Integration of molecular regulatory networks: The *Drosophila* insulin response and mouse embryonic stem cell differentiation dynamics.

Inaugural-Dissertation

zur

Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Robert Sehlke
aus Neubrandenburg

Köln, 2019

Berichterstatter (Gutachter):	Prof. Dr. Linda Partridge
	Prof. Dr. Andreas Beyer
Tag der mündlichen Prüfung:	17. Mai 2018

This dedication is self-referential.

Acknowledgements

I want to thank Linda Partridge, whom I first approached about doing graduate work in ageing research. She gave me the opportunity to work on the first and major project presented in this thesis: the intricate and complex insulin response of long-lived fruit flies. My thanks also go to Andreas Beyer, who was my principal supervisor for most of my graduate work. I appreciated his open-door policy, insightful advice, and willingness to let me pursue own ideas; though many of them ultimately failed, they were good failures, and I learned a lot from them.

My sincere gratitude goes to Luke Tain, who worked tirelessly in the lab to generate the data my analysis relied on (he does not, in fact, sleep a lot). Through numerous patient discussions, he taught me a lot about the biology of the system I was working with.

I also like to thank Martin Leeb, with whom I collaborated closely on the second project in this thesis, the early differentiation of mouse embryonic stem cells. His keen sense for detail has made this project both more challenging and more rewarding.

Not less importantly, I want to thank all members of the Beyer lab, who accepted my at times stoic fixation on my projects, for their help, offbeat discussions, onbeat discussions, and everything in between. In particular I want to thank Manopriya Chokkalingam, with whom I worked on the *Drosophila* insulin signalling project from the beginning. I also want to thank the members of the former Dieterich lab, who hosted me for the first year of my graduate work.

Lastly, I want to thank the Cologne Graduate School for Ageing Research, in particular the coordinators Daniela Morick and Jenny Ostermann, for their support both material and immaterial. Many thanks also go to Ruth Willmott for being an excellent coach through many helpful soft skill courses.

1 Summary

Systems biology relies on considering processes on multiple levels of the organism—e.g. tissues, transcripts and proteins, developmental stages, and age—to generate and test hypotheses that are not amenable to a single-gene approach. This is of particular importance in ageing research, which deals with convoluted relationships between multiple disparate aspects of biology and medicine, as summarized recently in "The Hallmarks of Aging" (López-Otín et al. 2013). In this thesis, I present results from two large collaborative efforts that generated, integrated, and analysed systems level data. These projects focus on two of the nine hallmarks of ageing—deregulated nutrient sensing and stem cell exhaustion—, but touch on several others.

I. **The discovery of downstream processes of delayed ageing in *Drosophila* insulin mutants.** This project is further subdivided into four major parts:
    A. The tissue-specific proteome response to reduced insulin in the mNSC-ablation model of reduced insulin signalling (IIS). Furthermore, the dependency of these changes on the transcription factor dFOXO. As published in (Tain et al. 2017).
    B. The comparison of the tissue-specific proteome responses of mNSC-ablated flies and a second model of reduced IIS, $dilp^{2\text{-}3,5}$ mutant flies. Furthermore, the dependency of the $dilp^{2\text{-}3,5}$ response on the presence of the endosymbiont *Wolbachia.*
    C. The role of post-transcriptional regulation in the tissue-specific proteome to reduced IIS response of $dilp^{2\text{-}3,5}$ mutants, through the analysis of matching RNAseq data.
    D. The age-dependent proteome response to reduced IIS in $dilp^{2\text{-}3,5}$ mutants. Presented are preliminary results from two tissues (fat body and thorax).

II. **Understanding mouse embryonic stem cell pluripotency and early differentiation.** The focus of this project is the analysis of RNAseq data from 74 differentiation defect mESC lines under conditions promoting naive pluripotency or early differentiation. We also present supporting analysis of the specific effect of LIF on naive mESCs and insights from a screen of a inhibitor/knockout screen. As the project is ongoing, results presented in this thesis are preliminary.

Additionally, the appendix describes custom R tools that were developed to facilitate the analysis, visualization of results, and collaboration for these projects: CellPlot, SETHRO, and mESCexplorer.

Kurzzusammenfassung

Systembiologie betrachtet Prozesse gleichzeitig auf mehreren Ebenen des Organismus—Gewebe, Transkripte und Proteine, Entwicklungsstadien und Alter—um Hypothesen zu generieren und zu testen, die mit der Betrachtung einzelner Gene nicht adressiert werden können. Dies ist von besonderer Bedeutung in der Altersforschung, welche komplexe Beziehungen zwischen verschiedenen Aspekten von Biologie und Medizin betrachtet, wie kürzlich in "The Hallmarks of Ageing" zusammengefasst (López-Otín et al. 2013). In dieser Arbeit präsentiere ich die Ergebnisse von zwei großen kollaborativen Forschungsprojekten, in denen Daten auf Systemebene generiert, integriert und analysiert wurden. Diese Projekte konzentrieren sich auf zwei der neun Kennzeichen des Alterns—deregulierte Nährstoffsignale und die Erschöpfung von Stammzellen—, berühren jedoch mehrere andere.

I. **Die Identifizierung von Downstream-Effektoren der verzögerten Alterung in Drosophila Insulin-Mutanten.** Dieses Projekt ist weiter in vier Hauptteile unterteilt:
   A. Die gewebespezifische Proteom-Antwort auf verringertes Insulin im mNSC-Ablationsmodell des reduzierten Insulin-Signalweges (IIS). Darüber hinaus, die Abhängigkeit dieser Veränderungen vom Transkriptionsfaktor dFOXO. Ein Großteil dieses Projektteils wurde in (Tain et al. 2017) veröffentlicht.
   B. Der Vergleich der gewebespezifischen Proteom-Antworten von Fliegen mit mNSC-Ablation und einem zweiten Modell mit reduziertem IIS, *dilp*$^{2\text{-}3,5}$ Mutanten. Darüber hinaus wird die Abhängigkeit der *dilp*$^{2\text{-}3,5}$ Insulinantwort vom Endosymbionten *Wolbachia* betrachtet.
   C. Der Anteil der post-transkriptionellen Regulation im gewebespezifischen Proteom an der Verringerung der IIS-Antwort in *dilp*$^{2\text{-}3,5}$-Mutanten, mit Hilfe paralleler RNAseq.
   D. Die altersabhängige Proteom-Antwort auf reduzierten IIS in *dilp*$^{2\text{-}3,5}$-Mutanten. Präsentiert werden vorläufige Ergebnisse für zwei Gewebe (Fettkörper, Thorax).

II. **Förderung des Verständnisses der Pluripotenz von embryonalen Stammzellen der Maus (mESCs), sowie deren anfänglicher Differenzierungsprozess.** Der Fokus dieses Projektes liegt auf der Analyse von RNAseq-Daten von 74 mESC-Linien mit Differenzierungsdefekt unter Bedingungen, die entweder die naive Pluripotenz oder die Differenzierung fördern. Wir präsentieren zusätzlich Analysen der spezifischen Wirkung von LIF auf naive mESCs und Ergebnisse eines kombinierten Inhibitor/Knockout Screens. Dieses Projekt wird noch aktiv entwickelt, daher sind hier präsentierte Ergebnisse als vorläufig zu betrachten.

Darüber hinaus beschreibt der Anhang R-Skripte, die entwickelt wurden, um die Analyse, Visualisierung von Ergebnissen und die Kollaboration in den hier vorgestellten Projekten zu erleichtern: CellPlot, SETHRO und mESCexplorer.

2 Contents

3 Dynamics of fly insulin signalling across omics and age

3.1 Introduction: Insulin signaling in the long-lived fly

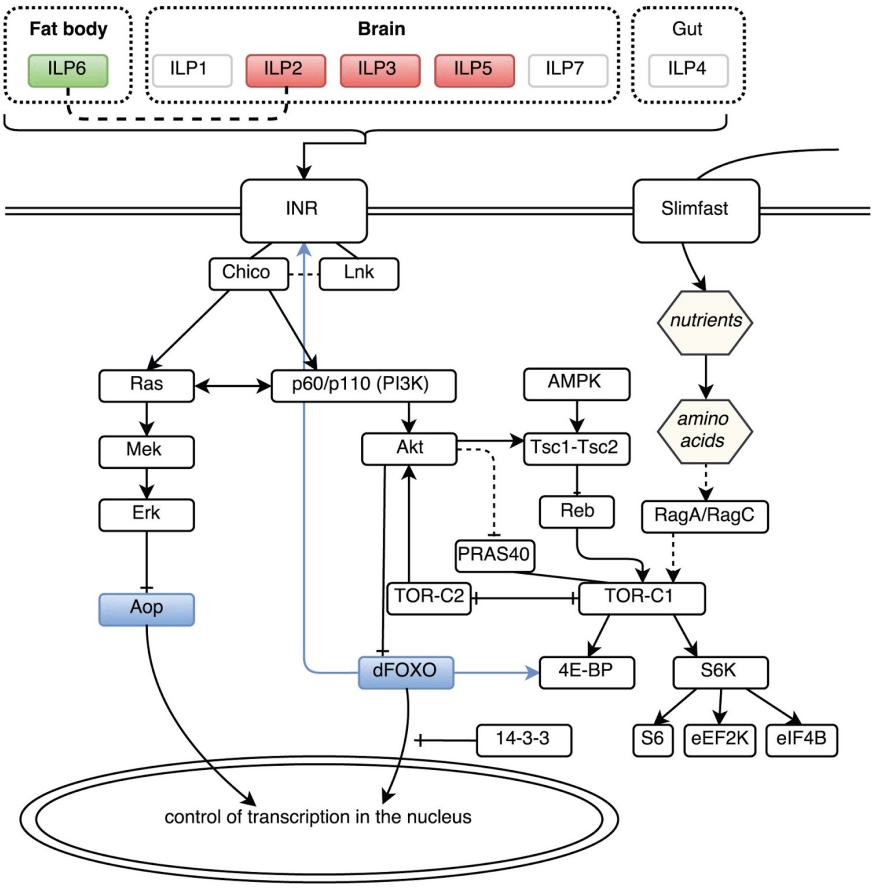


Fig 1: ***Schema of the Drosophila IIS/TOR signalling network.***
Schema showing elements of IIS/TOR signalling and their relationships. Arrows indicate activation. Blunted arrows signify suppression. Blue arrows indicate positive transcriptional regulation. Solid lines without arrows indicate physical association. Transcription factors are marked in blue. DILPs marked in red increase lifespan upon knockout and are targets of this study. DILP6 (green) over-expression in the fat body increases lifespan. DILP1, DILP4, and DILP7 are only expressed during embryonic development. Adapted from (Teleman 2009; Partridge et al. 2011; Slack et al. 2015).

The conserved insulin/IGF-like signalling pathway (IIS) is a central network of nutrient signalling, and consequently affects a variety of organismal phenotypes, including longevity (Partridge et al. 2011; Kenyon 2011). Reducing IIS extends longevity in a variety of model organisms, and may influence the lifespan of humans (Fontana 2007) as well as ameliorating negative effects of ageing and related pathologies (Saltiel & Ronald Kahn 2001; Guo 2014). IIS signalling integrates information on stress, nutrient availability, growth factors and other elements of energy

metabolism to affect a wide range of downstream processes: besides an increase in life expectancy, pleiotropic effects such as increased xenobiotic and oxidative stress resistance, reduced fecundity and growth, and delayed development occur (Bartke 2011). The interplay and downstream effectors of these phenotypes are still not fully understood.

In flies, IIS is activated through binding of insulin-like peptides (DILPs) by the insulin receptor (InR). Four of the DILPS, DILP1-3 and 5, are expressed in median neurosecretory cells (mNSC), a small cluster of cells in the brain of flies that functions analogously to mammalian pancreatic islet beta cells (Rulifson et al. 2002). However, only DILP2, 3, and 5 could be detected in the adult brain (Broughton et al. 2005). Similarly, DILP4 is only expressed in the gut during embryonic development (Brogiolo et al. 2001). Interestingly, DILP6, expressed in the fat body, can repress DILP2 secretion (Bai et al. 2012). Following binding of DILPs with InR, InR is phosphorylated and can bind its substrates Chico, Lnk, and Dreadlocks. Of these, Chico and Lnk are thought to be responsible for the majority of signal transduction though InR, as double mutants of these genes specifically are not viable (Teleman 2009). Downstream of this point, IIS splits into two branches: firstly, the Ras-Erk branch, which controls the activity of the transcriptional repressor Aop. Secondly, the PI3K-Akt branch, which regulates dFOXO, a central transcriptional activator (Jünger et al. 2003; Yamamoto & Tatar 2011). These branches are co-regulated, and disrupting them results in *Drosophila* lifespan extension through either activation of Aop (Slack et al. 2015), dFOXO (Kenyon et al. 1993; Yamamoto & Tatar 2011; Slack et al. 2011), or both (Slack et al. 2015).

Furthermore, IIS/PI3K-Akt interacts with the target of rapamycin (TOR) signalling pathway, which is responsible for sensing amino acid availability (Hara et al. 1998): firstly, Akt can directly phosphorylate Tsc1, alleviating the inhibition of Rheb and thereby increasing TORC1 activity (Manning et al. 2002). Secondly, the TORC1 substrate PRAS40 may be phosphorylated by either TORC1 itself or Akt, which causes it to dissociate and reduce TORC1 activity, thereby forming a negative feedback loop (Wiza et al. 2012). Additionally, Akt may be phosphorylated and activated by TORC2 (Hietakangas & Cohen 2007), which is antagonised by TORC1 (Yang et al. 2006). Downstream effects of insulin signalling are therefore intertwined with TOR signalling in various ways. One of these is the regulation of translation through S6K-eIF4B (Shahbazian et al. 2006). TOR signalling and IIS also converge on eIF4B binding protein (4E-BP), a repressor of translation that is inhibited by TORC1 but transcriptionally activated by dFOXO (Beretta et al. 1996; Miron et al. 2003; Jünger et al. 2003). Several other factors in both TOR signalling and IIS have been shown to affect lifespan (Partridge et al. 2011), making it a challenging task to disentangle downstream effects of interventions that affect the whole pathway.

In particular, interventions that reduce the levels of upstream regulators of IIS and thereby IIS activity can robustly extend lifespan: these include mutation of 3 of the 7 DILPs ($dilp^{2\text{-}3,5}$)(Grönke et al. 2010) and ablation of the mNSCs that produce those dilps (Broughton et al. 2008). Curiously, in $dilp^{2\text{-}3,5}$ deficient flies the presence of the ubiquitous intracellular symbiont

Wolbachia is required for lifespan extension (Grönke et al. 2010). Profiling of the mRNA transcriptome has identified genes and molecular mechanisms that may ameliorate ageing in IIS mutants in *C. elegans* (Murphy et al. 2003; Halaschek-Wiener et al. 2005; Oh et al. 2006; McElwee et al. 2007; Ewald et al. 2015; Kaletsky et al. 2016) and *Drosophila* (Teleman et al. 2008; Alic et al. 2011). Increased detoxification (Amador-Noguez et al. 2007; Selman et al. 2009; Ewald et al. 2015; Afschar et al. 2016) and reduced translation (Afschar et al. 2016; Kaletsky et al. 2016; Selman et al. 2009) are two functional signatures consistently associated with longevity across different model organisms (McElwee et al. 2007). However, in general RNA profiling in *C. elegans* has been carried out on whole worms, potentially leaving tissue-specific mechanisms undetected. Furthermore, a wide range of post-transcriptional mechanisms may modify expression of proteins (Barrett et al. 2012; Batista & Chang 2013; Cech & Steitz 2014a; Liu et al. 2016; MacInnes 2016), with recent reports of the correlation between expression levels in the transcriptome and the proteome ranging from 40 to 84% (Wolkow et al. 2000; Schwanhäusser et al. 2011; J. J. Li et al. 2014).

Tissue-specific modulation of IIS can also extend lifespan, in the worm, Drosophila and the mouse (Wolkow et al. 2000; Blüher et al. 2003; Giannakou et al. 2004; Demontis & Perrimon 2010). Conditional knockout models in mice, and studies in humans, have revealed that the responses of gene expression to reduced IIS are tissue-specific (Nandi et al. 2004; Rask-Madsen & Kahn 2012). Recently, tissue-specific, global regulation of RNA expression in IIS mutants has been undertaken in the Drosophila gut and fat body (Alic et al. 2014) and in neurons in C. elegans (Alic et al. 2014; Kaletsky et al. 2016). However, proteomic profiling of individual tissues in the key invertebrate model organisms Drosophila and C. elegans has been limited, due to their small size, and the requirement for large quantities of starting material. Initial investigations into the proteome of wild-type Drosophila focused on protein identification, using samples of low complexity, and either subcellular or biochemical fractionation to achieve increasing depth (Veraksa et al. 2005; Brunner et al. 2007; Aradska et al. 2015). Recent advances in mass spectrometry, sample preparation techniques and data analysis now allow the quantification of near complete proteomes and proteomic expression profiling (Nagaraj et al. 2012; Mann et al. 2013; Azimifar et al. 2014; M.-S. Kim et al. 2014; Deshmukh et al. 2015). These new developments have reduced the amount of starting material required for shotgun proteomics and created an opportunity to investigate how attenuated IIS affects the proteome of individual fly tissues and hence to identify candidate mechanisms that ameliorate the effects of ageing.

3.2 Project aims and contributions

Insulin signalling is a central element of metabolism conserved across a wide range of model organisms (Fontana et al. 2010). Concrete evidence that human lifespan and healthspan can be extended by either caloric restriction, dietary restriction or pharmacological interventions acting on IIS signalling, has emerged in recent years (Fontana et al. 2010; Barzilai et al. 2016; Lee & Longo 2016). Therefore, understanding the precise mechanisms of IIS that mediate these beneficial effects will be instrumental in extending human healthspan, while minimising negative side-effects. However, due to the central role of IIS in the regulation of metabolism, its downstream effects are highly convoluted. In particular, the understanding of the interplay of different tissues within the organism, as well as different levels of gene expression regulation in response to reduced IIS is still incomplete.

This project aims to advance the understanding of the system-wide insulin signalling response, integrating experimental data obtained by collaborators from two *Drosophila* model systems: mNSC-ablated flies and $dilp^{2\text{-}3,5}$ mutants. Importantly, it dissects the response to reduced insulin signalling both tissue-specifically and on different levels of regulation (the proteome, the transcriptome, and the translatome). To this end, this project makes use of computational methods to identify, classify, and integrate tissue-specific mechanisms of insulin signalling.

Contributions

Parts of this chapter, especially 3.3.2, have been published previously (Tain et al. 2017), including parts of the discussion.

I performed all bioinformatics analysis (with the exception of network propagation analysis, which was done by Manopriya Chokkalingam) and generated visualizations. In the course of the analysis I developed tools for visualization and functional analysis (CellPlot, SETHRO)(appendix). Work on CellPlot was facilitated by Sven Templer, who developed auxiliary functions and wrote documentation. I performed evaluation, interpretation, and discussion of analysis results, in close collaboration with Luke Tain, Manopriya Chokkalingam, and Andreas Beyer.

Experimental work leading up to and including proteomics and transcriptomics screens in the mNSC-ablation and $dilp^{2\text{-}3,5}$ systems was performed chiefly by Luke Tain, with contributions by Chirag Jain, Paul Essers, Sebastian Groenke, Jenny Froehlich, Nagara Nagarajuna, and Mark Rassner. Isolation of miRNA, its quantification and initial analysis was performed by Jennifer Werner and Corinna Klein (Werner 2014). Experimental work to obtain samples for proteomics time series analysis was coordinated and performed by Ralf Meilenbrock, with contributions by Sebastian Groenke, Luke Tain, Carina Weigelt, Yu Xuan.

3.3 Results

3.3.1 Experimental design overview

This section presents a brief summary of the data sources used in this project. Individual subsections contain introductions providing more detail and context.

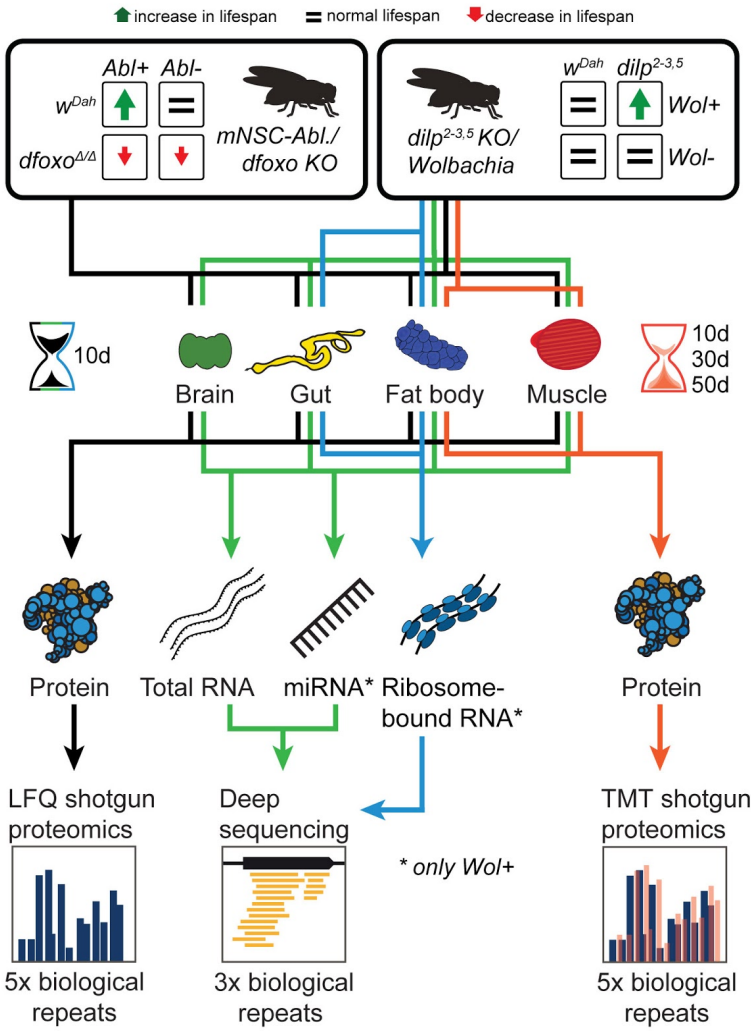


Fig 2*: **Overview: Experimental design of the project.***
Flowchart showing an overview over all fly model systems, their experimental factors and resulting groups, tissues, time points, and gene expression quantification levels, as analysed in the first project presented in this thesis.

To decipher the downstream mechanisms of IIS mediated lifespan extension, two *Drosophila* models of reduced IIS were used. Firstly, mNSC-ablated flies and their wild type controls (Broughton et al. 2005), with additional controls through mNSC-ablated and non-ablated flies in

a *dfoxo*-null background (Fig 2). Secondly, *dilp*$^{2\text{-}3,5}$ knockout mutants and *w*^{Dah} controls, both in the presence and absence of the endosymbiont *Wolbachia* (Grönke et al. 2010)(Fig 2). We quantified the proteomes of 10d old flies of both model systems and all eight experimental groups using LFQ shotgun proteomics (Fig 2). Additionally, transcripts from total RNA, ribosome-bound RNA and miRNA enriched samples were quantified in samples from 10d-old *dilp*$^{2\text{-}3,5}$ and control flies (Fig 2). Finally, we quantified protein content from samples of fat bodies and muscle of 10d, 30d, and 50d old flies using multiplexed TMT shotgun proteomics (Fig 2).

Section 3.3.2 describes tissue-specific effects of insulin signalling in the proteome of long-lived mNSC-ablated flies and identified *dfoxo*-dependent and -independent elements of that response. Large parts of this analysis have been published (Tain et al. 2017).

