's disease: an integrative analysis. *Int. J. Mol. Sci.* **23**, 5404 (2022).
- Kaye, J., Reisine, T. & Finkbeiner, S. Huntington's disease mouse models: unraveling the pathology caused by CAG repeat expansion. *Fac. Rev.* **10**, 77 (2021).
- Dovonou, A. et al. Animal models of Parkinson's disease: bridging the gap between disease hallmarks and research questions. *Transl. Neurodegener.* **12**, 36 (2023).
- Zhong, J. et al. Single-cell brain atlas of Parkinson's disease mouse model. *J. Genet. Genom.* **48**, 277–288 (2021).
- Jung, S. Y. et al. An anatomically resolved mouse brain proteome reveals Parkinson disease-relevant pathways. *Mol. Cell. Proteom.* **16**, 581–593 (2017).
- Tong, W. et al. Transcriptional profiling reveals brain region-specific gene networks regulated in exercise in a mouse model of Parkinson's disease. *Front. Aging Neurosci.* **14**, 891644 (2022).
- Zhang, T. D. et al. How well do rodent models of Parkinson's disease recapitulate early non-motor phenotypes? A systematic review. *Biomedicines* **10**, 3026 (2022).
- Translational Neurogenetics Unit, Wallenberg Neuroscience Center, Lund University, Lund, Sweden, Konnova, E. A., Swanberg, M., & Translational Neurogenetics Unit, Wallenberg Neuroscience Center, Lund University, Lund, Sweden. Animal models of Parkinson's disease. In *Parkinson's Disease: Pathogenesis and Clinical Aspects* (eds John Van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, UK, Stoker, T. B., Greenland, J. C., & John Van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, UK) 83–106 (Codon Publications, 2018).
- Fleming, S. M., Fernagut, P.-O. & Chesselet, M.-F. Genetic mouse models of parkinsonism: strengths and limitations. *NeuroRX* **2**, 495–503 (2005).
- Wittmann, C. W. et al. Tauopathy in *Drosophila*: neurodegeneration without neurofibrillary tangles. *Science* **293**, 711–714 (2001).
- Feany, M. B. & Bender, W. W. A *Drosophila* model of Parkinson's disease. *Nature* **404**, 394–398 (2000).