Section 3.3.3 introduces the long-lived *dilp*$^{2\text{-}3,5}$ model of reduced insulin signalling, comparing its proteome response to reduced IIS to that of mNSC-ablated flies. Additionally, we identify which parts of the response depend in the endosymbiont *Wolbachia*, which is required for lifespan extension in the *dilp*$^{2\text{-}3,5}$ background.

Section 3.3.4 identifies post-transcriptional effects of insulin signalling in the *dilp*$^{2\text{-}3,5}$ model system by integrating our transcriptome quantification with our analysis of the *dilp*$^{2\text{-}3,5}$ proteome.

Section 3.3.5 outlines preliminary findings of age-dependent differences in the proteomes of long-lived *dilp*$^{2\text{-}3,5}$ flies and their controls.

3.3.2 A tissue-specific proteomics atlas of insulin signalling

3.3.2.1 Section introduction and overview

To characterize changes in the tissue-specific proteomes of IIS mutant *Drosophila* and their responses to removal of the key dFOXO transcription factor, we profiled four key insulin-sensitive tissues: the brain, gut, fat body and muscle. The majority of the responses to a systemic reduction in IIS were highly tissue-specific, and 60% of them were not detected in previous transcriptional studies. Proteins associated with the ribosome and the mitochondrial electron transport chain were differentially expressed in the fat body, and led to a reduction in translation and increased mitochondrial respiration. The gut showed a proteomic signature of proteasome–ubiquitin-mediated catabolism, and was associated with elevated proteasomal assembly and activity. Increased respiration and proteasome activity in the fat body and gut, respectively, were required for the increased longevity of IIS mutants. Importantly, manipulation of mitochondrial biogenesis in the fat body and of proteasomal activity in the gut could each extend lifespan in wild-type flies. Our tissue-specific, proteomic analysis has thus revealed how individual tissues of an organism can act in concert, through diverse tissue-specific responses, to ameliorate the ageing of the whole organism in response to reduced IIS.

3.3.2.2 The tissue-specific proteome of wild-type flies

We first characterized the tissue-specific proteomes of w^{Dah} flies using label-free quantification in quintuplicate (Fig 3)(Ilian Atanassov), in four key insulin responsive tissues: brain, fat body, gut, and thorax. Missing values were then imputed by fitting a normal distribution to the left tail of the LFQ value distribution (Deeb et al. 2012)(Ilian Atanassov), to account for the left-censoring common in shotgun proteomics data (Karpievitch et al. 2012). We could quantify 6085 proteins, accounting for 44% of the predicted *Drosophila* proteome (Fig 3B). The four tissues shared a common core of 1916 proteins, while 6-26% were unique to each tissue (Fig 3C). To determine the tissue-specificity of our measurements, we performed principal component analysis (PCA) of all biological replicates. We found that replicates of each tissue cluster closely together, indicating that intra-tissue variability was much lower compared to inter-tissue-variability (Fig 3D). We then selected the top 5% of proteins that were most influential for placing replicates on each end of the first two principal components according to their PCA loading scores. A major determinant for separation along the first principal component (PC) were proteins associated to neuronal GO terms like *synaptic vesicle* and *terminal bouton*, separating brain samples from the other tissues. The second PC highlighted differences in transcriptional regulation, respiration, development and metabolism (Fig 3D).

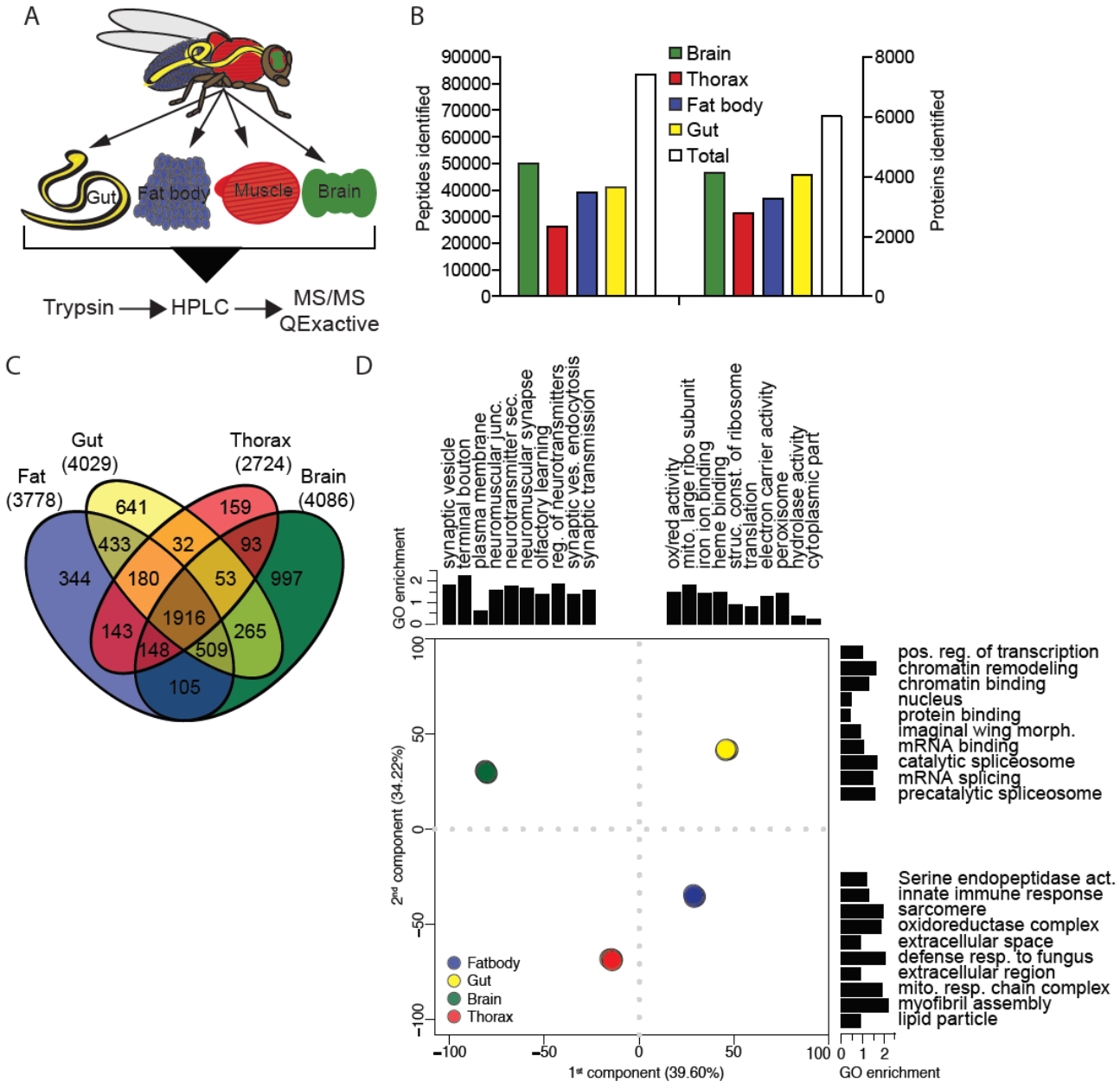


Fig 3***: Drosophila tissues contain common and tissue-specific proteomes.***

(A) *Schematic of experimental design.* ***(B)*** *Tissue-specific and total number of detected peptides and proteins.* ***(C)*** *Venn diagram of proteins detected in wild-type fly tissues.* ***(D)*** *Principal component analysis of the four wild-type tissues GO terms enriched in the 5% tails of contribution along each dimension. Replicates show little variation and overlay each other.*

3.3.2.3 Reduced IIS remodels tissue-specific proteomes

Ablation of median neurosecretory cells late in development is known to abrogate the expression of dilps 2, 3, and 5, which results in lifespan extension (Broughton et al. 2010). To determine the effect of reduced insulin signalling on tissue-specific proteomes, we carried out differential expression analysis, comparing the wild type (w^{Dah}) phenotype to that of mNSC ablated flies. For this, we first fitted a linear model to all expression data, including $dfoxo^{\Delta/\Delta}$ controls, explaining the expression vector *E* of each protein at each sample *i* in terms of the genotype (Eq 1)(error terms omitted). Here, each genotype is represented by a feature vector of the same length as *E* that equals 1 if sample *i* is of that genotype, and 0 otherwise. The genotypes are *W* (w^{Dah}), *WA* (w^{Dah}*/Ablation*), *D* ($dfoxo^{\Delta/\Delta}$) and *DA* ($dfoxo^{\Delta/\Delta}$*/Ablation*). As no intersect was fitted, each feature coefficient β takes on the mean of protein expression across samples of the respective genotype.

$$E_i = \beta_1 W_i + \beta_2 WA_i + \beta_3 D_i + \beta_4 DA_i \qquad (1)$$

In the statistical programming environment R, the more compact Wilkinson-Rogers notation is used to specify linear models (Wilkinson & Rogers 1973). The model used here (Eq 1) can then be written in the following form (Eq 2). In this case, *Expression* is the protein expression vector to be modeled, 0 indicates that no intercept is fitted, and *Genotype* is a multi-level factor that incorporates all four experimental groups.

$$Expression = 0 + Genotype \qquad (2)$$

Using the limma framework for R, we then used the fitted model to calculate statistics for individual linear combinations of experimental groups, called contrasts, in a flexible manner. A contrast sums up to 0 by design and is assumed to be normally distributed, i.e. the significance of differences encoded by the contrast can be determined by testing against the null hypothesis of no change via a t-test. To compare expression of genes between w^{Dah} and w^{Dah}*/Abl* conditions, the necessary contrast takes on the following simple form (Eq 3):

$$\boldsymbol{w^{Dah}}\textbf{/Abl } \textit{\textbf{vs.}} \textbf{ } \boldsymbol{w^{Dah}} = w^{Dah}/\text{Abl} - w^{Dah} \qquad (3)$$

We could then make use of limma's implementation of empirical Bayes statistics for differential expression analysis, which exploits information across all protein expression vectors to obtain a more robust estimate of each protein's variance (Smyth 2004). This is achieved by shrinking sample towards to a common value and calculating moderated t-statistics with increased power (i.e. degrees of freedom)(Smyth 2004). Limma then calculates p-values from these statistics

against the null hypothesis of no change between groups, and corrects for multiple hypothesis testing using the Benjamini-Hochberg (BH) method (Benjamini & Hochberg 1995). While considerable co-dependence between protein expression profiles is expected, the BH method is known to be robust against violations of its independence assumption (Benjamini & Yekutieli 2001).

Thus, at a significance level of $\alpha = 0.1$, we identified 2372 proteins that were differentially regulated in response to reduced IIS (Fig 4), 982 of which showed absolute fold changes larger than two in at least one tissue. 60% of these changes were not previously identified at the RNA transcript level in whole IIS mutant flies (McElwee et al. 2007; Alic et al. 2011). The brain showed the strongest response with 1312 differentially regulated proteins, followed by the gut and fat. Only two proteins, *betaTub60D* and *l(1)G0230*, were regulated across all tissues. However, 87% of proteins that were regulated by IIS in a tissue-specific manner were detected across more than one tissue, indicating that the specificity of the response is not merely a byproduct of tissue-specific protein quantification.

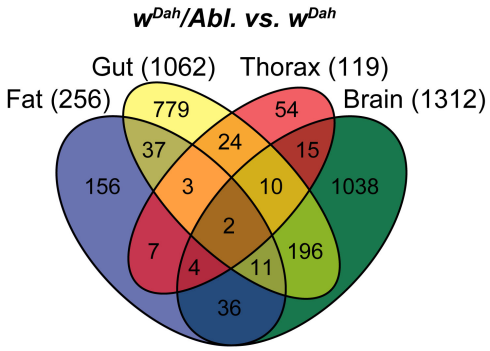


Fig 4***: Proteome changes induced by mNSC ablation.***
Venn diagram of differentially expressed proteins upon reduced IIS, comparing tissues from control (w^{Dah}) and mNSC-ablated flies. Total number of differentially expressed, tissue-specific proteins is shown in parentheses.

We next performed GO enrichment analysis using the topGO R package (Alexa & Rahnenführer 2009) to identify general tissue-specific responses to reduced insulin signalling. TopGO uses a classical enrichment testing approach, using Fisher tests to determine if the proportion of annotated proteins from a GO term is higher in the protein set of interest (significantly changed proteins) than in the background of all other proteins. Additionally, an elimination procedure is used to promote the most specific enriched terms and down-weight general, less informative annotations (Alexa et al. 2006). Importantly, the choice of background protein or gene set has a strong influence on the outcome of the analysis (Tipney & Hunter 2010). To avoid biases arising from sample preparation and detection bias in our proteomics screen, as well as tissue-specific

biases, we therefore used tissue-specific backgrounds (i.e. all detected proteins in that tissue) in our enrichment analyses.

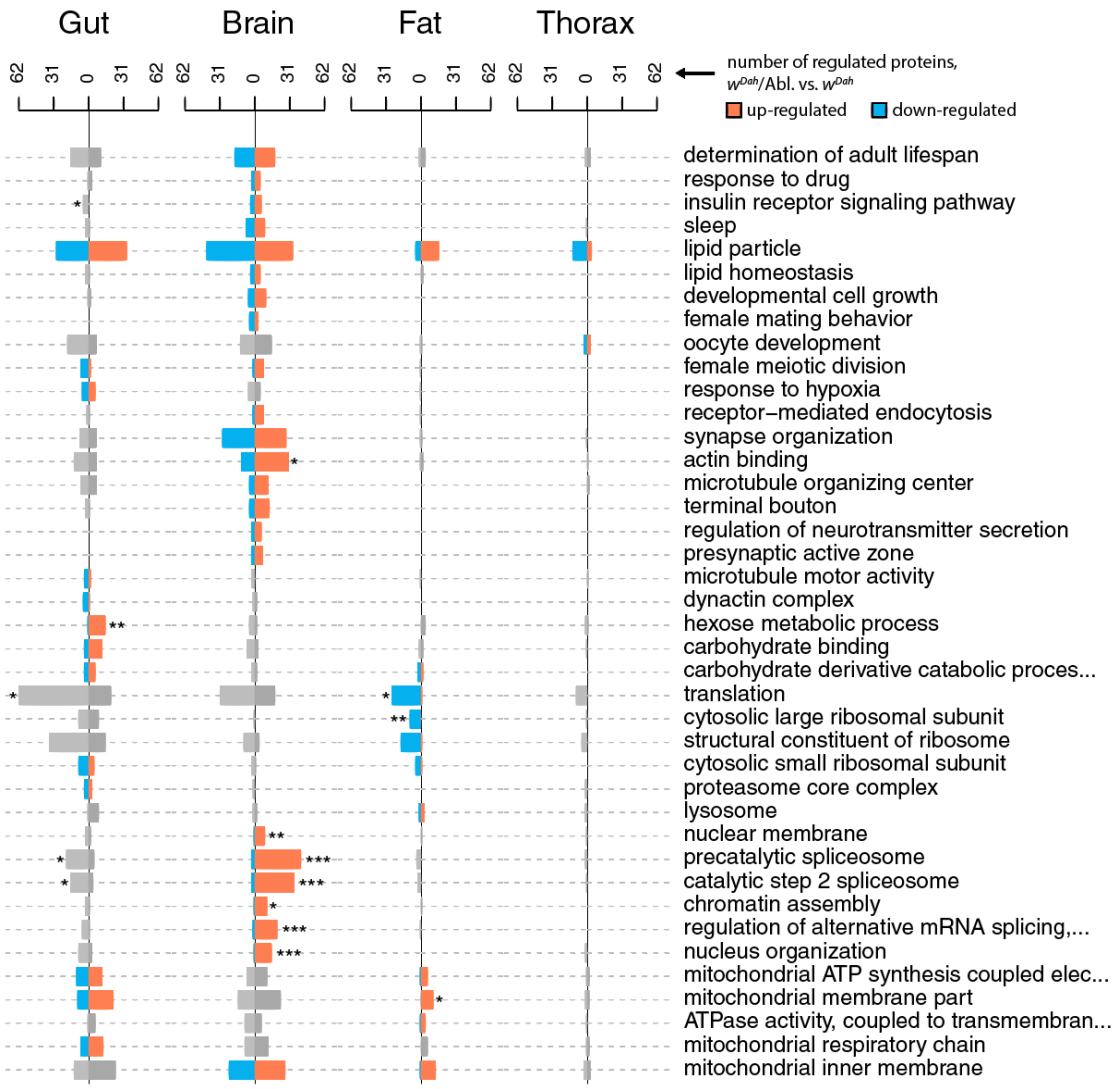


Fig 5: ***Tissue-specific functions of proteins changed in mNSC-ablated flies.*** *Curated list of functional gene ontology terms that were significantly enriched in the differentially expressed protein set of at least one tissue. Asterisks indicate significance level of a one-sided t-test performed on log-fold-changes, indicating coordinated up- (right) or down-regulation (left).*

Confirming the success of our experimental setup, proteins belonging to the insulin receptor pathway in the brain were significantly enriched in the insulin response (Fig 5). Further brain-specific functional enrichments included an up-regulation of splicing-associated proteins, as well as sleep (Fig 5). Alternative splicing contributes to shaping the fly neuronal landscape (Mohr & Hartmann 2014) and is a key part in circadian rhythm regulation (Hughes et al. 2012). It is also known that lowered IIS activity improves sleep quality in *Drosophila* (Metaxakis et al. 2014). In the gut, several functions of carbohydrate and energy metabolism were enriched, along with a small but significant proteostasis response (Fig 5). In the fat body, proteins with

functions in mRNA translation, in particular ribosomal proteins, were enriched and down-regulated, while mitochondrial components were up-regulated (Fig 5). To narrow down on functions that are most likely to be related to the lifespan extension phenotype, we decided to further focus our analysis on identifying responses that were conditional on the presence of the *dfoxo* transcription factor.

3.3.2.4 Identifying tissue-specific, differential protein expression potentially causal in extension of lifespan

dFOXO is a key transcription factor that is required for lifespan extension and increased xenobiotic resistance through reduced IIS (Slack et al. 2011; Yamamoto & Tatar 2011). However, it is not required for other pleiotropic phenotypes of such a reduction (Slack et al. 2011). We could therefore leverage dependence of the IIS response on dFOXO to identify proteins more likely to be causal for the longevity phenotype. To identify proteins whose expression was significantly affected by the interaction between ablation of mNSCs and *dfoxo* deletion, we fit an interaction contrast to our linear model (Eq 1). We considered two types of interaction contrast: the first type tests for a difference in the expression of the long-lived condition, w^{Dah}/Abl, compared to the mean of all controls (Eq 4).

$$Interaction_{all} = w^{Dah}/\text{Abl} - (w^{Dah} + dfoxo^{\Delta/\Delta}/\text{Abl} + dfoxo^{\Delta/\Delta})/3 \qquad (4)$$

This form is reasonable if little difference is expected to occur between the individual controls. However, we found that in our analysis a strong difference to just two out of the three controls could result in detecting a significant difference overall. Moreover, this form does not allow for shifts in base level expression that might result from the strong $dfoxo^{\Delta/\Delta}$ phenotype. E.g., an insulin responsive protein might still be relevant to the longevity phenotype if its absolute expression level is similar to that in $dfoxo^{\Delta/\Delta}$ controls, but the response to mNSC ablation is abrogated in the *dfoxo*-negative background. Therefore, we decided to use a second type of interaction term which quantifies the difference of differences between experimental groups (Eq 5). Note that this form is equivalent to including the interaction term *Ablation:Foxo* in the model itself, where *Ablation* is a factor with two levels (*ablated* and *non-ablated*) and *Foxo* is a factor with two levels (*dfoxo-knockout* and w^{Dah}).

$$Interaction_{diff} = (w^{Dah}/\text{Abl} - w^{Dah}) - (dfoxo^{\Delta/\Delta}/\text{Abl} - dfoxo^{\Delta/\Delta}) \qquad (5)$$

The $Interaction_{diff}$ term can identify two classes of interaction: one, changes that occur in the *dfoxo*-null background upon mNSC ablation, but are not present in the wild type. Two, changes that are present upon mNSC ablation in the wild type, but abrogated in the *dfoxo*-null background. Because of the central role of *dfoxo* in lifespan extension through many IIS interventions (Slack et al. 2011; Yamamoto & Tatar 2011), we reasoned that changes in the second category were more likely relevant for the longevity phenotype of mNSC ablated flies. Thus, we pre-selected proteins that showed a significant insulin response in the wild type, as

identified earlier. We then determined their overlap with proteins that were significantly differentially expressed in the interaction contrast (BH-corrected, $\alpha = 0.1$). We also excluded from the resulting set seven proteins whose response to reduced IIS was exaggerated, as opposed to abrogated, in the *dfoxo*-null background—i.e. if their change upon mNSC ablation was in the same direction and greater in the *dfoxo*-null background compared to the wild type background. We thus identified 361 proteins that required *dfoxo* for their change in expression in IIS mutant flies (Fig 6A). Most of these were expressed in the brain, but in relative terms, a high proportion of the insulin response in the fat body was *dfoxo*-dependent (Fig 6B). To assess whether those 361 proteins were more likely to be direct or indirect targets of *dfoxo*, we searched for predicted *dfoxo*-binding motifs within 1 kb of their transcriptional start sites using MEME (Bailey et al. 2009). 45% of the 361 proteins came from genes with *dfoxo*-binding motifs and may therefore be directly regulated by dFOXO (Tain et al. 2017). However, the remaining 55% of those 361 proteins came from genes lacking *dfoxo*-binding motifs which suggest that although their expression was *dfoxo*-dependent, they were not directly regulated transcriptionally by dFOXO (Tain et al. 2017).

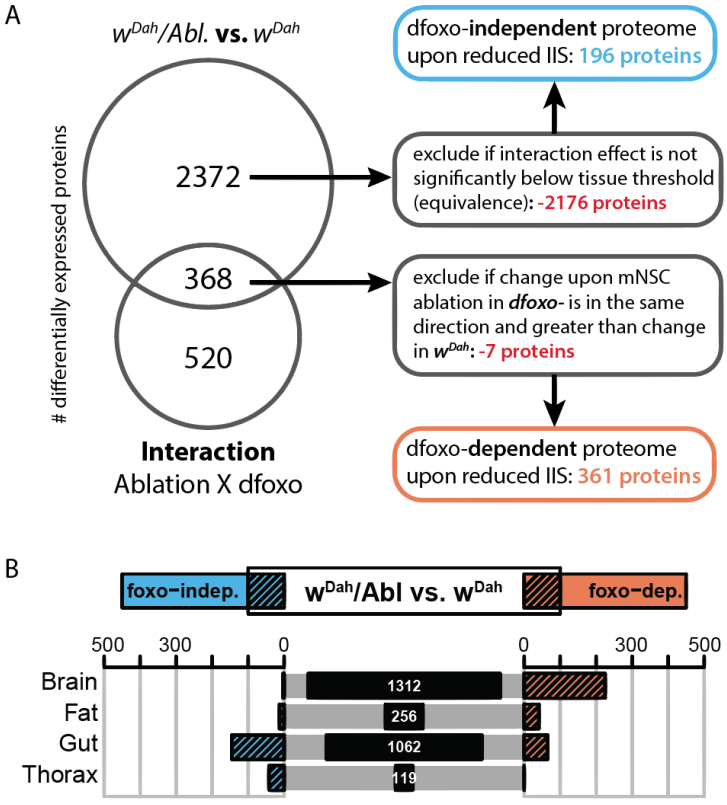


Fig 6: **dfoxo-dependence of insulin-responsive proteins.**

(A) *Genotypes defined as w^{Dah}, $dfoxo^{\Delta/\Delta}$ (dfoxo Δ94 mutant), w^{Dah}/Abl (InsP3-Gal4/UAS-rpr) and $dfoxo^{\Delta/\Delta}$/Abl (InsP3-Gal4/UAS-rpr, dfoxo Δ94). Reduced IIS-mediated changes in the wild-type background (w^{Dah}/Abl vs. w^{Dah}) and in the $dfoxo^{\Delta/\Delta}$ background ($dfoxo^{\Delta/\Delta}$/Abl vs. $dfoxo^{\Delta/\Delta}$) or their interaction term (w^{Dah}/Abl - w^{Dah}) vs. ($dfoxo^{\Delta/\Delta}$/Abl - $dfoxo^{\Delta/\Delta}$). dfoxo-dependent proteins were defined as proteins regulated between w^{Dah}/Abl and w^{Dah} and within the interaction term, but not if the response in the dfoxo- background was the same direction but amplified compared*

to the w^{Dah} background. Proteins were identified as dfoxo-independent if they were regulated between w^{Dah}/Abl and w^{Dah}, and were not subject to interaction effects (equivalence given a tissue-specific threshold). Also included were proteins if they were regulated in the same direction in both backgrounds, but amplified in the same direction in with a stronger same directional response in $dfoxo^{\Delta/\Delta}$/Abl vs. $dfoxo^{\Delta/\Delta}$. ***(B)*** *Symmetric barplots (symplots) showing the number of dfoxo-dependent (orange) and dfoxo-independent (blue) proteins in each tissue. The middle column shows the total number of proteins regulated between w^{Dah}/Abl and w^{Dah} .*

We were also interested to identify responses to reduced IIS that were independent of *dfoxo*. Such a response may be involved in pleiotropic, non-lifespan related parts of the reduced IIS phenotype (Fontana et al. 2010; Kenyon 2010), or mediate lifespan extension through a different branch of IIS (Slack et al. 2015). For this, we needed to identify proteins that responded to IIS but had an interaction magnitude (Eq 5) close to zero. A common ad-hoc practice in differential expression analysis is to assert the absence of a difference when failing to reject the null hypothesis of no change. However, this carries a high risk of type II errors (false negatives), as statistical tests for differential expression usually only control type I errors (false positives). Therefore, it was necessary to first define a threshold of interaction magnitude below which we consider no biologically meaningful interaction to be present. For this, we determined the lowest absolute interaction change detected as significant in each tissue, *t*, and used it to define an equivalence interval around zero ([-t, t]). Using a two one sided t-test procedure (TOST) (Westlake 1976), we then determined whether the 95% confidence interval of the interaction coefficient of each protein fell within this interval. Thus, we guaranteed that the proteins identified as non-interacting in each tissue were significantly distinct from any proteins identified as *dfoxo*-dependently regulated. In this way we identified 196 *dfoxo*-independently regulated proteins among the IIS responsive set (Fig 6A), most of which were regulated in the gut (Fig 6B).

To further increase the power of our analysis of *dfoxo* dependency, we employed network propagation (Manopriya Chokkalingam), which combines differential expression with physical interactors of individual proteins to identify altered subnetworks in each tissue (Vanunu et al. 2010). Differentially expressed *dfoxo*-dependent or *dfoxo*-independent proteins were mapped independently onto the *Drosophila* protein–protein interaction network (Murali et al. 2011), and the associated *P*-value of each protein was negative logarithmic transformed and propagated to adjacent interacting proteins. We then clustered the *dfoxo*-dependent and *dfoxo*-independent responses to reduced IIS in each tissue and identified functional categories of these clusters by GO enrichment analysis using topGO (Tain et al. 2017)(Fig 7). We focused on clusters whose regulation was predicted as *dfoxo*-dependent within the network propagation analysis, since these should include processes causally linked to longevity. *dfoxo*-dependent responses in the brain were enriched for mitochondrial electron transport chain, mRNA splicing and nucleosome components , while in the gut they were enriched for proteasome and ubiquitin-mediated protein catabolism (Fig 7). Mitochondrial electron transport chain was also enriched within *dfoxo*-dependent responses in the fat body, as were ribosomal constituents and

proteins involved in nucleosome assembly. Among the functional groups identified by network propagation, we further focused on three strong candidates for mediating the longevity of IIS mutant flies. First, we chose the ribosome, which we identified as an almost uniquely fat body-specific response. Second, we focused on the mitochondrial electron transport chain, the only response identified as common to multiple tissues, the brain, gut and fat body. Finally, we examined proteasome–ubiquitin protein catabolism as the most enriched cluster in the gut.

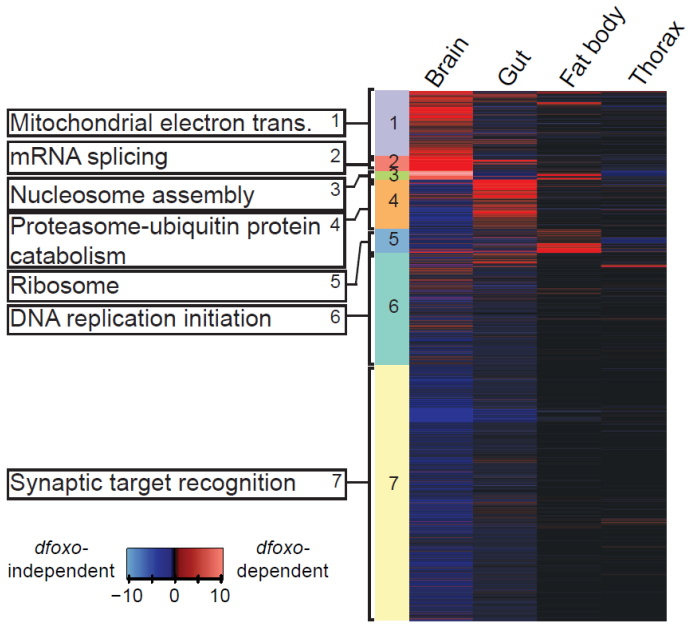


Fig 7***: Hierarchical clustering and GO enrichment analysis of dFOXO-dependent and dFOXO-independent IIS-mediated regulation of the proteome***

Tissue-specific heatmap of significantly regulated dfoxo-dependent (red) and dfoxo-independent (blue) proteins in response to reduced IIS. Coloured side bars represent network propagated score clustering and associated most significantly enriched GO terms.

3.3.2.5 Reduced IIS alters fat body-specific, dfoxo-dependent protein expression to regulate translation

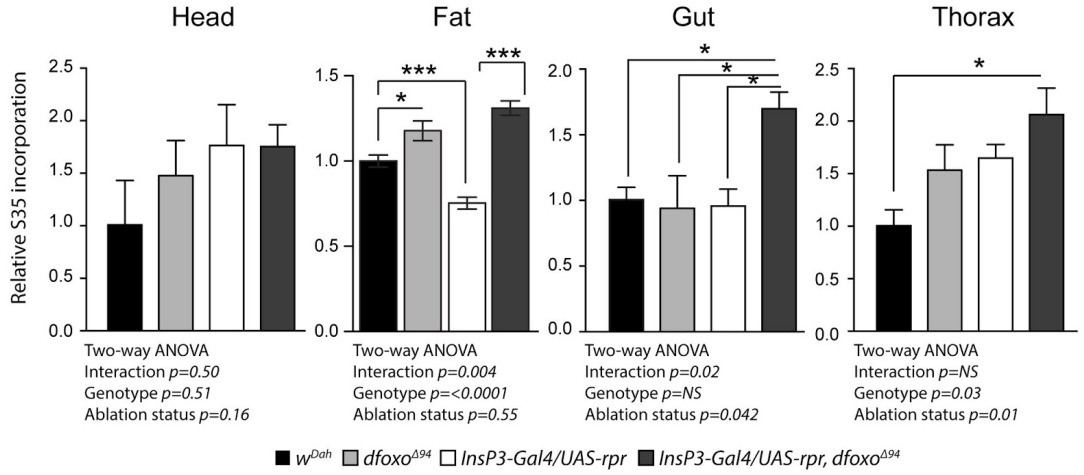


Fig 8**: Translational activity in the fat body under reduced IIS is reduced in a dfoxo-dependent manner**

*De novo protein synthesis as measured by incorporation of 35S into proteins from ex vivo fat body, head, thorax and gut tissue. Equal levels of protein were loaded per lane. Representative gel exposures show fat body 35S incorporation from w^{Dah}, $dfoxo^{\Delta 94}$, InsP3-Gal4/UAS-rpr and InsP3-Gal4/UAS-rpr, $dfoxo^{\Delta 94}$ flies. Quantification of gel exposures normalized to total protein and corresponding two-way ANOVA, as well as Bonferroni corrected post hoc tests are shown alongside each gel exposure. Error bars show SEM, n = 11 for w Dah and InsP3-Gal4/UAS-rpr fat body samples, for all other genotypes and tissues n = 5, *P < 0.05, ***P < 0.001. Experimental results and statistics by Luke Tain. Figure adapted from our paper, see (Tain et al. 2017) for the full figure.*

We next determined if the tissue-specific, dfoxo-dependent proteomic responses led to tissue-specific functional changes. Our bioinformatics analysis identified a fat body-specific, dfoxo-dependent reduction in ribosome-associated proteins in response to reduced IIS (Fig 7). To determine whether this signature was reflected in altered physiology, we quantified tissue-specific translation, using ^{35}S incorporation to detect de novo protein synthesis. Individual tissues were dissected and incubated in medium containing ^{35}S-labelled methionine and cysteine. We separated samples by SDS–PAGE and quantified de novo labelled proteins (Tain et al. 2017). No significant change was detected in head, thorax or gut (Fig 8), but there was significantly less incorporation of labelled amino acids into proteins in the fat body of the IIS mutants, and that reduction was dependent on dfoxo (two-way ANOVA $P < 0.05$, Fig 8)(Tain et al. 2017). This finding suggests that IIS-mediated, *dfoxo*-dependent reduction in translation in the fat body may contribute to longevity.

3.3.2.6 Reduced IIS tissue-specifically and dfoxo-dependently regulates respiration

Changes in mitochondrial function have been linked to ageing in several model organisms (Bratic & Larsson 2013). Paradoxically, however, both reduced mitochondrial respiratory chain protein levels (Copeland et al. 2009) and increased mitochondrial biogenesis (Rera et al. 2011) can extend lifespan. Our analysis highlighted a tissue-specific regulation of mitochondrial proteins that was *dfoxo*-dependent in the brain and fat body and independent of *dfoxo* in the gut (Fig 7). In total, we detected 60 proteins belonging to the annotation mitochondrial electron transport chain, with 6–35% of them in different tissues showing significantly altered levels upon IIS reduction (Fig 9). They were coordinately up-regulated in the fat body, with no clear coordinated change in other organs (Tain et al. 2017). Using Western blotting, we directly confirmed that NDUFS3, a complex I subunit, was up-regulated in a *dfoxo*-dependent manner in the fat body (Tain et al. 2017).

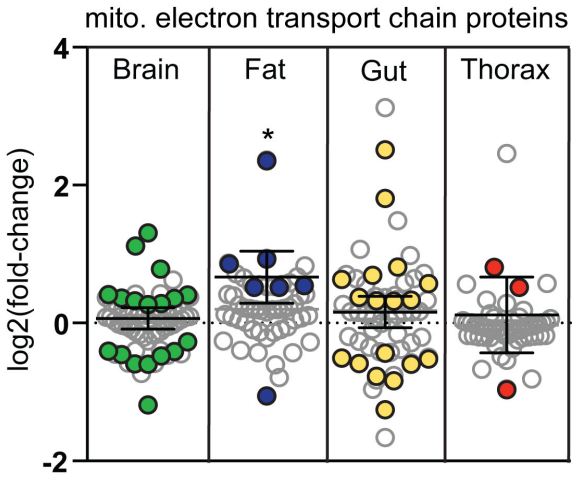


Fig 9*: **Reduced IIS differentially regulates tissue-specific mitochondrial function and biogenesis.***
Log2 fold changes of IIS-regulated proteins associated with the functional term mitochondrial electron transport chain. One-sample t-test shows directional significance (p-value ≤ 0.05). Grey circles show all detected proteins, and coloured circles show significantly regulated proteins. See (Tain et al. 2017) for the full figure with additional experimental results.*

To determine the functional significance of the observed changes in protein levels, we provided substrates for the individual complexes of the electron transport chain to measure respiration in actual respiratory state (PGMP3), the potential maximum respiratory state (Vmax) and role of complex I (rotenone sensitive). In agreement with our bioinformatic analysis, endogenous respiration (PGMP3) was up-regulated by reduced IIS in the fat body, a response that was

completely dependent on *dfoxo* (two-way ANOVA interaction term $P < 0.01$)(Tain et al. 2017). This was also the case for the Vmax and rotenone-sensitive respiratory states, suggesting that the increased respiration was due to increased capacity of the respiratory chain, at least in part through complex I (Tain et al. 2017). In contrast, endogenous (PGMP3) and Vmax respiratory states in the gut were reduced in mNSC-ablated flies and, as predicted by our analysis of the proteome, this occurred independently of *dfoxo* (Tain et al. 2017). Respiration in the head and thorax was unchanged by reduced IIS (Tain et al. 2017). We confirmed the increase in endogenous respiration of the fat body, and the reduced endogenous respiration of the gut, in an independent model of reduced IIS, long-lived *dilp2-3,*5 mutant flies (Grönke et al. 2010; Tain et al. 2017).

To further understand the mechanistic basis for increased respiration in the fat bodies of mNSC-ablated flies, we examined mitochondrial biogenesis. We first examined mitochondrial DNA (mtDNA) levels in the fat body, in the presence and absence of *dfoxo,* and found that they were increased in the MNC-ablated flies in a *dfoxo*-dependent manner (two-way ANOVA interaction term $P = < 0.01$)(Tain et al. 2017). Mitochondrial biogenesis is driven by nuclear transcription factors, including nuclear respiratory factors 1 and 2 (nrf)(Evans & Scarpulla 1990; Virbasius et al. 1993), which are co-activated by PGC-1 to regulate the expression of mitochondrial proteins (Puigserver & Spiegelman 2003; Scarpulla 2008; Tiefenböck et al. 2010). *Drosophila* has single homologs of PGC-1 and nrf-2 (Gershman et al. 2007), named *spargel* and *delg,* respectively (Tiefenböck et al. 2010), and loss of either reduces transcription of nuclear encoded mitochondrial genes (Tiefenböck et al. 2010). *Spargel* and *delg* are both expressed at low levels in adult *Drosophila* fat body (Graveley et al. 2011), and we did not detect them, or their regulation in our analysis (Tain et al. 2017). However, as with many TFs and cofactors, the activity of Spargel and delg could be regulated on the post-translational level (Li et al. 2007). We tested whether *spargel* or *delg* were required for the increased respiration in the fat bodies of IIS mutant flies. To achieve fat body-specific knockdown of *spargel* and *delg,* we performed this analysis in $dilp^{2\text{-}3,5}$ mutants. We found that increased respiration in the fat body of the IIS mutant was entirely dependent on *Spargel* and *delg* expression in the fat body (two-way ANOVA interaction term $P = < 0.001$)(Tain et al. 2017). Furthermore, knockdown of expression of *Spargel* or *delg* in the fat body of $dilp^{2\text{-}3,5}$ mutants reduced the extent of their increased longevity, but did not affect the lifespan of wild-type controls (Tain et al. 2017). Experimentally increased expression of *Spargel* in the adult fat body, using the GeneSwitch driver S106-GS, was sufficient to increase respiration and lifespan, while over-expression in the gut alone was not (Tain et al. 2017).

Overall, our data reveal that lowered systemic IIS coopts different mechanisms to regulate respiration in different tissues. Reduced IIS decreased respiration in the gut independently of *dfoxo*, whilst simultaneously increasing respiration in the fat body in a *dfoxo- and*

spargel/delg-dependent manner. Furthermore, increased mitochondrial biogenesis, and thus respiration, in the fat body is both necessary and sufficient to extend lifespan.

3.3.2.7 Reduced IIS alters the dfoxo-dependent gut proteome, increasing proteostasis to maintain gut health

The proteasome is a major site of protein degradation in the cell (Lecker et al. 2006), and its dysfunction is associated with ageing and disease (Vilchez et al. 2014). Our bioinformatics analysis identified a prominent, gut-specific, *dfoxo*-dependent regulation of proteasomal proteins in response to reduced IIS (Fig 7). The 26S proteasome consists of 33 proteins, 14 of which (7-α & 7-β subunits) constitute the 20S proteasome core, while 19 constitute the 19S regulatory particle (Tomko & Hochstrasser 2013). In total, we detected 94% of the proteasomal subunits, and 36% were differentially regulated specifically in the gut in response to reduced IIS (Fig 10). However, we did not detect a consistent pattern or direction of regulation (4 up- and 8 down-regulated, Fig 10). Proteasomal subunit Rpt6R showed the greatest degree of regulation, increasing 2.6-fold in the gut. We confirmed this change and its *dfoxo* dependency, by Western blot analysis of the guts of mNSC-ablated flies compared to controls (Tain et al. 2017). Rpt6, a 19S proteasomal subunit, plays a key role, along with Rpn6, in regulating the assembly and activity of the 26S proteasome holoenzyme (Park et al. 2011; Pathare et al. 2012; Sokolova et al. 2015). Hence, up-regulation of this subunit could indicate an up-regulation of proteasomal assembly. Indeed, the guts of mNSC-ablated flies showed a threefold increase in assembly of the 26S proteasome compared to controls (Tain et al. 2017), as measured using an in-gel assay (Vernace *et al*, 2007). These data suggest that reduced IIS results in increased proteasomal assembly in the gut, possibly through increased levels of Rpt6.

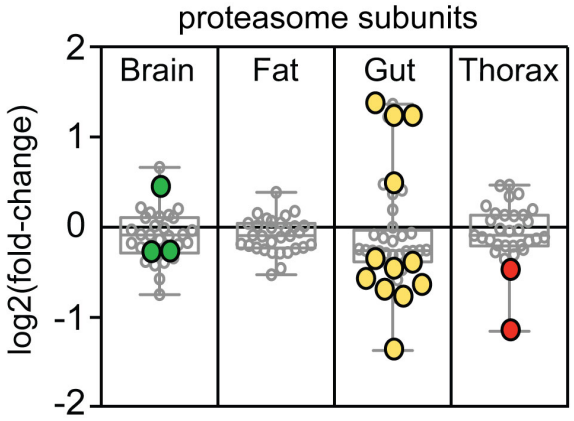


Fig 10*:* ***IIS-dfoxo-dependent regulation of proteasome activity in the gut enhances proteome maintenance***
Tissue-specific log2-fold-changes of IIS-regulated proteins associated to the proteasome. See (Tain et al. 2017) for the full figure with additional experimental results.

3.3.3 Conserved and model-specific mechanisms of the proteomic IIS response

3.3.3.1 Section introduction and overview

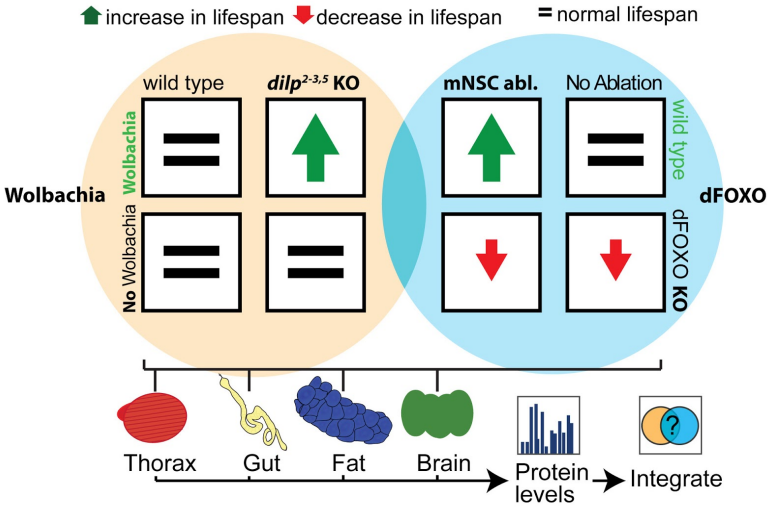


Fig 11: **Integrative analysis of two model systems of reduced insulin signalling.**
We quantified tissue-specific protein levels in two fly model systems of reduced insulin signalling: mNSC-ablation and dilp$^{2-3,5}$ knockout. In the mNSC ablation model, we have further used a dfoxo-null background to delineate the dfoxo-dependent insulin response (Tain et al. 2017). In the dilp$^{2-3,5}$ model, lifespan extension due to knockouts of DILPs was dependent on the presence of the endosymbiont Wolbachia. We therefore quantified proteomes in the absence of Wolbachia as additional controls. We then proceeded to integrate our data to identify shared and specific responses to insulin signalling.

To determine responses to reduced IIS that are conserved or divergent on the proteome level, we investigated an independent model of reduced IIS, *dilp*$^{2-3,5}$ knock-out flies. In contrast to mNSC-ablated flies, *dilp*$^{2-3,5}$ lack expression of Dilps even during early development (Grönke et al. 2010). While differences between models may arise from different courses of development, all major phenotypes of IIS reduction, such as reduced fecundity and growth, increased resistance against toxins, and increased lifespan, are also exhibited by *dilp*$^{2-3,5}$ flies (Grönke et al. 2010). Curiously, the increase in lifespan was dependent on the presence of *Wolbachia*, an intracellular symbiont ubiquitous in arthropod species (Jeyaprakash & Hoy 2000; Werren & Windsor 2000)(Fig 11). Resistance to DDT, generally increased in *dilp*$^{2-3,5}$ flies, was improved further in the presence of the endosymbiont, while fecundity or oxidative stress resistance are independent phenotypes (Grönke et al. 2010). Individual strains of *Wolbachia* can adversely affect lifespan (Min & Benzer 1997) or increase lifespan in interaction with mutant phenotypes (Toivonen et al. 2007). *Wolbachia* can manipulate the immune system, potentially protecting the host from viral infection (Teixeira et al. 2008), while itself evading the immune response of constitutive hosts (Zug & Hammerstein 2015). Furthermore, host diet affects *Wolbachia* titers in flies, and this effect is dependent on TORC1 activity (Serbus et al. 2015). Artificially suppressing

TORC1 activity by rapamycin treatment increased titers, indicating that IIS/TOR signalling generally acts to suppress *Wolbachia* proliferation (Serbus et al. 2015). This is consistent with an earlier study on the effects of the endosymbiont on IIS which found that *Wolbachia* generally causes an increase in insulin signalling (Ikeya et al. 2009), indicating a compensatory response. However, in *dilp*$^{2\text{-}3,5}$ mutants, *Wolbachia* did not result in an increase of 4E-BP (Grönke et al. 2010), suggesting that the endosymbiont affects different, possibly downstream mechanisms of IIS. Here, we analysed the insulin response of *dilp*$^{2\text{-}3,5}$ flies on the level of the proteome and compare it to the mNSC ablation response (Fig 11). We delineate parts of the response that are dependent or independent of *Wolbachia* infection, and use the *Wolbachia* negative background as a control to screen for proteins that are more likely causal for lifespan extension upon reduced IIS. Our analysis provides additional evidence for the regulation of proteostasis in the gut and provides additional insight into how gut barrier function is improved through strengthening of septate junctions in long-lived flies. We also show that up-regulation of respiration in the fat body is a conserved response to reduced IIS. On the other hand, we uncover an improved DNA damage response that is unique to the *dilp*$^{2\text{-}3,5}$ model and largely conditional on *Wolbachia*.