26. Chouhan, A. K. et al. Uncoupling neuronal death and dysfunction in *Drosophila* models of neurodegenerative disease. *Acta Neuropathol. Commun.* **4**, 62 (2016).
27. Suzuki, M., Sango, K. & Nagai, Y. Roles of  $\alpha$ -synuclein and disease-associated factors in *Drosophila* models of Parkinson's disease. *Int. J. Mol. Sci.* **23**, 1519 (2022).
28. Auluck, P. K., Chan, H. Y. E., Trojanowski, J. Q., Lee, V. M.-Y. & Bonini, N. M. Chaperone suppression of  $\alpha$ -synuclein toxicity in a *Drosophila* model for Parkinson's disease. *Science* **295**, 865–868 (2002).
29. Ordóñez, D. G., Lee, M. K. & Feany, M. B.  $\alpha$ -synuclein induces mitochondrial dysfunction through spectrin and the actin cytoskeleton. *Neuron* **97**, 108–124.e6 (2018).
30. Delandre, C., McMullen, J. P. D. & Marshall, O. J. Dynamic changes in neuronal and glial GAL4 driver expression during *Drosophila* aging. *Genetics*, iyaf014, <https://doi.org/10.1093/genetics/iyaf014> (2025).
31. Guerreiro, P. S. et al. LRRK2 interactions with  $\alpha$ -synuclein in Parkinson's disease brains and in cell models. *J. Mol. Med.* **91**, 513–522 (2013).
32. Park, S. S. & Lee, D. Selective loss of dopaminergic neurons and formation of Lewy body-like aggregations in  $\alpha$ -synuclein transgenic fly neuronal cultures. *Eur. J. Neurosci.* **23**, 2908–2914 (2006).
33. Volpicelli-Daley, L. A. et al. Exogenous  $\alpha$ -synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron death. *Neuron* **72**, 57–71 (2011).
34. Langfelder, P. et al. Integrated genomics and proteomics define huntingtin CAG length-dependent networks in mice. *Nat. Neurosci.* **19**, 623–633 (2016).
35. Dick, F., Tysnes, O.-B., Alves, G. W., Nido, G. S. & Tzoulis, C. Altered transcriptome-proteome coupling indicates aberrant proteostasis in Parkinson's disease. *iScience* **26**, 105925 (2023).
36. Vastrik, I. et al. Reactome: a knowledge base of biologic pathways and processes. *Genome Biol.* **8**, R39 (2007).
37. Ashraf, D., Khan, M. R., Dawson, T. M. & Dawson, V. L. Protein translation in the pathogenesis of Parkinson's disease. *Int. J. Mol. Sci.* **25**, 2393 (2024).
38. Garcia-Esparcia, P. et al. Altered machinery of protein synthesis is region- and stage-dependent and is associated with  $\alpha$ -synuclein oligomers in Parkinson's disease. *Acta Neuropathol. Commun.* **3**, 76 (2015).
39. Lu, B., Gehrke, S. & Wu, Z. RNA metabolism in the pathogenesis of Parkinson's disease. *Brain Res.* **1584**, 105–115 (2014).
40. Benoit, S. M. et al. Expanding the search for genetic biomarkers of Parkinson's disease into the living brain. *Neurobiol. Dis.* **140**, 104872 (2020).
41. Johnson, E. C. B. et al. Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat. Med.* **26**, 769–780 (2020).
42. Herrero, M.-T., Estrada, C., Maatouk, L. & Vyas, S. Inflammation in Parkinson's disease: role of glucocorticoids. *Front. Neuroanat.* **9**, 32 (2015).
43. Mangleburg, C. G. et al. Integrated analysis of the aging brain transcriptome and proteome in tauopathy. *Mol. Neurodegener.* **15**, 56 (2020).
44. Barbeau, A. Etiology of Parkinson's disease: a research strategy. *Can. J. Neurol. Sci.* **11**, 24–28 (1984).
45. Levy, G. The relationship of Parkinson disease with aging. *Arch. Neurol.* **64**, 1242 (2007).
46. Li, S. et al. The accelerated aging model reveals critical mechanisms of late-onset Parkinson's disease. *BioData Min* **13**, 4 (2020).
47. Bobela, W., Aebischer, P. & Schneider, B. Alpha-synuclein as a mediator in the interplay between aging and Parkinson's disease. *Biomolecules* **5**, 2675–2700 (2015).
48. Collier, T. J., Kanaan, N. M. & Kordower, J. H. Aging and Parkinson's disease: different sides of the same coin? *Mov. Disord.* **32**, 983–990 (2017).
49. Mann, D. M. A. & Yates, P. O. Pathogenesis of Parkinson's disease. *Arch. Neurol.* **39**, 545–549 (1982).
50. Bridi, J. C. et al. Presynaptic accumulation of  $\alpha$ -synuclein causes synaptopathy and progressive neurodegeneration in *Drosophila*. *Brain Commun.* **3**, fcab049 (2021).
51. Lee, H.-J. et al. Direct transfer of  $\alpha$ -synuclein from neuron to astroglia causes inflammatory responses in synucleinopathies. *J. Biol. Chem.* **285**, 9262–9272 (2010).
52. Freeman, M. R. *Drosophila* central nervous system glia. *Cold Spring Harb. Perspect. Biol.* **7**, a020552 (2015).
53. Fan, H.-H. et al. ALDH1A1 genetic variations may modulate risk of Parkinson's disease in Han Chinese population. *Front. Neurosci.* **15**, 620929 (2021).
54. Grünblatt, E. & Riederer, P. Aldehyde dehydrogenase (ALDH) in Alzheimer's and Parkinson's disease. *J. Neural Transm.* **123**, 83–90 (2016).
55. Masato, A., Plotegher, N., Boassa, D. & Bubacco, L. Impaired dopamine metabolism in Parkinson's disease pathogenesis. *Mol. Neurodegener.* **14**, 35 (2019).
56. Calabrese, V. et al. Aging and Parkinson's disease: inflammaging, neuroinflammation and biological remodeling as key factors in pathogenesis. *Free Radic. Biol. Med.* **115**, 80–91 (2018).
57. Franceschi, C. et al. Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans. *Mech. Ageing Dev.* **128**, 92–105 (2007).
58. Gangemi, S. et al. Age-related modifications in circulating IL-15 levels in humans. *Mediators Inflamm.* **2005**, 245–247 (2005).
59. Gerli, R. et al. Chemokines, sTNF-Rs and sCD30 serum levels in healthy aged people and centenarians. *Mech. Ageing Dev.* **121**, 37–46 (2001).
60. Meghadri, S. H. et al. Loss of Serpina1 in mice leads to altered gene expression in inflammatory and metabolic pathways. *Int. J. Mol. Sci.* **23**, 10425 (2022).
61. Shahaf, G. et al.  $\alpha$ -1-Antitrypsin gene delivery reduces inflammation, increases T-regulatory cell population size and prevents islet allograft rejection. *Mol. Med.* **17**, 1000–1011 (2011).
62. Chen, T. et al. Research progress on the SERPINE1 protein and chronic inflammatory diseases of the upper respiratory tract: a literature review. *Int. Arch. Allergy Immunol.* **182**, 1097–1102 (2021).
63. Wei, Y.-N. et al. Transcript and protein expression decoupling reveals RNA binding proteins and miRNAs as potential modulators of human aging. *Genome Biol.* **16**, 41 (2015).
64. Koussounadis, A., Langdon, S. P., Um, I. H., Harrison, D. J. & Smith, V. A. Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. *Sci. Rep.* **5**, 10775 (2015).
65. Zhou, Z. D., Selvaratnam, T., Lee, J. C. T., Chao, Y. X. & Tan, E.-K. Molecular targets for modulating the protein translation vital to proteostasis and neuron degeneration in Parkinson's disease. *Transl. Neurodegener.* **8**, 6 (2019).
66. Sarkar, S. et al. Comparative proteomic analysis highlights metabolic dysfunction in  $\alpha$ -synucleinopathy. *Npj Park. Dis.* **6**, 40 (2020).
67. Dunn, L. et al. Dysregulation of glucose metabolism is an early event in sporadic Parkinson's disease. *Neurobiol. Aging* **35**, 1111–1115 (2014).
68. Dai, C. et al. Glucose metabolism impairment in Parkinson's disease. *Brain Res. Bull.* **199**, 110672 (2023).
69. Ali, M. Z. & Dholaniya, P. S. Oxidative phosphorylation mediated pathogenesis of Parkinson's disease and its implication via Akt signaling. *Neurochem. Int.* **157**, 105344 (2022).
70. Devi, L., Raghavendran, V., Prabhu, B. M., Avadhani, N. G. & Anandatheerthavarada, H. K. Mitochondrial import and accumulation of  $\alpha$ -synuclein impair Complex I in human dopaminergic neuronal cultures and Parkinson disease brain. *J. Biol. Chem.* **283**, 9089–9100 (2008).