3.3.3.2 Accurate and tissue-specific quantification of the *dilp*$^{2\text{-}3,5}$ model proteome

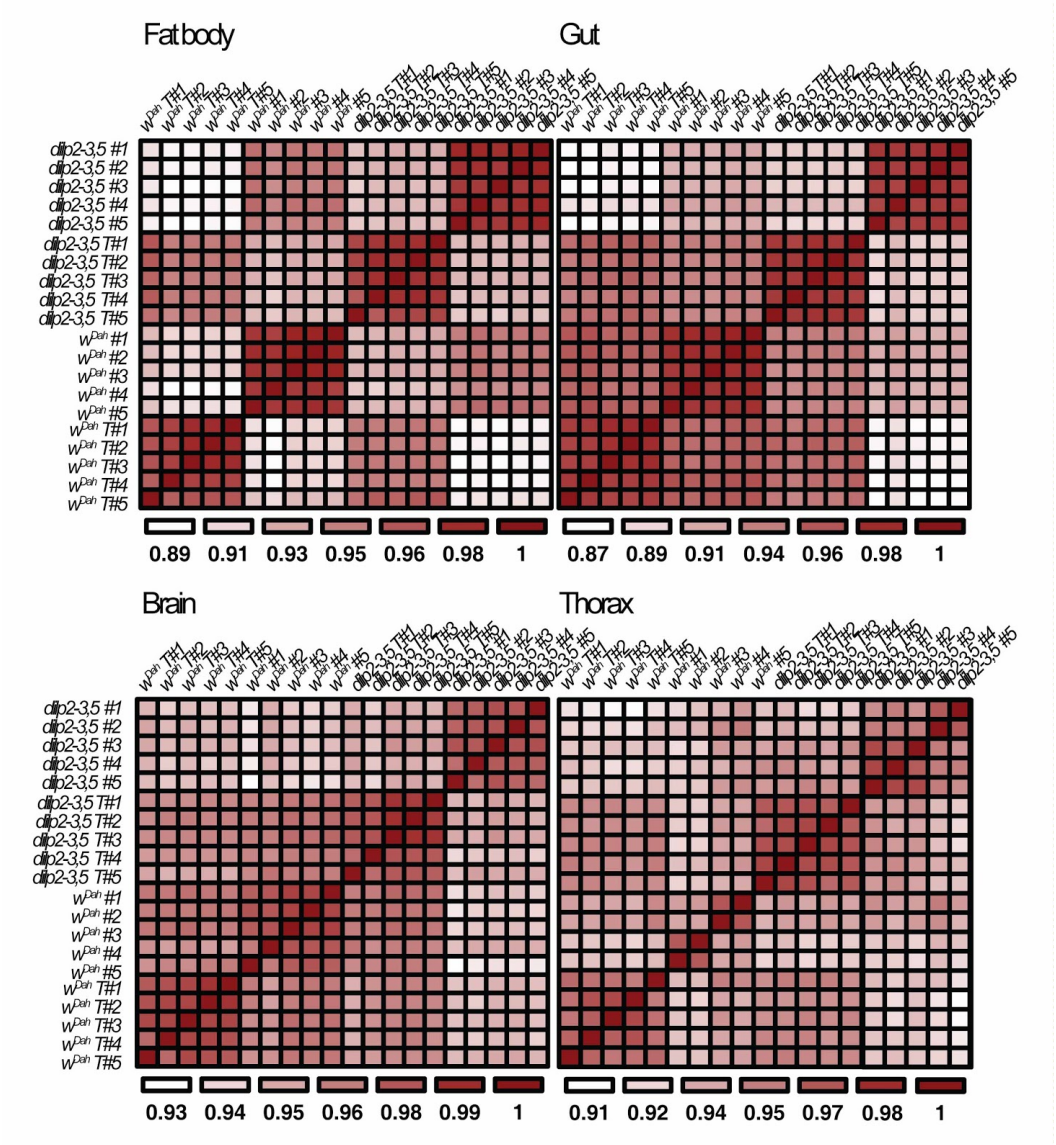


Fig 12*: **Reproducibility of dilp$^{2\text{-}3,5}$ model proteomics.***
Correlation heatmaps of protein quantification in all replicates and tissues. T indicates absence of Wolbachia.

We characterized the tissue-specific proteome of *w*^{Dah}, and long-lived *dilp*$^{2\text{-}3,5}$ mutant flies lacking 3 of the 7 Drosophila insulin-like peptides (Grönke et al. 2010), in brain, thorax, gut, and fat body. Proteomic analysis was performed using single-shot, label-free mass-spectrometry with a Q-exactive benchtop quadrupole-Orbitrap mass spectrometer. To maximise the depth of our analysis and minimise variation measurements were carried out on five biological replicates. All proteomic measurements were highly reproducible between biological replicates (Fig 12).

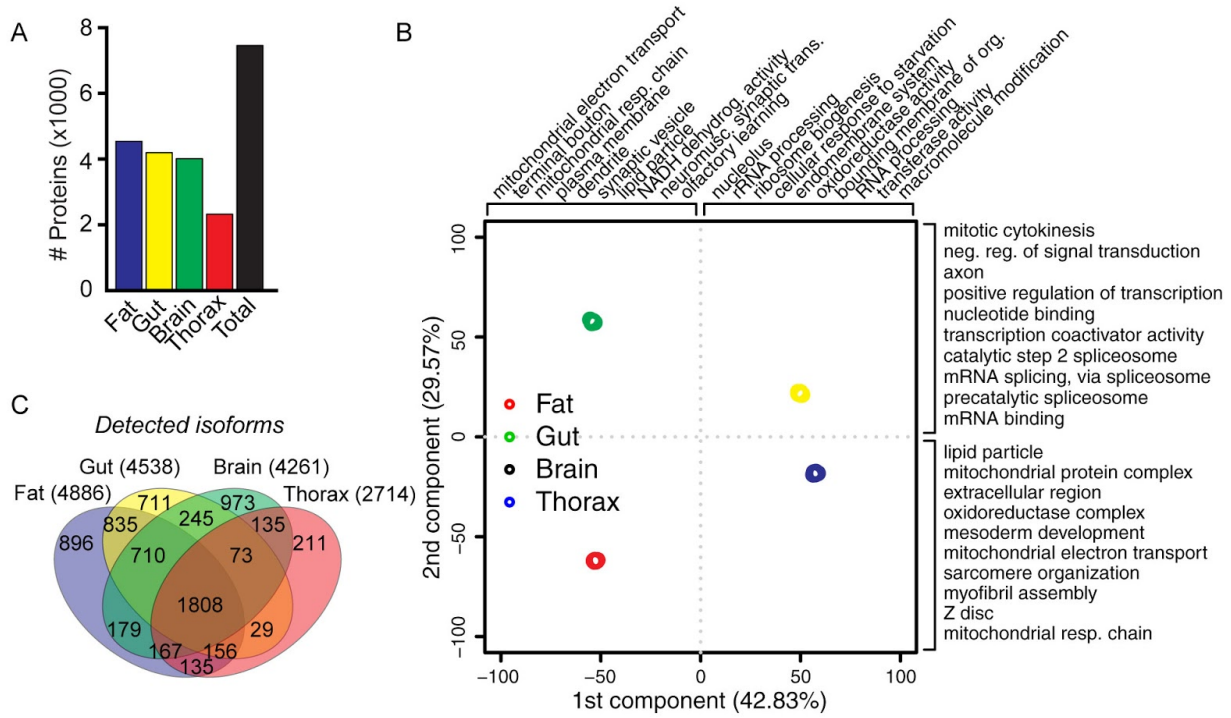


Fig 13*: **Tissue-specificity of the dilp$^{2\text{-}3,5}$ model proteome.***
***(A)** Number of proteins detected in each tissue of the dilp$^{2\text{-}3,5}$ model, and total unique (black). **(B)** PCA on all replicates of the dilp$^{2\text{-}3,5}$ model proteomics quantification. Annotation shows most enriched GO terms of proteins with the top 5% of loading scores of each component in each direction. **(C)** Venn diagram showing detected proteins and isoforms in each tissue, and their overlap between tissues.*

In total, we identified 100805 peptides (at FDR 0.01) representing 7234 proteins and their isoforms, 340 of which originated from *Wolbachia* (Fig 13A). Among the remainder we identified proteins originating from 6693 unique fly gene models, or 48% of protein-coding *Drosophila* gene models. To visualize the extent to which proteomes vary from one tissue to another we performed principal component analysis (PCA), capturing 72% of the variation in the first two PCs (Fig 13B). Individual tissues cluster closely together suggesting intra-tissue variability was much lower than variability between tissues (Fig 13B). Emphasising this tissue-specificity, we identified only 1808 proteins that were common to all four tissues, and 2791 which were present in only one tissue (Fig 13C).

3.3.3.3 Integration of *dilp*$^{2\text{-}3,5}$ and mNSC-ablation proteomics reveals conserved and specific responses to reduced IIS

To assess how reduced IIS affects tissue-specific gene expression in the *dilp*$^{2\text{-}3,5}$ model, we compared protein levels from all four extracted tissues of *dilp*$^{2\text{-}3,5}$ mutants to wild type controls (*w*^{Dah}). We again fitted a linear model (Eq 6). This model was formally equivalent to the model used in the mNSC ablation analysis (Eq 2), but here *Genotype* is a factor with four levels representing the experimental groups in the *dilp*$^{2\text{-}3,5}$ experiment (*dilp*$^{2\text{-}3,5}$, *w*^{Dah}, *dilp*$^{2\text{-}3,5}$*/Wol+*, *dilp*$^{2\text{-}3,5}$*/Wol-*)(Fig 11).

$$Expression = 0 + Genotype \qquad (6)$$

After correcting for multiplicity using the BH method, we found a total of 3738 proteins that were regulated in at least one tissue at a significance level of α = 0.1 (Fig 15A). Importantly, only 22 proteins showed a significantly altered expression across all four tissues, and 50% of the differentially expressed proteins in each tissue were significant in only that tissue, highlighting the tissue-specific response to reduced IIS (Fig 15A).

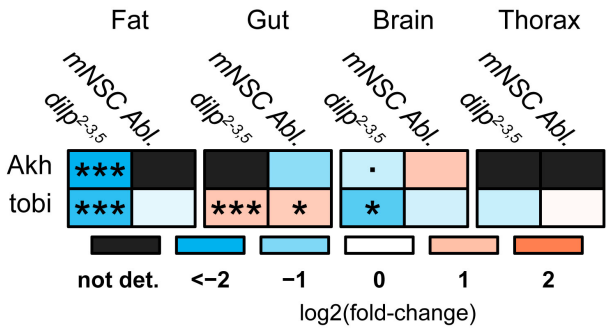


Fig 14***: Regulation of adipokinetic hormone and its target Tobi in response to reduced IIS.***
Heatmap showing the tissue-specific regulation of Akh and Tobi in both models of reduced insulin signalling (dilp$^{2\text{-}3,5}$ *vs. w*^{Dah}*, w*^{Dah}*/Abl vs. w*^{Dah}*). Asterisks indicate BH-corrected significance of the limma moderated t-test (.p ≤ 0.1, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001)*

As an initial check of whether insulin reduction showed the expected effects metabolic signalling across tissues, we inspected the regulation of adipokinetic hormone (*Akh*). Akh is a *Drosophila* glucagon homolog that acts opposite to dILP's to regulate carbohydrate metabolism (Rulifson et al. 2002). It is produced in a subset of the corpora cardiaca, which is physically connected to mNSC's and has been shown to accumulate *dilp2* under normal conditions (Rulifson et al. 2002). We found Akh to be significantly down-regulated in the brains of long-lived *dilp*$^{2\text{-}3,5}$ flies, as was its target Tobi (Buch et al. 2008)(Fig 14), confirming a reduction in metabolic signalling in our mutants (Fig 14), but we failed to detect regulation in mNSC-ablated

flies. Both Akh and Tobi were also down-regulated in the *dilp*$^{2\text{-}3,5}$ fat body (Fig 14), but not mNSC-ablated flies. The stronger phenotype of *dilp*$^{2\text{-}3,5}$ flies with regard to *Akh* might be partially explained by developmental effects of the constitutive knockout that do not occur the mNSC ablation model (Grönke et al. 2010). Curiously, while Akh was reduced in the guts of mNSC-ablated flies (adj. P-Value = 0.11), Tobi was significantly up-regulated in the guts of both insulin models (Fig 14). This suggests the existence of an additional, tissue-specific mechanism of Tobi induction.

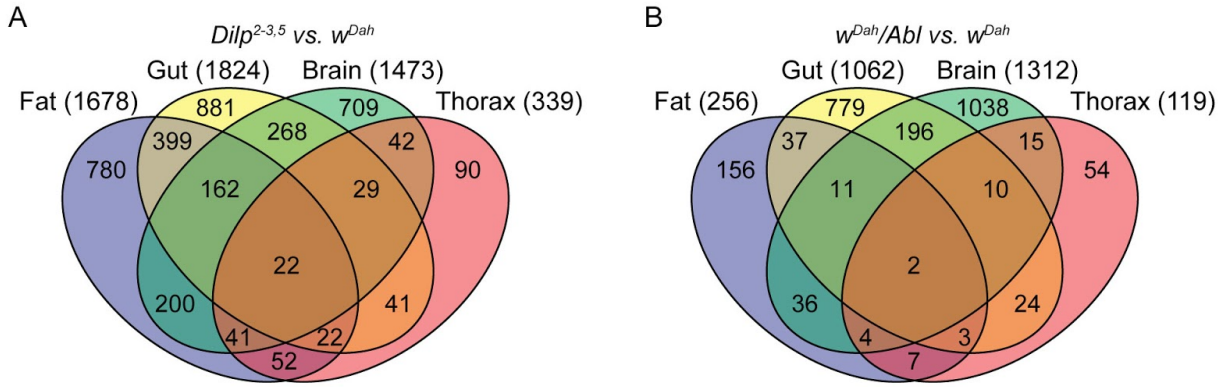


Fig 15: ***Proteome changes induced by mNSC ablation and dilp2-3,5 knockout.***
Venn diagram of differentially expressed proteins upon reduced IIS, ***(A)*** *comparing tissues from control (w*^{Dah}*) and mNSC-ablated flies. Figure repeated from* Fig 4*.* ***(B)*** *comparing tissues from control (w*^{Dah}*) and dilp*$^{2\text{-}3,5}$ *mutants. Total number of differentially expressed, tissue-specific proteins is shown in parentheses.*

To determine which responses to reduced IIS were shared or distinct in both reduced IIS models, we partitioned proteins based on significance and direction of regulation in the comparisons *dilp*$^{2\text{-}3,5}$ vs. *w*^{Dah} (Fig 15A) and mNSC-ablated vs. *w*^{Dah} (Fig 16B)(File 2). For proteins identified as differentially expressed in both models (adj. P-Value ≤ 0.1), we saw a moderate but highly significant rank correlation of log-fold-changes (Fat: 0.36, Gut: 0.29, Brain: 0.35, Thorax: 0.63). To be more permissive in our initial integration, we combined the p-values of *dilp*$^{2\text{-}3,5}$ and mNSC-ablation responses using the Stouffer method (Whitlock 2005; Stouffer et al. 1949), correcting for multiple hypothesis testing using the BH method. By design, this combined p-value is lower for proteins that change significantly in both models, and more conservative for those regulated in only one model. We intentionally did not penalise or exclude proteins that are regulated in opposite directions in each model, as these might represent interesting model-specific responses.

We continued by dividing the response into four quadrants based on their expression changes: increase in both models (I), decrease in both models (III), increase in *dilp*$^{2\text{-}3,5}$ but decrease in mNSC-ablation (II), and decrease in *dilp*$^{2\text{-}3,5}$ but increase in mNSC-ablation (IV). We then

performed GO enrichment analysis on significant proteins in each of these four quadrants per tissue using Fisher tests (Rivals et al. 2007). We used tissue-specific backgrounds (all proteins detected in the respective tissue). To reduce their redundancy, GO terms were clustered together if they showed a high overlap of significant proteins, i.e. if they differed by fewer than 5 proteins. In this case, the smallest, most specific term is chosen to represent the cluster. This method, coined SETHRO (simple enrichment testing for highly redundant ontologies), is described in detail in the appendix of this thesis. After multiplicity correction using the BH method, we selected significantly enriched GO terms below an adj. p-value cutoff of 0.05.

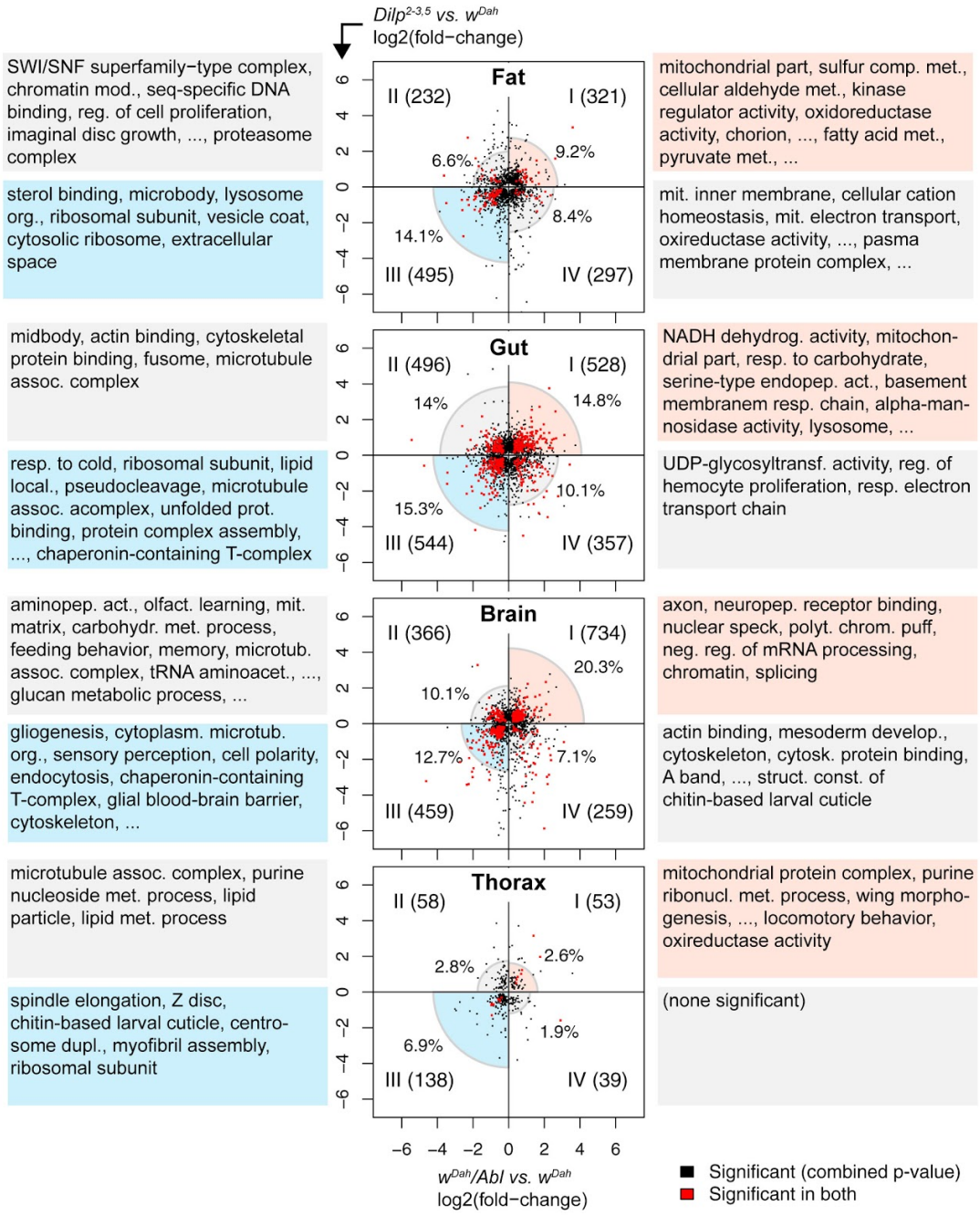


Fig 16**: Integration of the mNSC-ablation and dilp$^{2\text{-}3,5}$ insulin responses.**

Scatterplots showing log-fold-changes from the dilp$^{2\text{-}3,5}$ vs. w^{Dah} (y-axis) and w^{Dah}/Abl vs. w^{Dah} (x-axis) comparisons. Black dots show proteins with a significant combined p-value after multiplicity adjustment ($p \leq 0.1$). Numbers in quadrants show the number of these significant proteins in each quadrant. Circle-segments and percentages represent the fraction of significant proteins over total proteins in the given tissue. Red dots signify proteins that are individually significant in both proteomics comparisons. Representative GO terms for each quadrant are shown (also see File 4*).*

In the fat body, functional analysis revealed that proteins associated to ribosomal subunits were enriched among those down-regulated in both long-lived fly models (Fig 16), indicating that a fat-specific reduction in translational activity was a conserved response to decreased IIS. We could confirm that the tissue-specific reduction in translation rate in the fat body of mNSC-ablated flies (Fig 8) was recapitulated in *dilp*$^{2\text{-}3,5}$ mutants (Tain et al. 2017). Likewise, proteins of the mitochondrial matrix in were up-regulated across both models (Fig 16). We have previously shown that respiration was increased in a tissue-specific manner in the fat body of mNSC-ablated flies (Tain et al. 2017) and could confirm that this phenotype was mirrored in long-lived *dilp*$^{2\text{-}3,5}$ flies (Tain et al. 2017). Proteins associated to the DNA damage response, however, were up-regulated specifically in the fat body of long-lived dilp$^{2\text{-}3,5}$ flies and not in the mNSC ablation (Fig 16). Increased DNA damage is associated with age and transposon burden is known to correlate with an increase in double-strand breaks (Wood et al. 2016).

In both gut and brain of long-lived flies functions relating to proteostasis were regulated (Fig 16). Proteins belonging to the chaperonin-containing T-complex, a mediator of protein folding in the cytosol (Kubota et al. 1995), were diminished in both tissues.

In the gut specifically, we observed a regulation of protein folding and protein complex assembly (Fig 16), confirming the increase in proteasome efficiency in mNSC ablated flies we reported previously (Tain et al. 2017). We also observed an enrichment of mitochondrial and ETC proteins among down-regulated proteins of both models, confirming that this is a robust response in both models (Tain et al. 2017)(Fig 16). Furthermore, the expression of most lysosomal alpha-mannosidases, which are required for the breakdown of glycoproteins and interact with the proteasome to regulate homeostasis of N-linked glycoproteins (Winchester 2005), was strongly increased in long-lived flies of both models (Fig 16). This may contribute to the robust and conserved proteostasis response to reduced IIS in the gut.

The brain showed a decreased expression of proteins associated to gliogenesis and the glial blood-brain barrier (BBB) in both long-lived fly models (Fig 16). This is surprising, as a compromised BBB is expected to have an adverse phenotype (Limmer et al. 2014). While it is known that endosymbionts such as *Wolbachia* act to circumvent host defenses (Zug & Hammerstein 2015), further studies will be needed to determine whether *Wolbachia* can affect blood brain barrier permeability of its host. Additionally, components of the mRNA processing and splicing machinery were up-regulated in long-lived flies (Fig 16). Splicing homeostasis has a role in lifespan extension via dietary restriction and TORC1 pathway components in *C. elegans* (Heintz et al. 2016) and was reported to change with age in human brains (Mazin et al. 2013). Regulation of RNA processing is therefore likely to be a conserved mechanism of longevity assurance in response to IIS or DR that is of particular importance in neuronal tissues.

In the thorax, we observed a decreased expression of proteins with muscle related annotations such as Z-disc and myofibril assembly (Fig 16). However, proteins involved in mitochondrial

functions and locomotory behavior control were up-regulated (Fig 16). Age-related loss of locomotory ability is common across model organisms and can be mitigated by increased dFOXO/4E-BP activity in flies (Demontis & Perrimon 2010). To determine if the observed changes led to functional changes we therefore examined age-related muscle function by performing a climbing assay. *Dilp*$^{2\text{-}3,5}$ mutants flies out performed control flies at all ages (Fig 17), suggesting that reduced IIS may influence the proteome of muscle cells to prevent age-related loss of muscle function and structure with age (Brandt et al. 2017; Rhodenizer et al. 2008; Ferguson et al. 2005).

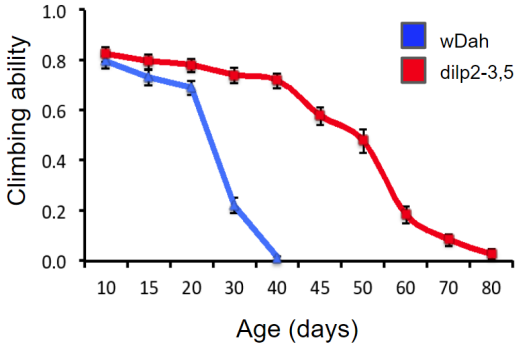


Fig 17: ***dilp*$^{2\text{-}3,5}$ mutants retain their climbing ability for longer than wild type flies.**
Line chart showing decline of climbing ability with age, in w^{Dah} and *dilp*$^{2\text{-}3,5}$. Experimental results and plot by Luke Tain. Climbing assays were performed by using a countercurrent apparatus. The Climbing index was calculated as the weighted average of the chamber into which the flies climbed divided by four times the number of flies in the assay (n>80). Error bars represent the standard error of the mean.

3.3.3.4 The Wolbachia-dependent response to reduced IIS

To further narrow down on candidate proteins mediating lifespan extension in *dilp*$^{2\text{-}3,5}$ flies, we investigated the role of *Wolbachia* in IIS regulation. Presence of the endosymbiont *Wolbachia* is required for lifespan extension in *dilp*$^{2\text{-}3,5}$ *(Grönke et al. 2010)*, though it remains unclear how this interaction is mediated. One possibility is that *Wolbachia* rescues a deleterious side-effect of the *dilp*$^{2\text{-}3,5}$ mutant that, in the absence of *Wolbachia*, negates benefits of reduced insulin signalling. Therefore, not all *dilp*$^{2\text{-}3,5}$ induced protein expression changes connected to the regulation of lifespan extension are expected to depend on *Wolbachia*. However, if such a dependence were to be detected, it would suggest a more likely involvement in crucial downstream pathways mediating lifespan.

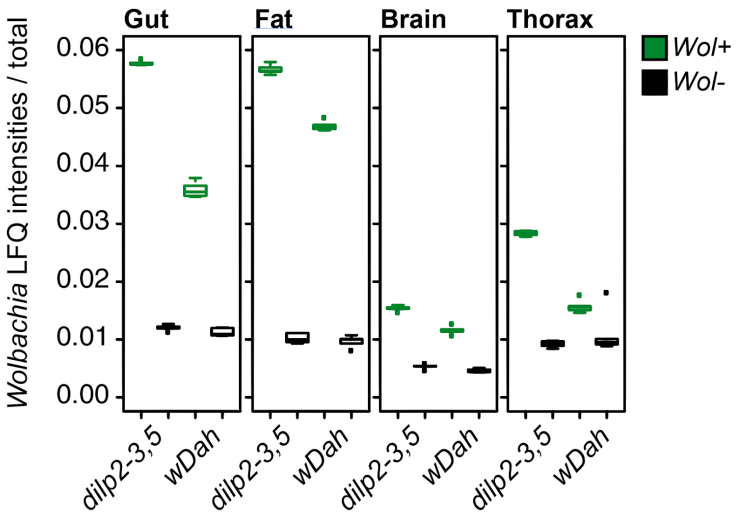


Fig 18**: Abundance of Wolbachia in wDah tissues.**

*Boxplots showing the abundance of Wolbachia in each tissue and genotype, expressed as the fraction of Wolbachia protein intensities over the sum of all protein intensities in each replicate. Wolbachia-positive genotypes are represented by green boxes, Wolbachia-negative genotypes by black boxes. Box widths represent the interquartile range (IQR), box whiskers extend to the farthest data point up to 1.5*IQR from the box border.*

We first checked for differences in abundance of *Wolbachia* between our experimental groups. In total, we identified 340 proteins originating from *Wolbachia* in our $dilp^{2\text{-}3,5}$ proteomics screen. We then calculated the intensity of all *Wolbachia* proteins as a fraction of total intensity in each sample. A very low degree of putative *Wolbachia* protein expression could be detected even in *Wolbachia*-negative samples, indicating false positive protein identifications or mis-annotations (Fig 18). However, we found that the endosymbiont had the highest relative abundance in long-lived mutants, particularly in fat and gut (Fig 18), consistent with reports that reducing IIS/TOR signalling increases *Wolbachia* titers (Serbus et al. 2015).

We next asked which *Drosophila* proteins differed in their response to reduced IIS depending on *Wolbachia*. Starting from the set of proteins that were significantly regulated in response to reduced IIS in $dilp^{2\text{-}3,5}$ flies, we selected those that showed a significant interaction effect with regard to *Wolbachia*-negative controls. As previously for the identification of *dfoxo*-dependently regulated proteins in the mNSC-ablation model (Eq 5), we implemented a difference-of-differences interaction term (Eq 7). Proteins that were significantly interacting but not changed in the *Wolbachia*-positive background were disregarded in the following analysis.

$$Interaction_{diff} = (dilp^{2\text{-}3,5} - w^{Dah}) - (dilp^{2\text{-}3,5}/\text{Wol-} - w^{Dah}/\text{Wol-}) \qquad (7)$$

In total, we identified 453 *Wolbachia*-dependent proteins, most of them in the fat body (Fig 19). We were also interested in pleiotropic elements of the reduced IIS response, i.e. effects that are

not necessarily mediating the observed lifespan extension, or mechanisms that do so independently of *Wolbachia*. We defined proteins as independent of *Wolbachia* if they responded to *dilp*$^{2\text{-}3,5}$ knockout in the same manner regardless of infection status, i.e. showed no difference-of-differences (Eq 7). This was done by explicitly testing whether the absolute magnitude of the interaction effect with regard to *Wolbachia*-negative controls was less than a low threshold, using a TOST procedure (Westlake 1976). We could identify 949 such proteins.

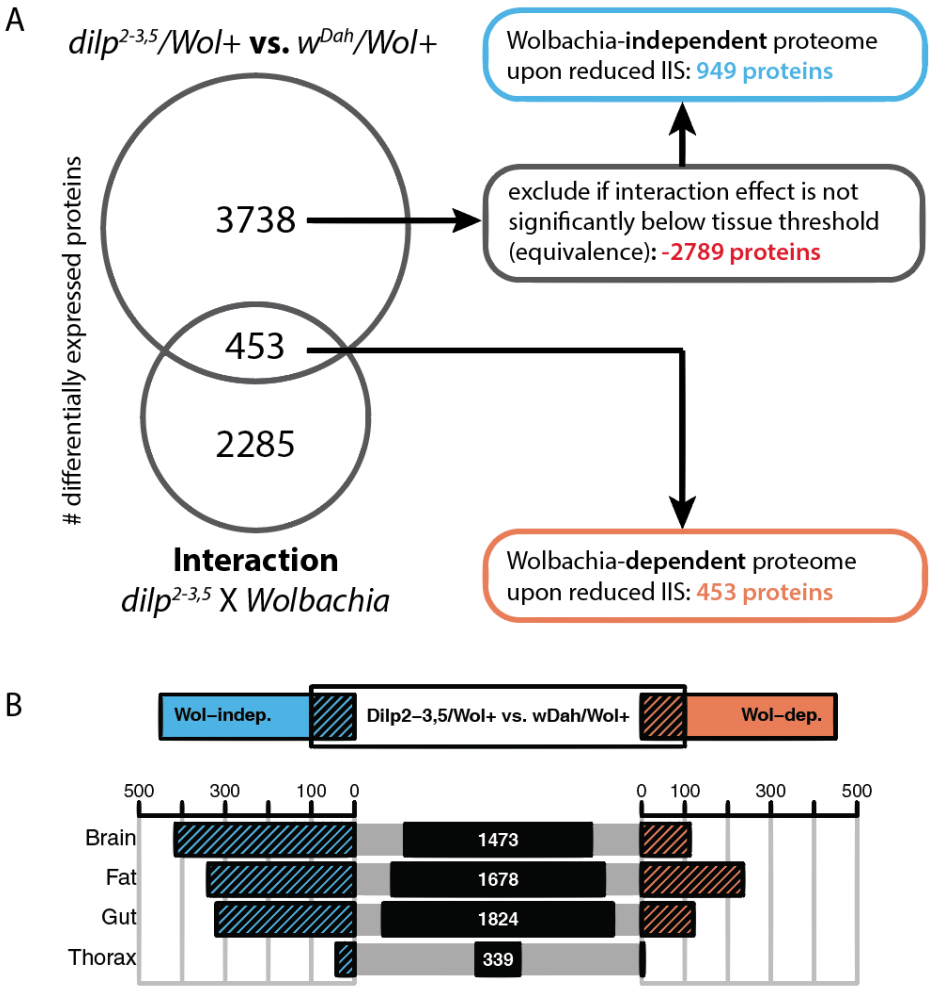


Fig 19: ***Wolbachia*-dependence of insulin-responsive proteins.**
(A) *Reduced IIS-mediated changes in dilp*$^{2\text{-}3,5}$ *vs. w*^{Dah} *(Wolbachia present, Wol+) or their interaction term with Wolbachia-negative controls (dilp*$^{2\text{-}3,5}$*/Wol+ - w*^{Dah}*/Wol+) vs. (dilp*$^{2\text{-}3,5}$*/Wol- - w*^{Dah}*/Wol-). Wolbachia-dependent proteins were defined as proteins regulated between dilp*$^{2\text{-}3,5}$ *vs. w*^{Dah} *and within the interaction term. Proteins were identified as Wolbachia-independent if they were regulated between dilp*$^{2\text{-}3,5}$ *vs. w*^{Dah}*, and were not subject to interaction effects (equivalence given a tissue-specific threshold).* ***(B)*** *Symmetric barplots (symplots) showing the number of Wolbachia-dependent (orange) and Wolbachia-independent (blue) proteins in each tissue. The middle column shows the total number of proteins regulated between dilp*$^{2\text{-}3,5}$ *and w*^{Dah}*.*

We then carried out GO enrichment analysis of *Wolbachia*-dependent proteins to identify functions that were regulated specifically in the presence of the endosymbiont (Fig 20).

In the fat body, many proteins involved in the previously identified, *dilp*$^{2\text{-}3,5}$-specifically elevated DNA damage response (Fig 16) were regulated in a *Wolbachia*-dependent manner (Fig 20). Other dependent responses included parts of carbon metabolism, as well as proteins involved in chorion formation (Fig 20).

In the gut, we observed that the up-regulation of many structural chitin proteins was *Wolbachia*-dependent (Fig 20). Chitin is a component of the peritrophic matrix that is an integral part of the gut barrier in insects (Lehane 1997). Its up-regulation may therefore have a beneficial effect on gut barrier function. Furthermore, five out of seven lysosomal α-mannosidases were increased specifically in *Wolbachia*-positive *dilp*$^{2\text{-}3,5}$ flies (LManI, LManII, LManIII, LManV, LManVI)(Fig 20). Lysosomal α-mannosidases cleave Asn residues from n-linked glycoproteins to allow their recycling in the lysosome (Aronson & Kuranda 1989). Thus their up-regulation suggests that presence of *Wolbachia* might contribute to modulating IIS to achieve improved proteostasis as reported by us in (Tain et al. 2017).

In the brain, 22 cuticular proteins (*Cpr*'s) were down-regulated in a *Wolbachia*-specific manner in *dilp*$^{2\text{-}3,5}$ flies (Fig 20). However, only one of these, *Cpr47Ef*, was regulated in the same direction in mNSC-ablated flies. Furthermore, the annotation of *Cpr's* is based almost entirely on the results of one study which used sequence homology with a known cuticular protein motif (Karouzou et al. 2007). Their similarity to structural proteins of the extracellular matrix might indicate that *Cpr's* serve more general functions in the extracellular matrix of brain cells, such as participating in the regulation of the blood brain barrier, but additional experiments are needed to determine this.

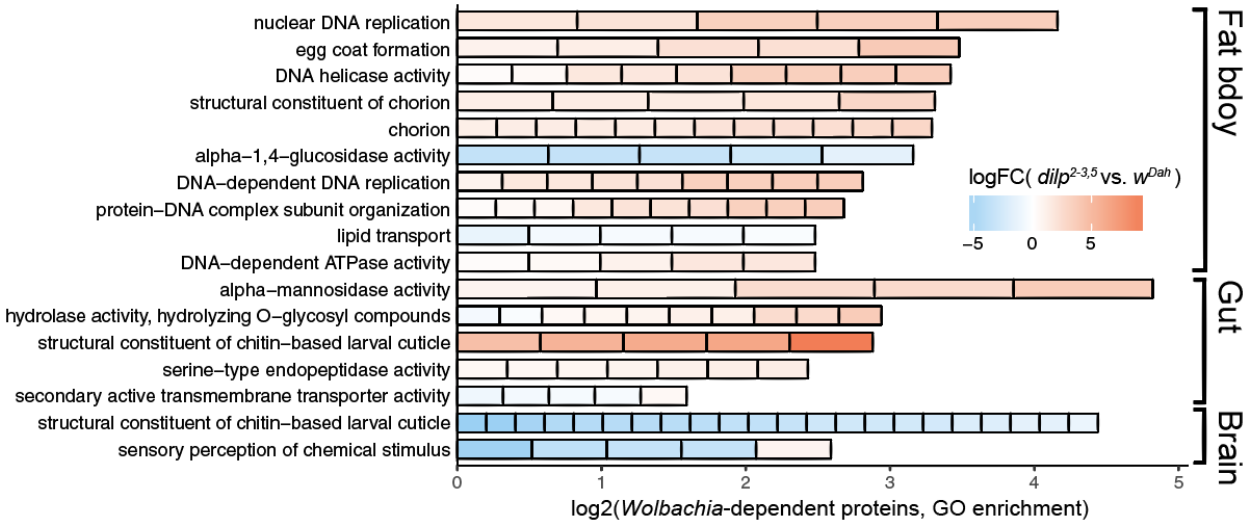


Fig 20: **Functions of Wolbachia-dependent proteins.**
Segmented bar charts showing the significant results from GO enrichment analysis on the set of

Wolbachia-dependent proteins in each tissue (none significant in thorax). Length of each bar shows the log2-transformed enrichment over the tissue-specific background. Number of cells of each bar equals the number of Wolbachia-dependent proteins annotated with the term in the respective tissue. Colours of cells represent log-fold-changes of proteins between dilp$^{2\text{-}3,5}$ and w^{Dah}.

3.3.3.5 The *dfoxo*- and *Wolbachia*-dependent conserved insulin response

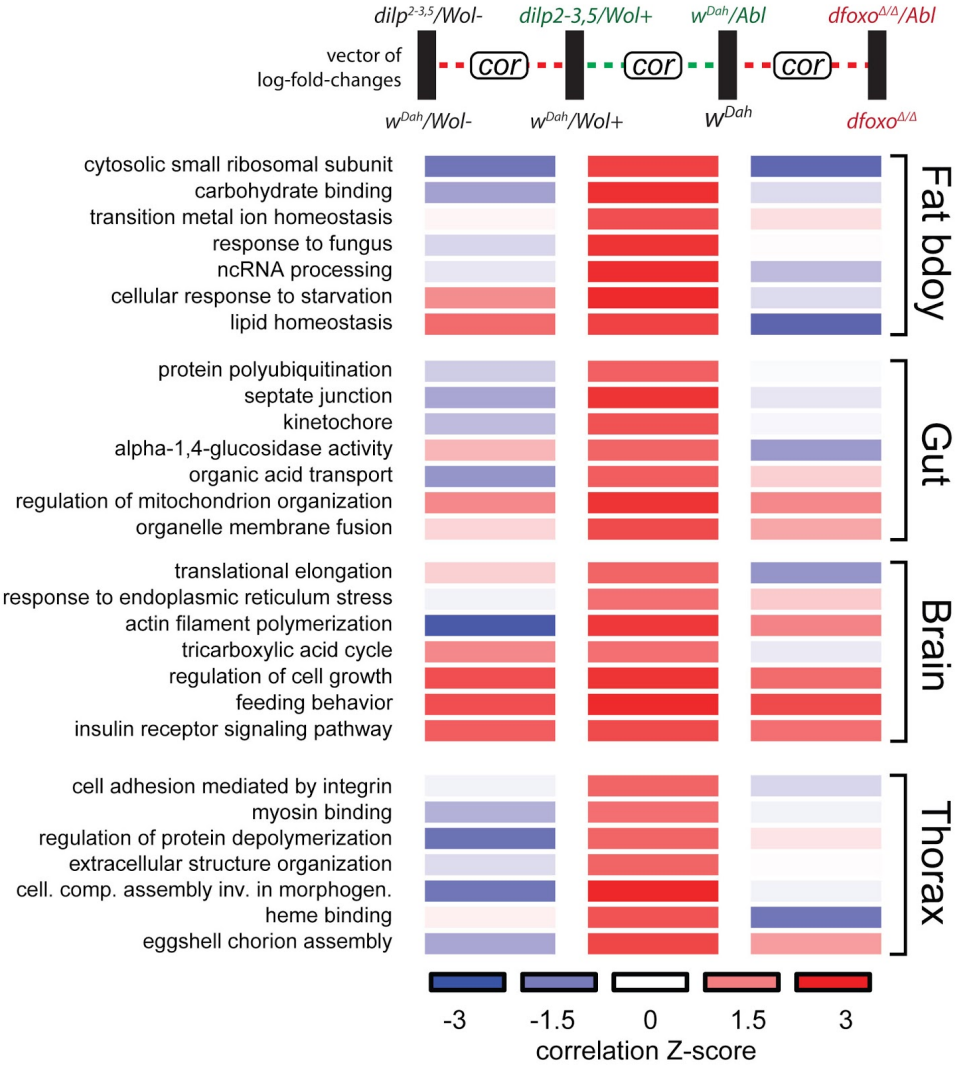


Fig 21**: Functional correlation analysis to integrate *dilp$^{2\text{-}3,5}$* and mNSC-ablation models.**
Heatmap representing correlations between log-fold-changes of proteins annotated with GO terms in fat, gut, and brain tissue. Log-fold-changes were derived from the following comparisons: dilp$^{2\text{-}3,5}$/Wol- vs. w^{Dah}/Wol- (control), dilp$^{2\text{-}3,5}$/Wol+ vs. w^{Dah}/Wol+ (long-lived), w^{Dah}/Abl vs. w^{Dah} (long-lived), dfoxo$^{\Delta/\Delta}$/Abl vs. dfoxo$^{\Delta/\Delta}$ (control). Correlations were z-transformed using empirical distributions obtained by permutation. Up to seven representative, high scoring GO terms (z-value $\geq$ 2 in central column) were selected in each tissue and sorted in descending order of their maximal difference to controls.

To discriminate between parts of the insulin response that are conserved in long-lived *dilp$^{2\text{-}3,5}$* and mNSC-ablation flies, while simultaneously taking into account their *Wolbachia*-negative or

dfoxo-knockout controls, we carried out a correlation analysis. We based the analysis on the differences between long-lived flies and normal lived controls in each of the two models (w^{Dah}*/Abl vs.* w^{Dah}, and $dilp^{2\text{-}3,5}$ *vs.* w^{Dah}). We selected the log-fold-changes of proteins annotated within a size-filtered list of GO terms (between 10 and 400 annotated genes). A coordinated, conserved insulin response of a functional term will manifest in a high positive correlation coefficient between the long-lived proteome responses.

To reduce redundancy of GO terms and increase interpretability, we employed an iterative filtering scheme: descending the list of GO terms from most highly correlated between long-lived responses to most negatively correlated, we added all annotated proteins from each term to an exclusion list. As correlations show higher variances at lower numbers of observations, this procedure will select smaller, more specific GO terms first. We then required each subsequent term to contain at least 10 proteins that were not observed in any previous term (i.e. on the exclusion list), or be discarded. To make terms of different sizes directly comparable and account for prior positive correlation of log-fold-changes, empirical distributions of correlations for terms of any observed size (number of proteins in the term) were calculated by repeated shuffling of all observed log-fold-changes in each comparison. The actually observed correlations were then z-transformed using the parameters of the empirical distribution. We further selected terms that showed a high correlation, i.e. z-scores in excess of 2 standard deviations above the mean (File 2). In addition to the correlations between long-lived comparisons, comparisons performed in two control backgrounds were then included in the analysis: $dilp^{2\text{-}3,5}$ vs. w^{Dah} in absence of *Wolbachia*, as well as mNSC ablation vs. w^{Dah} in a *dfoxo*-null background. Both the absence of *Wolbachia* and dfoxo-knockout lead to an abrogation of the lifespan-extension phenotype in their respective model system, but may play limited roles in pleiotropic phenotypes (Slack et al. 2011). A high correlation to these controls does not exclude an involvement in lifespan extension entirely, but a low correlation lends additional confidence that a functional term is involved in the extension of lifespan observed in the respective model.

In the fat body, we observed several functional terms that were correlated between the long-lived insulin responses, but not with regard to either of the controls. These included regulation of the cytosolic ribosome (Fig 21), supporting our observation that fat body translation is modulated specifically in long-lived flies (Fig 8). The concordant regulation of ncRNA processing suggests that non-coding RNA may play a role in the longevity-specific insulin response (Fig 8). Interestingly, while carbohydrate metabolism was longevity-specifically regulated, proteins annotated with lipid homeostasis and starvation response were induced in a correlated manner independently of *Wolbachia* (Fig 21).

In the gut, several functional terms were closely associated between long-lived comparisons, but dysregulated in normal lived controls. These included protein polyubiquitination, a process that is integral to protein turnover and recycling by the proteasome. We have previously shown

that proteasome activity, along with clearance of polyubiquitinated proteins, is improved in the gut of mNSC-ablated flies (Tain et al. 2017). We could confirm that this was also the case in *dilp*$^{2\text{-}3,5}$ mutants (Tain et al. 2017). In addition to that, septate junction proteins were regulated in a consistent manner between long-lived fly models.

In the brain, proteins belonging to the insulin receptor signalling pathway showed a coordinated response in long-lived phenotypes, reflecting the similarity of the interventions that result in the loss of *dilp* expression in both model systems (Fig 21). This pattern was independent of *Wolbachia* and *dfoxo* status, as was regulation of cell growth (Fig 21). Regulation of the tricarboxylic acid cycle was consistent between long-lived contrasts independently of *Wolbachia*, but altered in the *dfoxo*-knockout (Fig 21). The same was true for translational elongation. We failed to detect a change in translation rate in heads of long-lived mNSC-ablated flies (Fig 8); however, a more precise essay using only brain tissue might yield different results. Additionally, actin filament polymerization was highly correlated between long-lived responses and with regard to *dfoxo*-null controls, but anti-correlated in *Wolbachia*-negative controls (Fig 21). *Wolbachia* relies on the host actin cytoskeleton for effective transmission and can directly manipulate it (Newton et al. 2015; Sheehan et al. 2016). In particular, we observed *Wolbachia*-dependent up-regulation of the actin interacting protein Flr and WASp, as well as down-regulation of the highly conserved Arp2/Arp3 complex (Kang et al. 2010)(File 2). Arp2/3 induction facilitates Akt membrane recruitment upstream of dFOXO (Zhao et al. 2015), linking actin cytoskeleton organization to insulin signalling. This suggests that *Wolbachia* may modulate insulin signalling through manipulation of cytoskeleton components in the brain.