71. Xu, W. et al. Reactivation of nonsense-mediated mRNA decay protects against C9orf72 dipeptide-repeat neurotoxicity. *Brain* **142**, 1349–1364 (2019).
72. Zuniga, G. et al. Tau-induced deficits in nonsense-mediated mRNA decay contribute to neurodegeneration. *Alzheimers Dement.* **19**, 405–420 (2023).
73. Ortega, J. A. et al. Nucleocytoplasmic proteomic analysis uncovers eRF1 and nonsense-mediated decay as modifiers of ALS/FTD C9orf72 toxicity. *Neuron* **106**, 90–107.e13 (2020).
74. Ho, W. Y. et al. Dysfunction in nonsense-mediated decay, protein homeostasis, mitochondrial function, and brain connectivity in ALS-FUS mice with cognitive deficits. *Acta Neuropathol. Commun.* **9**, 9 (2021).
75. Hallacli, E. et al. The Parkinson's disease protein alpha-synuclein is a modulator of processing bodies and mRNA stability. *Cell* **185**, 2035–2056.e33 (2022).
76. Brogna, S. & Wen, J. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat. Struct. Mol. Biol.* **16**, 107–113 (2009).
77. Petrić Howe, M. & Patani, R. Nonsense-mediated mRNA decay in neuronal physiology and neurodegeneration. *Trends Neurosci* **46**, 879–892 (2023).
78. Li, J. et al. Integration of transcriptome-wide association study with neuronal dysfunction assays provides functional genomics evidence for Parkinson's disease genes. *Hum. Mol. Genet.* **32**, 685–695 (2023).
79. Onur, T. S. et al. Downregulation of glial genes involved in synaptic function mitigates Huntington's disease pathogenesis. *eLife* **10**, e64564 (2021).
80. Johnson, E. C. B. et al. Large-scale deep multi-layer analysis of Alzheimer's disease brain reveals strong proteomic disease-related changes not observed at the RNA level. *Nat. Neurosci.* **25**, 213–225 (2022).
81. Dammer, E. B., Seyfried, N. T. & Johnson, E. C. B. Batch correction and harmonization of -Omics datasets with a tunable median polish of ratio. *Front. Syst. Biol.* **3**, 1092341 (2023).
82. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
83. Davie, K. et al. A single-cell transcriptome atlas of the aging *Drosophila* brain. *Cell* **174**, 982–998.e20 (2018).
84. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
85. McGinnis, C. S. et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).
86. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
87. Özel, M. N. et al. Neuronal diversity and convergence in a visual system developmental atlas. *Nature* **589**, 88–95 (2021).
88. Wu, T. et al. Tau polarizes an aging transcriptional signature to excitatory neurons and glia. *eLife* **12**, e85251 (2023).
89. Li, H. et al. Fly Cell Atlas: a single-nucleus transcriptomic atlas of the adult fruit fly. *Science* **375**, eabk2432 (2022).
90. Skene, N. G. & Grant, S. G. N. Identification of vulnerable cell types in major brain disorders using single cell transcriptomes and expression weighted cell type enrichment. *Front. Neurosci.* **10**, 16 (2016).

## Acknowledgements

We would like to thank Jennifer Deger, Pinghan Zhao, Tom Lee, and Hui Ye for their help in *Drosophila* brain dissection and Vicky Brandt for their thoughtful comments and edits to the manuscript. This work was funded by the Huffington Foundation, NIH U01 AG072439, and NIH R01 AG057339.

## Author contributions

Conceptualization: J.M., J.B., J.M.S., T.W. Data generation: M.A., O.E.-F., A.L., A.P., B.A., A.T., K.F.Z., T.W., D.D., E.B.D., N.T.S., I.A.-R. Data processing: T.W., J.M. Data analysis and investigation: J.M. Writing: J.M., J.B. Reviewing and editing: J.M., J.B. Funding acquisition: J.B., J.M.S. Supervision: J.B.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41531-025-00899-z>.

**Correspondence** and requests for materials should be addressed to Juan Botas.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025