In the thorax, the longevity-specific concordant regulation of proteins involved in cell adhesion, myosin binding, and structural organization point towards an effect on muscle function. We observed a beneficial effect of reduced insulin on climbing ability of ageing *dilp*$^{2\text{-}3,5}$ mutants, as shown earlier in this thesis (Fig 17). While our essay did not encompass *dfoxo*-null flies, our correlation analysis supports *dfoxo*-dependence of improved muscle function upon reduced IIS, corroborating previous reports (Demontis & Perrimon 2010).

3.3.3.6 Increased gut barrier function in long-lived insulin mutants

With age, gut barrier integrity is weakened (Rera et al. 2012), leading to more frequent or chronic inflammation. One important determinant of integrity of epithelia in general, and therefore gut barrier permeability, are septate junctions (Farquhar & Palade 1963; Bonnay et al. 2013; Lee 2015). While there was no clear trend towards either up- or down-regulation of septate junction proteins in long-lived flies, key proteins were regulated in a way that predicts reduced permeability: Yurt (*yrt*), a positive regulator of septate junction permeability (Laprise et al. 2006), was reduced by more than 2-fold in both *dilp*$^{2\text{-}3,5}$ and mNSC-ablated flies. Conversely,

Lac, previously reported to be required for formation of the trans-epithelial diffusion barrier in the *Drosophila* tracheal epithelium (Llimargas et al. 2004), was up-regulated in both models. Additionally, we investigated expression of proteins that may be involved in the peritrophic matrix (PM). In insects the PM acts as a physical, mesh-like, barrier between the contents of the gut lumen and the epithelial cells of the gut (Lehane 1997). Diminishing the PM negatively affects gut function and integrity (Kuraishi et al. 2011). We observed a strong and consistent up-regulation of proteins involved in chitin-based cuticle development in the gut (Fig 16), which might suggest that long-lived IIS mutants possess a denser peritrophic matrix.

To directly determine the functional consequence of these changes, we quantified age-associated changes in gut integrity in control and long-lived flies. Feeding non-absorbable blue food dye to flies allows direct assessment of gut integrity, with loss of integrity resulting in blue dye leaking into the hemolymph giving “smurf” flies (Rera et al. 2012). As previously reported (Rera et al. 2012), control flies show an age-associated loss of gut integrity, with the proportion of “smurfs” in the population increasing with age. However, this was suppressed in mNSC-ablated flies, and independently in $dilp^{2\text{-}3,5}$ mutants (Tain et al. 2017). We hypothesised that the improved gut integrity of long-lived flies may be, at least in part, caused by the observed *dfoxo*-dependent increase in proteasomal activity (Tain et al. 2017)(Fig 10, Fig 7). To test this hypothesis, we orally administered bortezomib, a potent inhibitor of the proteasome, to flies with reduced IIS and controls, for the duration of their lifespan. At a concentration of 2 µM, bortezomib did not reduce the lifespan of wild-type flies, but significantly reduced the lifespan-extension of mNSC-ablated flies and, independently, of $dilp^{2\text{-}3,5}$ mutants (Tain et al. 2017). Using Cox proportional hazards, we confirmed that treatment with bortezomib had a significantly different effect on survival in the two models of reduced IIS flies compared to their wild-type controls ($P < 0.01$)(Tain et al. 2017). Importantly, the treatment also dramatically increased the proportion of “smurfs” in both mNSC-ablated flies and $dilp^{2\text{-}3,5}$ mutants (Tain et al. 2017). Together, these results indicate that increased proteasomal activity in the gut is a conserved response to reduced IIS that leads to increased gut integrity, and is required for the lifespan extension of IIS mutants.

3.3.3.7 Improved DNA damage response in long-lived *dilp*$^{2\text{-}3,5}$ mutants

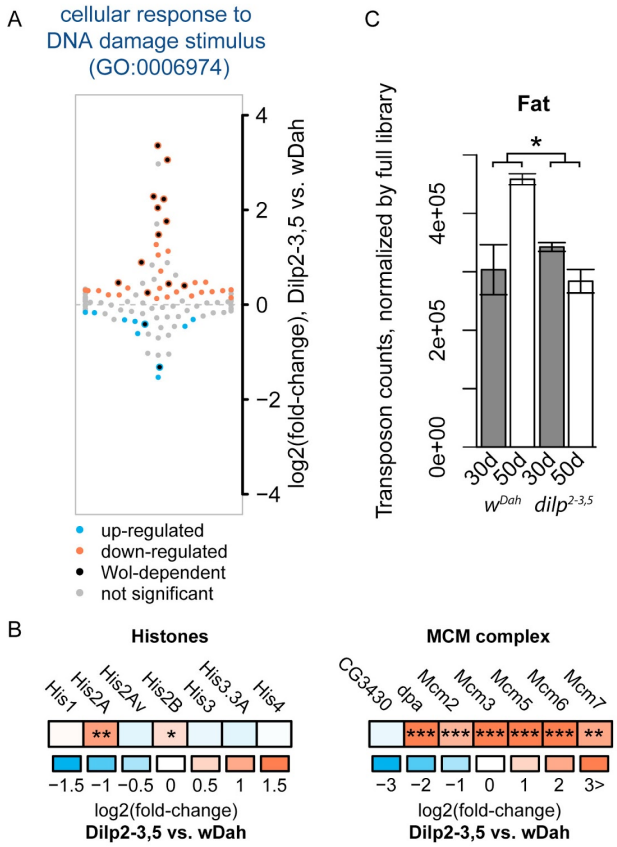


Fig 22**: DNA damage response in the *dilp*$^{2\text{-}3,5}$ model fat body.**
(A) *One-dimensional scatter plot showing the response of proteins involved in the DNA damage response, dilp*$^{2\text{-}3,5}$ *vs. w*^{Dah}*. Up-regulated proteins are represented by orange dots, down-regulated proteins by blue dots, and not significantly changed proteins by grey dots. Wolbachia-dependent proteins are marked by an additional black dot.* ***(B)*** *Regulation of histones and MCM complex proteins. Asterisks indicate significance levels.* ***(C)*** *Transposon burden in 30d and 50d old flies, dilp*$^{2\text{-}3,5}$ *and w*^{Dah}*, expressed in total counts after library normalization. Asterisk indicates significance of the two-way ANOVA interaction term (* $p \leq 0.05$).*

We discovered that key elements of the DNA damage response were strongly induced in the fat body of long-lived *dilp*$^{2\text{-}3,5}$ flies (Fig 22A), but not the independent mNSC-ablation model of insulin signalling (Fig 20). Curiously, much of the detected DNA damage response was conditional on the presence of the intracellular symbiont *Wolbachia* (Fig 22A). All members of the MCM complex, which regulates replication initiation and DNA repair (Bailis & Forsburg 2004) were increased in long lived IIS mutants (Fig 22B). Furthermore, we found that H2A was increased in *dilp*$^{2\text{-}3,5}$ flies (Fig 22B), but failed to detect any corresponding regulation of the H2A variant H2Av (Fig 22B), which is assembled into nucleosomes at the site of DNA double-strand

break repair (Rogakou et al. 1998; Madigan et al. 2002). Thus suggests that the ratio of H2Av to H2 nucleosomes was decreased in long-lived flies, possibly reflecting an overall decrease in DNA damage as a result of reduced IIS.

Previous reports have indicated that transposable elements as well as DNA damage accumulate with age in the *Drosophila* fat body (Wood et al. 2016). In support of this, we note that Piwi, known to facilitate transposon silencing in the fat body of Drosophila (Jones et al. 2016), was up-regulated more than 16-fold in long-lived $dilp^{2\text{-}3,5}$ mutants. We therefore decided to investigate if the observed changes correlated with a decrease of transposon burden in the $dilp^{2\text{-}3,5}$ fat body. For this, we probed preliminary data from the RNAseq quantification of fat bodies from 30d and 50d old $dilp^{2\text{-}3,5}$ mutant and their controls for signs of transposon regulation. We obtained a list of known fly transposon sequences from flybase (Gramates et al. 2017) and used RSEM to quantify reads mapping to these sequences in our RNAseq data. RSEM was chosen because it is a transcript-level alignment tool (Li & Dewey 2011), making it ideally suited for this task. However, to compare transposon burden across experimental groups, it was further necessary to normalize transposon expression using a global reference. For this, we decided to pool the transposon read counts and separately mapped *Drosophila* gene read counts using HISAT2. We then calculated sample correction factors using TMM normalization and applied them to transposon counts – thereby normalizing them with regard to overall *Drosophila* transcript expression. We summed up the expression of transposons for each sample and compared transposon burden in 30d and 50d old $dilp^{2\text{-}3,5}$ and w^{Dah} flies. For this, we fitted a linear model (transposon burden ~ day * genotype). We then performed ANOVA to determine the significance of the interaction term. Our preliminary findings indicate that transposon burden did not change in the fat body of 50d-old $dilp^{2\text{-}3,5}$ mutants compared to their younger (30d) counterparts, but was increased in the wild type (Fig 22C). Due to concerns about the overall quality of these RNAseq data, and for additional validation, a targeted assay of specific transposons is underway at the time of writing. However, our preliminary findings suggests that reducing IIS in the *Drosophila* fat body by $dilp^{2\text{-}3,5}$ knockout affects DNA maintenance mechanisms to maintain transposon repression and potentially delay genomic instability in ageing flies.

3.3.4 Post-transcriptional mechanisms of insulin signalling

3.3.4.1 Section introduction and overview

The majority of studies investigating the molecular mechanisms of reduced IIS have examined gene expression on only one level of gene expression, such as transcription (Murphy et al. 2003; Halaschek-Wiener et al. 2005; Oh et al. 2006; Teleman et al. 2008; Alic et al. 2011; Ewald et al. 2015), or more recently the proteome (Stout et al. 2013; Narayan et al. 2016). Although informative, studies that focus on one level of gene expression inevitably fail to take into account the complex interplay of different levels of regulation. Thus, these studies have not examined the considerable impact of post-transcriptional regulation (PTR) on the functional output of gene expression, i.e. protein expression (Cech & Steitz 2014b). Indeed, patterns of post-transcriptional regulation , such as reduced translation, are associated with increased longevity (Afschar et al. 2016; Kaletsky et al. 2016; Selman et al. 2009). The exact proportion of transcript- vs. protein-level control is itself heavily context dependent, with protein-level dynamics gaining in importance as the cell responds to stimuli or stresses (Liu & Aebersold 2016; Cheng et al. 2015; Brockmann et al. 2007).

One way post-transcriptional regulation in response to IIS may be mediated is through non-coding RNA, such as microRNA (miRNA) or long non-coding RNA (lncRNA). Mature miRNAs are short (22 nt) RNAs that can influence gene expression through targeted degradation of mRNA or may suppress translation of target mRNAs through complementary antisense binding (Lee et al. 1993). miRNAs can play major roles in metabolic signalling (Nolte-'t Hoen et al. 2015), including the control of insulin secretion and sensitivity (Poy et al. 2004; Trajkovski et al. 2011; Karolina et al. 2011) and other components of insulin or TOR signalling (Jordan et al. 2011; Yu et al. 2015). Another form of non-coding RNA, lncRNAs, can physically bind to proteins involved in the regulation of gene expression, such as transcription factors, histones, and ribosomes (Long et al. 2017). It has been hypothesized that lncRNA-ribosome interactions may affect translation, however their precise functions remain unclear (van Heesch et al. 2014; Carlevaro-Fita et al. 2016). lncRNA may also regulate transcription through chromatin modification (Saxena & Carninci 2011; Long et al. 2017), directly affect translation (Carrieri et al. 2012) or regulate the function of receptors in response to environmental cues such as nutrient availability (Kino et al. 2010).

Regulation of transcripts can further be affected through RNA editing: miRNA primary transcripts are processed to mature miRNA through Drosha and Dicer (Ha & Kim 2014). Activation of mTOR has been shown to suppress Drosha, thereby reducing miRNA biogenesis (Ye et al. 2015). RNA editing by adenosine deamination affects protein-coding transcripts, potentially altering splicing signals (Athanasiadis et al. 2004), and can also be under metabolic control (Gan et al. 2006). Stability of transcripts is another important entry point for regulation

(Guhaniyogi & Brewer 2001), which in turn can be subject to regulation by miRNA (Bagga et al. 2005). mRNA stability and translation are affected by polyadenylation (Proudfoot 2011), alternative polyadenylation (Elkon et al. 2013), or capping (Mukherjee et al. 2012). Insulin receptor (INR) mRNA stability depends on the cell cycle (Levy & Hug 1992) and the stability of insulin mRNA is modified through binding of *PTB* protein under the control of mTOR (Tillmar & Welsh 2002; Tillmar & Welsh 2004).

The conversion of mRNA into protein via translation is another major point of PTR. Translation itself is regulated at different levels at initiation, elongation, termination, and ribosome biogenesis (Sonenberg & Hinnebusch 2009). Translation initiation is the main determinant of translational rate and canonically proceeds by binding of the mRNA 5'-cap to the eIF4F initiation complex (cap-dependent initiation)(Sonenberg & Hinnebusch 2009). However, cap-independent mechanisms of regulating translation, for example via binding to the internal ribosome entry site (IRES), exist (Komar & Hatzoglou 2011). Decreased ribosomal protein, rRNA expression, and reduced nucleolar size are associated with longevity in fly and mouse models undergoing dietary restriction, and possibly in humans undergoing dietary interventions (Tiku et al. 2016). Recently, it has been shown that ribosomes can vary in their core protein composition, making them specific for subpools of mRNA and further contributing to translational regulation (Shi et al. 2017). It is known that IIS regulates cap-dependent translation through the phosphorylation of translation eukaryotic initiation factor 4B (eIF4B)(Shahbazian et al. 2006) and translation initiation factor 4E-binding protein (4E-BP), which sequesters eukaryotic initiation factor 4E form the translational machinery (Zid et al. 2009). The translation of IRES-containing transcripts such as INR may also be regulated in a cap-independent manner, circumventing the eIF4 complex (Marr et al. 2007). This level of regulation allows the cell to adjust protein synthesis in response to extrinsic and intrinsic stimuli, such as nutrient availability or IIS (Kaeberlein et al. 2005; Proud 2002; Cardenas et al. 1999). Furthermore, the response of reducing translation under low IIS is an evolutionarily conserved response (Essers et al. 2016).

Lastly, there is considerable regulation at the level of the proteome. Protein stability and turnover are important determinants of protein expression, and play a major role in PTR. Increased protein turnover may be beneficial where fast adaptation to stimuli is essential, e.g. in response to cellular stress (Hochstrasser & Varshavsky 1990; Lackner et al. 2012). On the other hand, protein levels may be directly modulated through transcription, particularly if stoichiometry of complexes needs to be preserved (Jüschke et al. 2013; G.-W. Li et al. 2014). Conversely, this implies that protein complexes may be efficiently regulated by varying individual members (de Lichtenberg et al. 2005). Beyond protein levels, post-translational modifications (PTM) are known to affect protein stability (Sadoul et al. 2008) as well as chromatin structure and transcription (Sterner & Berger 2000; Glozak et al. 2005). Improper folding or PTM of newly synthesised proteins may lead to their targeted degradation through specialized quality control pathways (Trombetta & Parodi 2003). Increased prevalence of protein

misfolding and damage is associated with age (Stadtman 2001; Kikis et al. 2010), and conversely, enhanced protein turnover associated with longevity (Lewis et al. 1985; Tavernarakis & Driscoll 2002). Reducing IIS/TOR signalling directly affects proteostasis (Schmidt et al. 1998; Bai et al. 2013; Tawo et al. 2017).

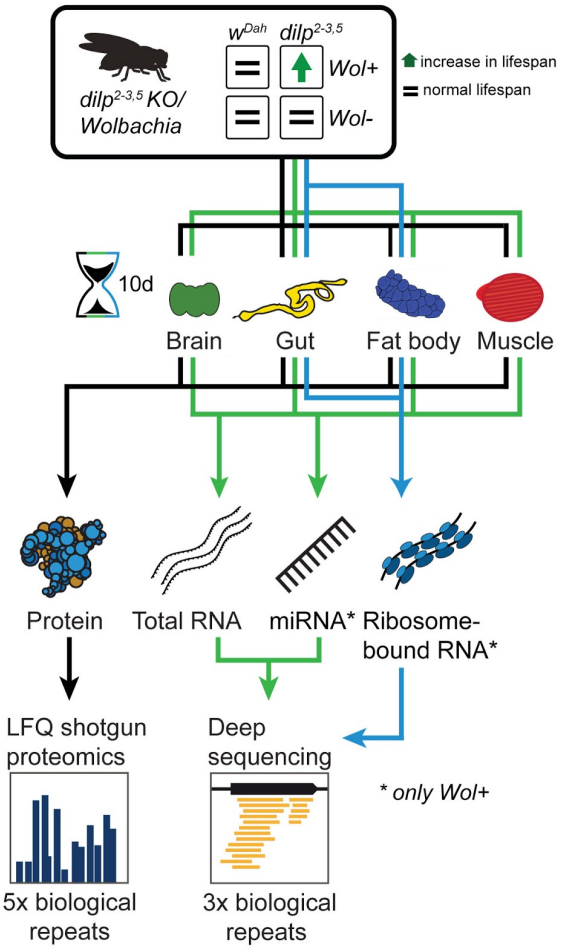


Fig 23**: Quantification of *dilp*$^{2-3,5}$ model proteomics and transcriptomics.**
Diagram showing the experimental design of the dilp$^{2-3,5}$ *screen. Muscle, gut, fat body, and brain samples were extracted from long-lived dilp*$^{2-3,5}$ *mutants and w*^{Dah} *controls. Protein levels were quantified using LFQ shotgun proteomics with 5 replicates per experimental group. Transcript levels were quantified using total RNA sequencing with 3 replicates per experimental group. Additionally, on gut and fat samples, RNA bound by two or more ribosomes was extracted and sequenced.*

Here, we investigate tissue-specific regulation of gene expression in a long lived IIS drosophila mutant concurrently on the level of the proteome and the transcriptome, in four key tissues: brain, gut, fat body, and muscle (Fig 23). We describe and functionally characterize post-transcriptional regulation in each tissue, including characteristic responses to reduced IIS at both the transcriptome and proteome level. We show that reduced translation in the *Drosophila* fat body (Essers et al. 2016) is the result of post-transcriptional down-regulation of membrane

and secretory proteins mediated by a reduction in co-translational protein import into the ER. Thus, reduced IIS modulates key functions of the fat body in metabolic signalling and immune response (Géminard et al. 2009). Furthermore, we identify an elevated DNA damage response that mediates a reduction of transposon burden in the fat body of aged insulin mutant flies.

3.3.4.2 Proteome and transcriptome are highly tissue-specific

To identify the tissue-specific gene expression profiles in *Drosophila* we characterized the transcriptome and proteome in wild type flies. Processing of $dilp^{2\text{-}3,5}$ model proteomics data is described in chapter 3.3.3.2; these data were re-used in this analysis. Additional samples from brain, fat body, gut, and thorax were dissected and divided for analysis of their RNA content. RNA quantification was performed by next-generation RNA-sequencing. To capture regulatory elements, specifically non-coding RNAs, our transcriptomic analysis was performed on rRNA depleted sample libraries. Additionally, we isolated and sequenced the polysome fractions of the two most translationally active tissues, the fat and the gut (Essers et al. 2016), in $dilp^{2\text{-}3,5}$ and w^{Dah} flies. The polysome fraction describes the part of the transcriptome that is bound by two or more ribosomes; its ratio to the total RNA is therefore a useful proxy of translational activity. Lastly, to capture and characterize small regulatory RNAs, we sequenced tissue-specific samples enriched for miRNA (Werner 2014).

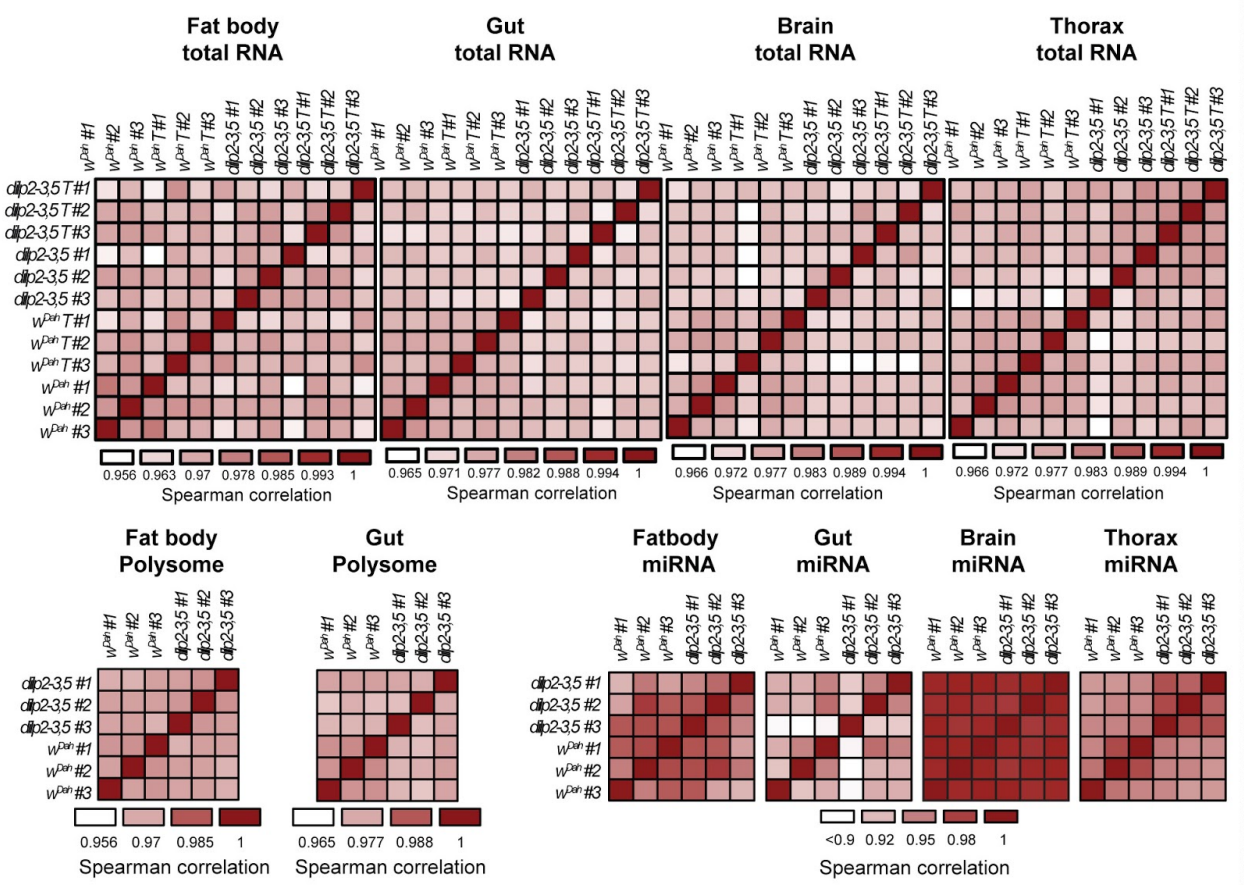


Fig 24: **Reproducibility of *dilp*$^{2\text{-}3,5}$ model transcriptomics.**
Correlation heatmaps of transcript quantification in all replicates and tissues, for total RNA, polysome fraction, and miRNA measurements. T indicates absence of Wolbachia.

After mapping the sequencing reads to the BDGP6 reference genome using tophat2, we processed the reads in R using the summarizeOverlaps function in the GenomicRanges package. This function is an R implementation of the widely used HTSeq framework (Anders et al. 2015). miRNA raw counts calculated from enriched samples were obtained directly from collaborators (Werner 2014). We then assessed the quality and reproducibility of our transcriptomics data. As RNAseq count data are not normally distributed, but rather follow a negative binomial distribution (Di et al. 2011), we calculated pairwise Spearman rank correlation coefficients which measure the monotonic association between variables. We found that tissue-specific transcriptomes and miRNA quantifications were highly correlated between all three biological replicates (Fig 24).

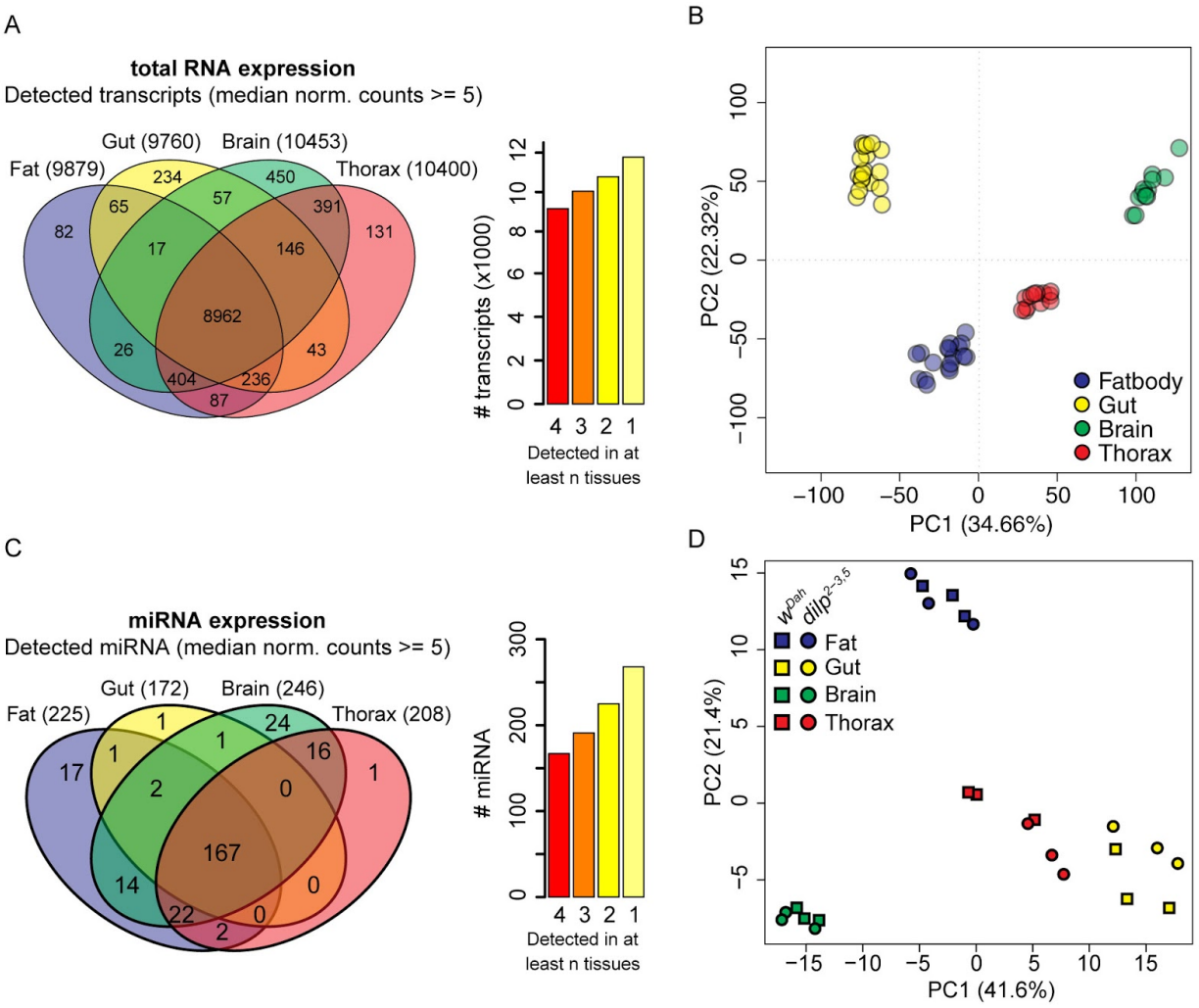


Fig 25***: Dilp$^{2-3,5}$ model transcriptomics and miRNA measurements are reproducible and tissue-specific.***
(A&C) Detected transcripts above a median of normalized counts of 5, from total RNA and miRNA quantifications. (B&D) PCA projections of experimental samples, using total RNA transcript expression and miRNA expression, respectively.

In total, we identified 11331 transcripts (median count >5), ranging from 9760 in the gut to 10453 in the brain (Fig 25A). Our analysis of miRNA-enriched samples identified detected between 172 and 246 known miRNA, most of these (246) in the brain (Fig 25C). To visualize the extent to which transcriptomes and regulatory elements such as miRNA vary from one tissue to another we performed principal component analysis. We could capture 57% of transcriptome variation (Fig 25B) and 63% of miRNA variation (Fig 25D) within the first two components of their PCA's, compared to 72% for the proteome (Fig 13). Intra-tissue variability was lower than variability between tissues, thus tissues cluster closely together based on both total RNA and miRNA expression (Fig 25B&D). However, the transcriptome showed a higher relative intra-tissue variability compared to the proteome (Fig 13). This may be due to the higher

dynamic range of next generation RNA-sequencing, allowing even low-abundant transcripts to be detected even if they are not functionally relevant in a given tissue. Mass-spectrometry based proteomics has yet to achieve the depth of coverage and dynamic range possible with RNAseq, despite dramatic progress in the last decade (Zubarev 2013). As a consequence, shotgun proteomics exhibits a bias for more highly abundant proteins.

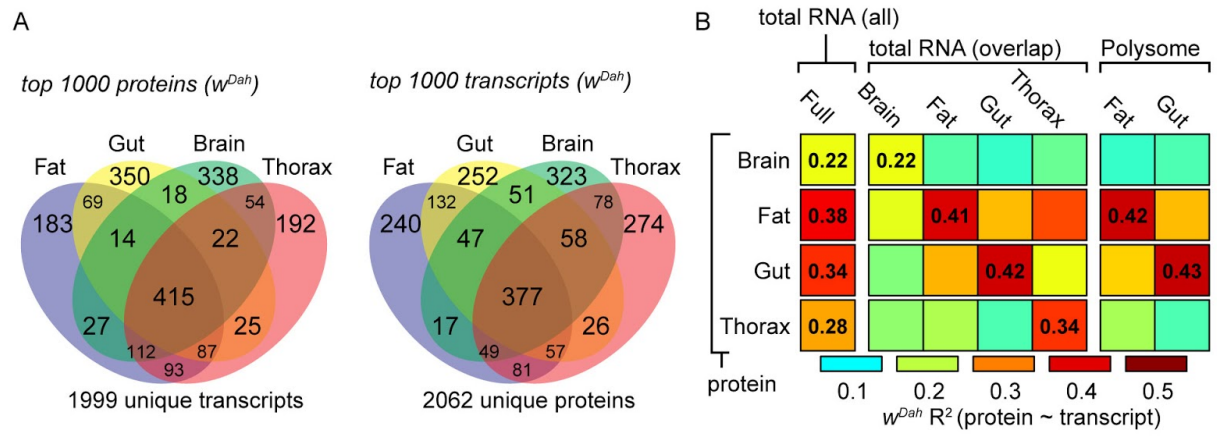


Fig 26***: Both proteome and transcriptomes are tissue-specific.***
(A) *Venn diagrams showing the 1000 most high-abundant proteins and transcripts, respectively, across w^{Dah} tissues. The total number of unique proteins or transcripts resulting from this selection is shown below the venn diagrams.* ***(B)*** *Heatmap showing R^2 values between mean protein expression and mean RPKM values of matching transcripts in w^{Dah}. First block shows correlation to all tissue-wise matching total RNA transcripts. Second block shows both intra- and inter-tissue correlations for 1545 genes that were identified on all levels and tissues (overlap). Third block shows correlations with transcript expression levels from polysome samples (for overlap genes), in fat and gut.*

To directly compare the tissue-specificity of the transcriptome to that of the proteome, we therefore took the 1000 highest expressed transcripts and the 1000 most abundant proteins in w^{Dah} samples of each tissue and visualized their overlap between tissues (Fig 26A). To determine if the proteome is more divergent between tissues than the transcriptome we performed a Fisher-test. We found no significant differences in the proportions of exclusive and shared proteins/transcripts (p-value = 0.4). Thus, the observed gene expression profiles are highly tissue-specific and are common to both transcript and protein levels.

To determine the agreement between protein and transcript expression, we first transformed transcript counts to RPKM values to account for gene length biases and library size differences in RNAseq (Mortazavi et al. 2008). We then calculated their pairwise Spearman rank correlations, squaring them to obtain R^2 values. Previous studies have shown considerable variation in the correlation between RNA and protein abundance, however, invariably the correlation between these two levels of gene expression is imperfect (Liu & Aebersold 2016). Our approach yielded R^2 values of 0.38 and 0.34 in fat and gut, 0.28 in the thorax, and 0.22 in the brain (Fig 26B). To determine the tissue-specificity of transcript-protein correlations, we also

calculated R^2 values for a set of 1545 genes that were identified in all tissues and on both levels of expression. These values were highest on the diagonal, confirming the tissue-specificity of transcriptome-proteome pairings (Fig 26B). Furthermore, R^2 values between protein and polysome transcript expression mirrored the values of the total RNA correlation (Fig 26B), suggesting that on a global level, polysome expression was not significantly more similar to protein expression. To side-step potential issues arising from label-free rather than absolute protein quantification, we focused on the comparison of relative gene expression changes between genotypes on the transcript and protein levels in the following analysis.

3.3.4.3 Reducing IIS modulates both the tissue-specific transcriptomic and proteomic landscapes

Reducing IIS remodels gene expression via downstream transcription factors (Kenyon 2010; Fontana et al. 2010; Slack et al. 2015). In chapter 3.3.3, we assessed the proteome response to *dilp*$^{2\text{-}3,5}$ knockout separately and compared it to the insulin response of mNSC-ablated flies. Here, we integrated our analysis of the *dilp*$^{2\text{-}3,5}$ model proteome response with an analysis of the reduced insulin response observed on the transcriptome level, to distinguish transcriptionally and post-transcriptionally regulated responses to reduced IIS.

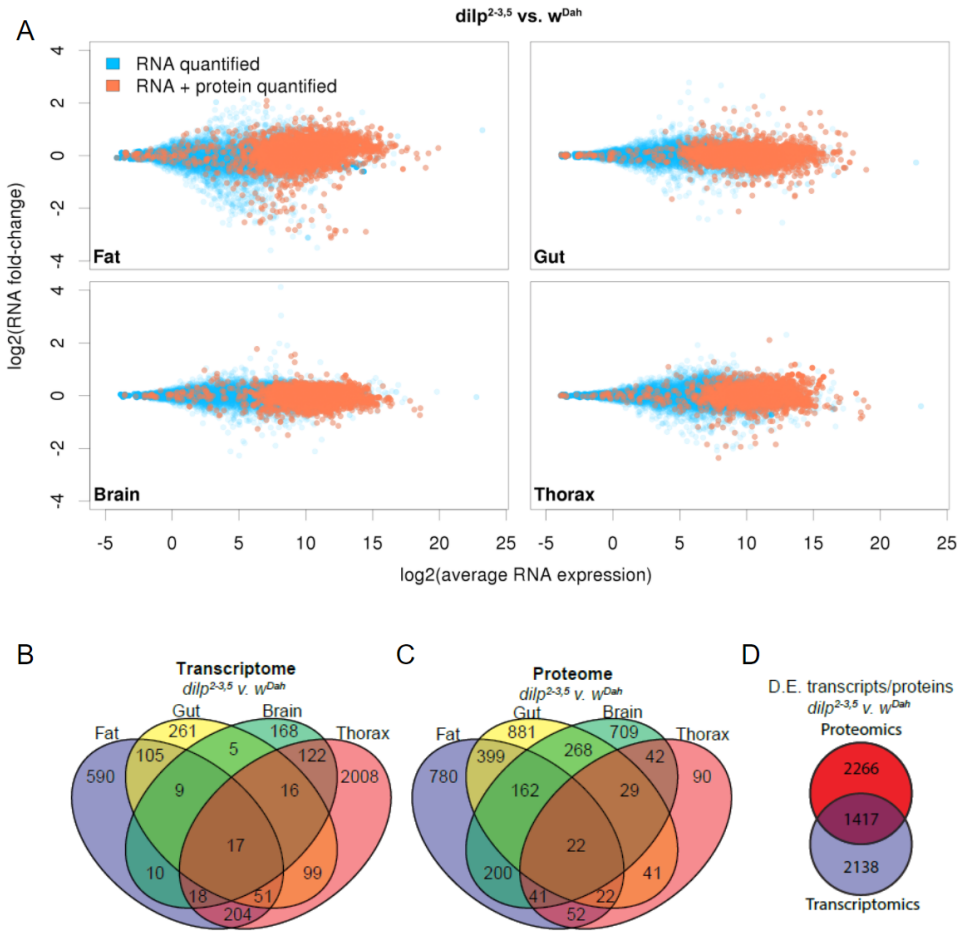


Fig 27**: Differential expression analysis of the *dilp*$^{2\text{-}3,5}$ transcriptome and proteome.**
(A) *MA-plots showing the log-fold-changes of transcripts depending on their average log-transformed expression levels. Blue dots indicate that only the transcript was measured, whereas orange dots indicate that both transcript and protein product were quantified.* ***(B)*** *Venn diagram showing numbers of transcripts differentially expressed between dilp*$^{2\text{-}3,5}$ *and w*^{Dah} *in each tissue.* ***(C)*** *Venn diagram showing numbers of proteins differentially expressed between dilp*$^{2\text{-}3,5}$ *and w*^{Dah} *in each tissue.* ***(D)*** *Venn diagram showing the total numbers of proteins and transcripts found to be differentially expressed between dilp*$^{2\text{-}3,5}$ *and w*^{Dah} *across all tissues, as well as their overlap.*

For this, we first implemented the same linear model as used in the analysis of our proteomics data (Eq 6), fitting transcript expression data across all experimental groups. After correcting for multiplicity using the BH method, we found a total of 3683 transcripts significantly regulated in at least one tissue at a significance level of $\alpha = 0.1$ (Fig 27)(File 1). Mapping transcripts to their protein products, we observed that proteomic measurements were biased towards products of more abundant genes (Fig 27A)(File 1). Also, we detected a greater number of differentially expressed proteins compared to differentially expressed transcripts in all tissues except the thorax, despite detecting lower numbers of proteins versus transcripts in each tissue (Fig 27 B&C). This may be due to a combination of technical factors, including differences in the

number of replicates, differences in precision between the quantification methods, and differences in the magnitude of changes of proteins compared to transcripts.

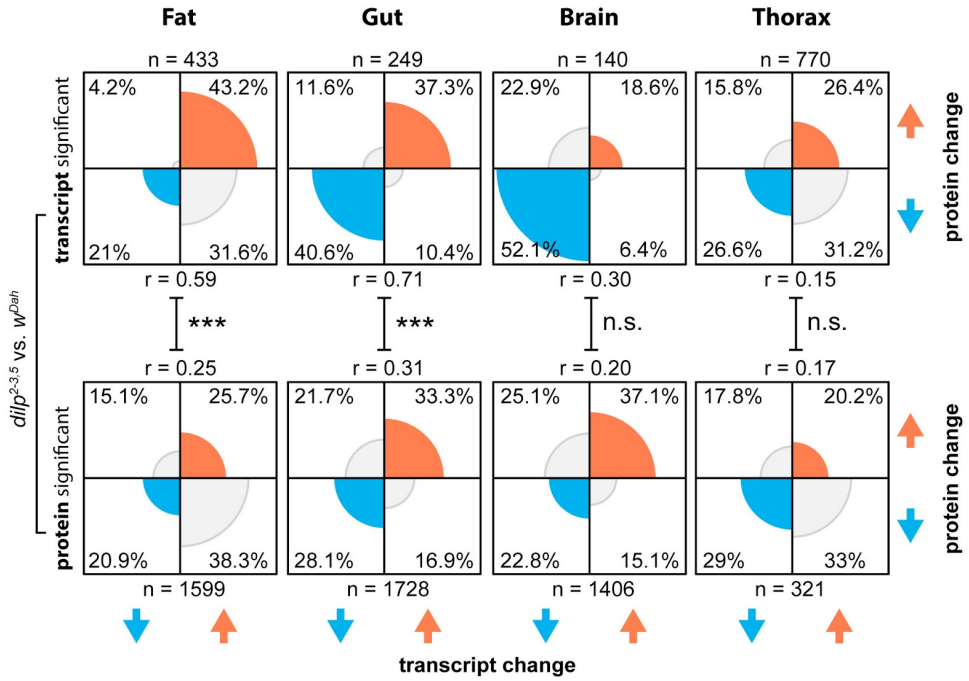


Fig 28: ***Coherence of the transcriptome and proteome IIS responses.***
*Radial bar charts showing the proportion of protein/transcript pairs that are concordantly regulated (both up or down) or oppositely regulated, in each of the four tissues (columns) and two gene subset (rows). Arrows show direction of regulation (up/down) of proteins (y-axis) and transcripts (x-axis). N-numbers for each plot show the number of protein/transcript pairs that were selected for that plot. Pearson correlations (r values) were calculated between the protein and transcript log-fold-changes of protein/transcript pairs in each plot. Asterisks indicate the significance of the difference between the indicated correlation coefficients (***P-value ≤ 0.001, n.s. not significant). First row: all protein/transcript pairs in the respective tissue where the transcript is independently significantly regulated (adj. P-value ≤ 0.1). Second row: all protein/transcript pairs in the respective tissue where the protein is independently significantly regulated (adj. P-value ≤ 0.1).*

We then asked whether changes that were first detected on the level of the transcriptome were more likely to be changed in the same direction on the protein level. Indeed, when compared to all protein/transcript pairs where changes were detected on the protein level, transcript-induced protein/transcript pairs were less likely to be post-transcriptionally regulated (Fig 28). This was especially apparent in the gut, where 78% of transcript-level changes evoked a protein response in the same direction, compared to 61% of protein-level changes that had a matching transcript response (Fig 28). In the thorax, no clear pattern of co-regulation could be identified under either filtering scheme (Fig 28).

$$Z_{diff} = \frac{\operatorname{atanh}(r_1) - \operatorname{atanh}(r_2)}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \quad (8)$$

Additionally, we calculated correlations between protein-level and transcript-level changes for all subsets (each tissue, detected on the transcriptome or proteome level)(Fig 28). In all tissues except the thorax, the correlation coefficient was higher for changes that originated from the transcriptome (Fig 28). In fat and gut, these differences were largest (Fig 28), suggesting a higher degree of PTR. In the brain, correlation coefficients were comparatively low, possibly due to increased prevalence of alternative splicing in this tissue (Mohr & Hartmann 2014). To determine the significance of differences between correlation coefficients within each tissue, we transformed the coefficients into z-values using the Fisher r-to-z transformation (also known as the *atanh* function)(Fisher 1915). The observed value of the normalized difference between correlations, Z_{diff}, was calculated, taking into account different sample sizes (Eq 8). We found that the observed differences in correlation coefficients within fat and gut were highly significant, but not those in brain and thorax (Fig 28). This suggests that fat and gut were more strongly affected by PTR.

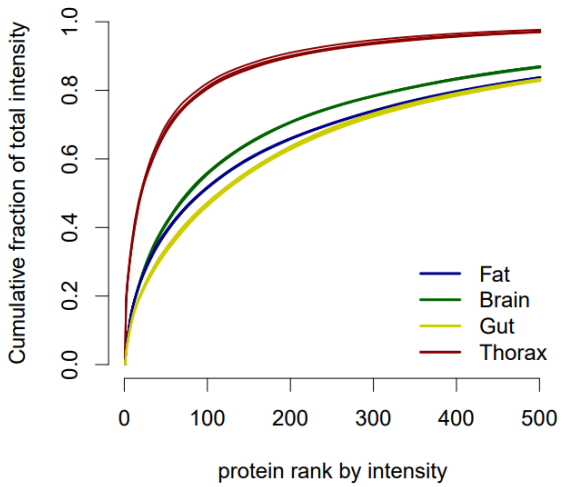


Fig 29: ***Proteomics dynamic range saturates faster in the thorax than in other tissues.***
Line chart showing the cumulative fraction of total protein intensity values in all w^{Dah} replicates, along the 500 most abundant proteins in each tissue.

Tissue composition may be a confounding variable contributing to the low agreement between protein and transcript levels in the thorax: the presence of abundant structural proteins can bias the performance of mass spectrometry measurements, and reduce dynamic range (Zubarev 2013). Indeed, in the thorax, 50% of the total intensity detected was represented by only 21 proteins, compared to between 75 and 114 in the remaining tissues (Fig 29). The most highly expressed protein in the thorax, Myosin heavy chain (Mhc), a major structural protein in muscle, was accounting for nearly 19% of the detected total intensity, followed by Act88F (3.1%). Thus

the poor correlation between significant protein/transcript pairs in the thorax may reflect technical challenges rather than biological regulation.

To filter transcript/protein pairs for significance, we calculated a single combined p-value using the Stouffer method (Whitlock 2005; Stouffer et al. 1949)(Fig 30). The combined p-value is lower for protein/transcript pairs that significant change at both the protein and transcript levels, and more conservative for those regulated on only one level. We intentionally did not penalise transcript/protein pairs that were regulated in an opposing manner, as these might represent some of the strongest examples of post-transcriptional regulation. We then carried out GO enrichment analysis using SETHRO, characterizing the four quadrants of the gene expression response to reduced IIS in each tissue: increase of both protein and transcript (I), decrease of both (III), increase on the protein level but decrease on the transcript level (II), and decrease on the protein level but increase on the transcript level (IV).

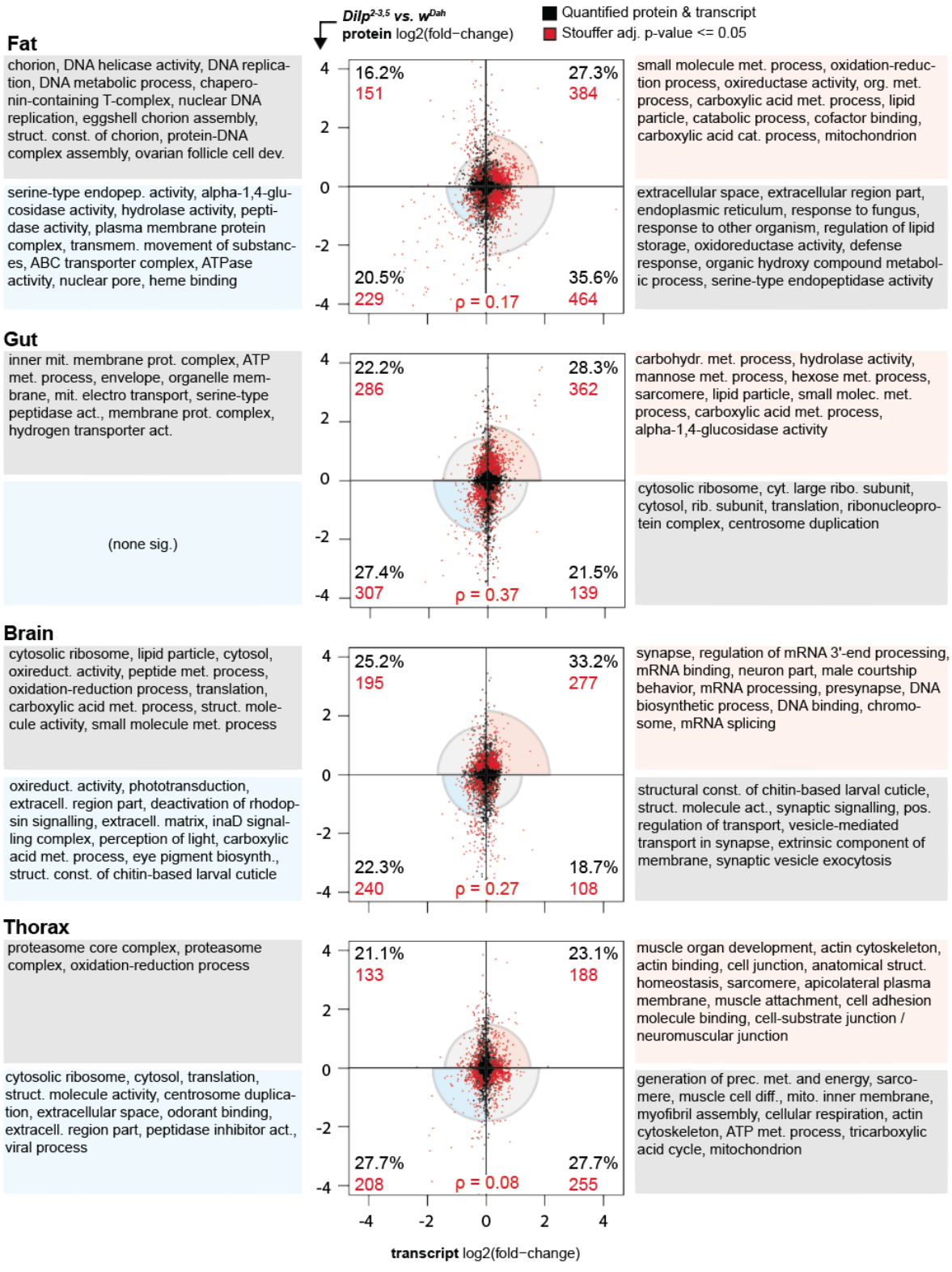


Fig 30**: Post-transcriptional regulation in response to reduced IIS.**

Scatter plots showing differential regulation between dilp$^{2\text{-}3,5}$ and w^{Dah} on the protein level (x-axis) and the transcript level (y-axis). Black dots indicate all fully quantified protein/transcript pairs. Red dots indicate protein/transcript

pairs with a significant combined adjusted p-value ($p \leq 0.05$). Percentages and circle segments in each quadrant indicate the fraction of protein/transcript pairs in that quadrant. Red numbers in each quadrant indicate the number of significant protein/transcript pairs. Values on the bottom of each scatter plot show the Spearman rank correlation of significant protein/transcript pairs in the respective tissue. Annotations beside each quadrant represent up to the 10 most significantly enriched GO terms based on significant protein/transcript in that quadrant (compared to the tissue-specific background)(File 1).

In the thorax, 188 and 208 protein/transcript pairs were concordantly up-regulated or down-regulated respectively (Fig 30). A total of 388 protein/transcript pairs were regulated in opposite directions (Fig 30). GO enrichment analysis of these significant protein/transcript pairs showed that muscle-related functions such as muscle attachment, M band, and sarcomere were enriched among up-regulated transcripts, regardless of the direction of protein regulation (Fig 30). Although the changes appeared uncoordinated on the protein level, age-related deterioration of muscle function occurs in Drosophila and other model organisms (Demontis & Perrimon 2010). We have previously observed an increased climbing ability of $dilp^{2\text{-}3,5}$ mutants compared to controls (Fig 17), suggesting that IIS counteracts age-related loss of muscle function.

In the brain, 517 protein/transcript pairs were regulated in the same direction, 240 down- and 277 up-regulated, respectively. Down-regulated protein/transcript pairs were enriched for the inaD signaling complex, phototransduction and rhodopsin signaling (Fig 30). However, a large fraction (33%) of protein/transcript pairs in the brain were concordantly up-regulated and enriched for GO terms associated with RNA processing (Fig 30). One major form of RNA processing, RNA splicing homeostasis, can influence lifespan in model organisms *C. elegans* (Heintz et al. 2016), and wide-reaching changes in splicing have been reported in ageing brains of humans (Mazin et al. 2013). Regulation of RNA processing may therefore be a conserved mechanism of longevity assurance in response to IIS or other longevity interventions such as DR, and may be of particular importance in neuronal tissues.

In the gut, 362 protein/transcript pairs were up-regulated, and were enriched for glucose and mannose metabolism (Fig 30). This suggests a gut-specific alteration of carbohydrate metabolism in response to reduced IIS. Although 307 protein/transcript pairs were down regulated in the gut in response to reduced IIS, they were not significantly enriched for any specific GO terms (Fig 30). Interestingly, post-transcriptionally down-regulated proteins were enriched in terms like translation and cytosolic ribosome (Fig 30), suggesting a reduction in translation in long-lived $dilp^{2\text{-}3,5}$ flies.

In the fat body, 384 protein/transcript pairs were up-regulated and enriched for genes associated with mitochondrial and lipid metabolism processes (Figure 2). In contrast, 229 protein/transcript pairs were concordantly down-regulated and enriched for genes encoding

membrane proteins, transporter complexes, as well as secreted serine-type endopeptidases (Fig 30). However, in contrast to the thorax, gut, and brain, an unusually high proportion of protein/transcript pairs in the fat body were regulated in opposite directions (Fig 30). Up to 50% of gene expression (615 protein/transcript pairs) may thus be post-transcriptionally regulated, indicating that PTR upon reduced IIS is more prevalent in the fat body than in other tissues. Post-transcriptionally up-regulated proteins were enriched for functions such as DNA replication/repair, cytosolic protein folding, and chorion-related functions (Fig 30). Interestingly, the majority of the 615 post-transcriptionally regulated protein/transcript pairs had reduced protein levels despite actively upregulated transcription (Fig 28) and were enriched for functions relating to the extracellular space and the endoplasmic reticulum (Fig 30). The role of ER in PTR is vital, with the ER acting as a major cellular compartment for translation (Reid & Nicchitta 2012; Reid & Nicchitta 2015). As such, our analysis suggested that changes translation in response to reduced IIS in the fat body (Fig 8) may be particular to proteins processed in the ER.

3.3.4.4 Translation modulates gene expression in the fat body in response to reduced IIS

Translation is a highly regulated process allowing cells to post-transcriptionally fine tune gene expression and thus rapidly respond to stimuli (Sonenberg & Hinnebusch 2009). We previously characterized *dfoxo*-dependent decrease of translation rate in the fat body (Fig 7)(Tain et al. 2017). To gain additional insight into the mechanisms involved, we examined how altered translation may influence the tissue-specific gene expression landscape in response to reduced IIS in the *dilp*$^{2\text{-}3,5}$ model.

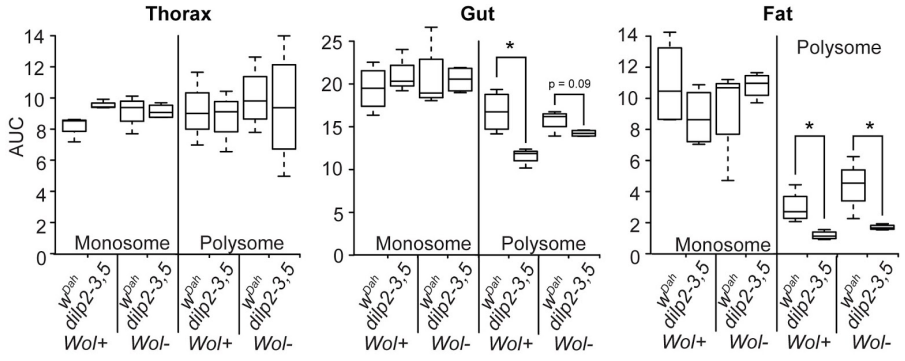


Fig 31*: **Changes in translation potential in response to reduced IIS.***
*Results and plots by Luke Tain. Box plots showing area under the curve from polysome profiles in thorax, fat, and gut, for the monosome fraction (transcripts with one or fewer bound ribosomes) and the polysome fraction (transcripts with two or more bound ribosomes). Error bars show minimum and maximum values. Asterisks indicate significant differences between groups, as determined using a t-test (*P-Value ≤ 0.05).*

We first quantified the level of tissue-specific translation using polysome profiling. Polysome profiling allows the quantitative visualization of mRNA-ribosome complexes as a proxy for translational activity (Arava et al. 2003). High levels of mRNA bound to multiple ribosomes, i.e. polysomes, is indicative of high translation, while a relative increase in mRNA bound to a single ribosome (the monosome fraction) indicates the opposite (Van Der Kelen et al. 2009). With the exception of the brain, we characterized clear, tissue-specific polysome profiles in both controls and IIS mutants. No significant differences in the monosome fraction were found for any tissue (Fig 31). However, the polysome fractions were decreased in long-lived *dilp*$^{2\text{-}3,5}$ flies in both gut and fat body (Fig 31). In the gut, this was in part conditional on presence of *Wolbachia* (Fig 31). This suggests that following reduced IIS signalling, the translation potential in specific tissues is lowered.

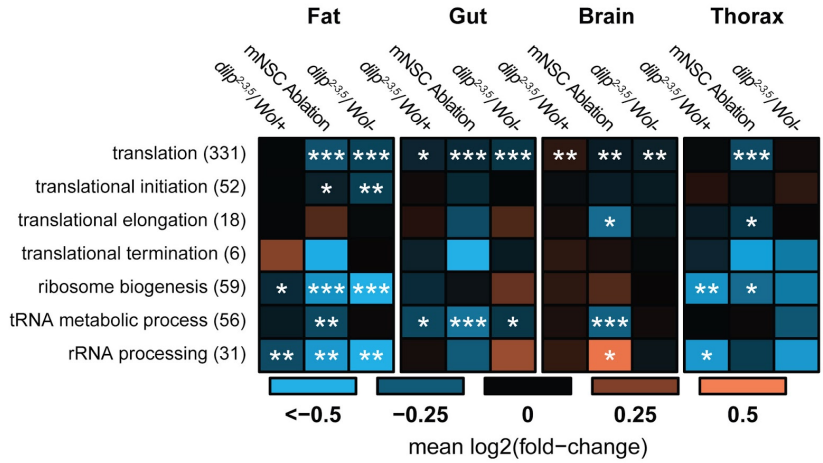


Fig 32***: Tissue-specific regulation of the translation machinery.***
*Mean of protein log-fold-changes of proteins annotated with selected translation-related GO terms, in the contrasts dilp2-3,5 vs. wDah (dilp2-3,5/Wol+) and InsP3-Gal4/UAS-rpr vs. wDah (mNSC Ablation), as well as the Wolbachia-negative control dilp2-3,5/Wol- vs. wDah/Wol- (dilp2-3,5/Wol-). All means were tested for significant difference from zero using a two-sided t-test (*p < 0.05, **p < 0.01, ***p ≤ 0.001).*

To determine if this was due to a down-regulation of translation machinery proteins we examined tissue-specific changes in their expression levels in response to reduced IIS. We compared changes observed in *dilp*$^{2\text{-}3,5}$ mutants, *dilp*$^{2\text{-}3,5}$*/Wol-* controls, and the proteome profiles of mNSC-ablated flies (Tain et al. 2017). Translation associated proteins were down regulated in all tissues of long lived flies, with the exception of the brain (Fig 32). In the gut, a significant down-regulation of genes associated with translation and tRNA metabolism was detected in all comparisons, and most strongly in mNSC ablated flies (Fig 32). This might indicate reduced activity of the tRNA metabolism acting through RNA polymerase III (RNApol III). Reduction of RNApol III in fly gut intestinal stem cells has recently been shown to extend lifespan in response to a reduction in TORC1 signalling (Filer et al. 2017). On the other hand,

ribosome biogenesis and rRNA processing were not significantly affected in the gut (Fig 32). The fat body and the thorax displayed consistent down-regulation of genes associated to ribosome biogenesis across both models of reduced IIS, along with genes associated to rRNA processing in *dilp*$^{2\text{-}3,5}$ flies (Fig 32). However, in the fat body a significant trend towards down-regulation was also observed in normal lived controls, while translation associated proteins showed coordinated regulation only in mNSC-ablated flies and *Wolbachia*-negative controls (Fig 32). This trend affected proteins involved in the rate limiting step of translational, initiation, but not translational elongation or termination (Fig 32).

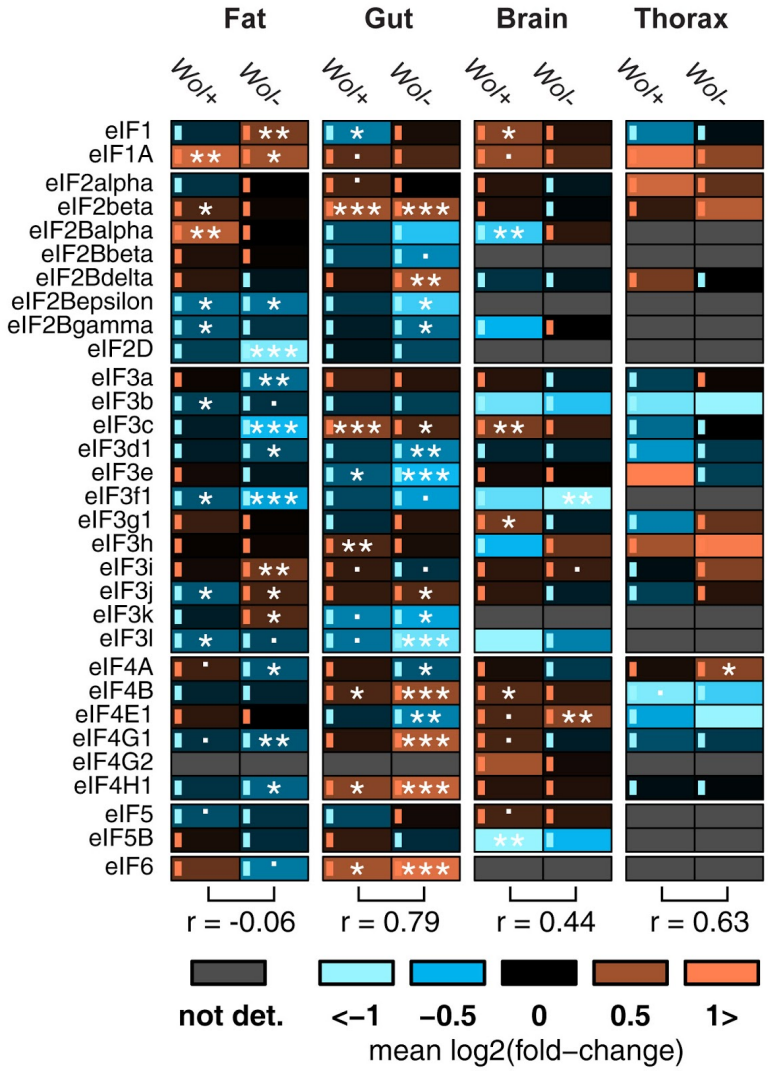


Fig 33***: Tissue-specific regulation of eIF subunits in response to reduced IIS and Wolbachia.***
Heatmap showing the regulation of all detected eukaryotic initiation factors (eIFs), in dilp$^{2\text{-}3,5}$ *mutants vs. w*^{Dah} *controls. Log-fold-changes and significances in Wolbachia-positive (long-lived)(Wol+) and Wolbachia-negative control (Wol-) backgrounds are shown. Inset box colours indicate positive (orange) or negative (blue) log-fold-change. Proteins not detected (not det.) in the respective tissue and background are marked in grey. Dots and asterisks indicate BH-corrected significance of the limma moderated t-test (.p ≤ 0.1, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001). Pearson correlations (r) between Wol+ and Wol- log-fold-changes are shown for each tissue.*

However, changes in the expression of single regulators of translation can affect global protein levels, e.g. through modulating the rate of translation (Essers et al. 2016; Jovanovic et al. 2015; Brockmann et al. 2007). In general, reduction of single components of protein complexes may affect total complex assembly and thus reduce overall activity (de Lichtenberg et al. 2005). Eukaryotic initiation factor (eIF) complexes were highly regulated upon reduced IIS in both fat and gut, but to a much lesser extend in brain and thorax (Fig 33). We found that eiF2B epsilon and gamma, the two core units of the eIF2B complex crucial for complex assembly (Gordiyenko et al. 2014), were reduced in the fat body of long-lived flies, but not in other tissues (Fig 33). We also observed changes upon reduced IIS in multiple subunits of eIF3 in both fat and gut (Fig 33). However, eiF3 diverged more strongly between *Wolbachia*-negative and -positive responses in the fat body; as a consequence, eIF correlation in the fat was the lowest observed (Fig 33). eiF3 serves as a scaffold for ribosomes and their bound mRNAs during translation, when it associates with other eiF's such as *eIF1A* and *eIF5* to stabilize the 43S and 48S pre-initiation complexes (Aitken et al. 2016). Mutations in single components of eIF3 can confer longevity and resistance to ER stress (Cattie et al. 2016) or impact the ability of the cell to translate sub-pools of mRNA (Shah et al. 2016).

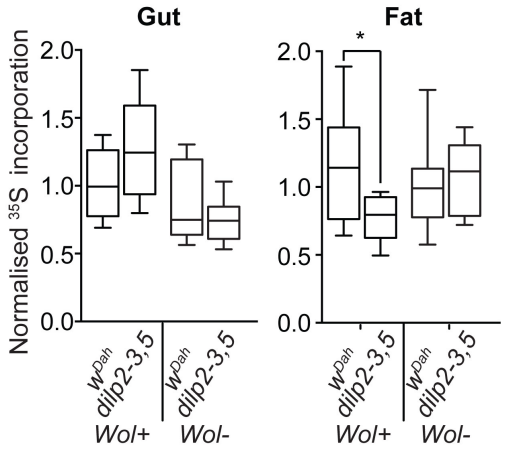


Fig 34*: **Altered translation rates in fat and gut as a result of reduced IIS.***
*Results and plots by Luke Tain. De novo protein synthesis as measured by incorporation of ^{35}S into proteins from ex vivo fat body and gut tissue. Quantifications of gel exposures normalized to total protein. Error bars show minimum to maximum values. Asterisks indicate significant differences between groups, as determined using a two-way ANOVA with Sidak's post-hoc test (*P-Value $\leq$ 0.05).*

To assess whether changes in the translation machinery observed in fat and gut had a tangible functional impact in *dilp*$^{2\text{-}3,5}$ mutants, we directly measured the ex-vivo tissue-specific rate of translation via ^{35}S incorporation analysis (Fig 34). The 35S assay quantifies the incorporation of ^{35}S -methionine into newly synthesized proteins as they are translated within a specific time period. We found that Incorporation of ^{35}S methionine and thus the rate of translation, was

unaltered in the gut in response to reduced IIS (Fig 34). However, the level of incorporation was significantly reduced in the fat body of long-lived *dilp*$^{2\text{-}3,5}$ mutants compared to both *w*^{Dah} and *Wolbachia*-negative *dilp*$^{2\text{-}3,5}$ controls (Fig 34). We previously found that translation rate was reduced in the fat body of long-lived mNSC-ablated flies in a *dfoxo*-dependent manner (Fig 8)(Tain et al. 2017). Together, this suggests that different tissues may regulate translation via different mechanism in response to reduced IIS. We decided to further investigate the changes we observed in the fat body.

3.3.4.5 The transcriptional response to reduced IIS is dampened on the level of translation

To determine if the reduction in polysome associated mRNA (Fig 31) was due to a global reduction in translation or associated to a specific subset of mRNA we performed sequencing on all mRNAs bound by two or more ribosomes (the polysome fraction). We sequenced the polysome fraction in fat and gut, in *dilp*$^{2\text{-}3,5}$ flies and their *w*^{Dah} controls, and compared transcript expression to that detected in our total RNA data. For this, we used DESeq2 to fit a linear model with an interaction term between RNA type (total or polysome RNA) and genotype (*dilp*$^{2\text{-}3,5}$ or *w*^{Dah})(Eq 9).

$$PolysomeInteraction = RNAtype{:}Genotype + RNAtype + Genotype \quad (9)$$

A positive interaction coefficient indicates a higher enrichment of the given transcript in the polysome fraction of *dilp*$^{2\text{-}3,5}$ mutants, and thus higher potential translation, while a negative coefficient implies the opposite.

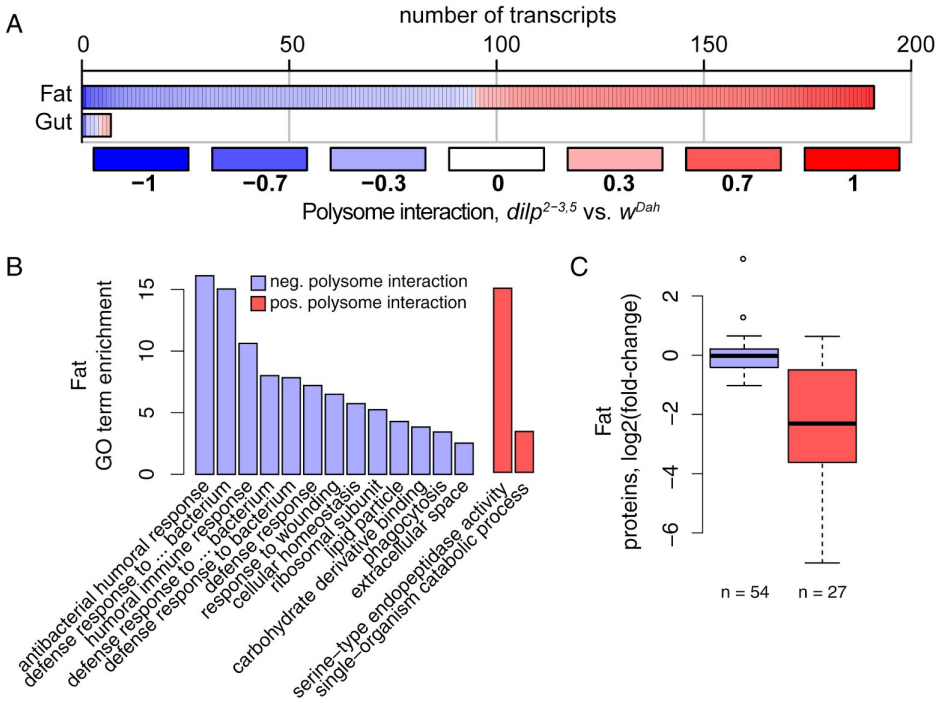


Fig 35***: Differential enrichment of transcripts in the polysome fraction in upon reduced IIS.***
(A) *Cell plot showing number of transcripts in fat and gut that are differentially enriched in the polysome fraction between dilp$^{2\text{-}3,5}$ and w^{Dah}. Colours of individual cells represent the polysome interaction coefficient of each transcript – positive coefficients indicate that a higher proportion of transcripts were found in the polysome fraction in dilp$^{2\text{-}3,5}$ mutants compared to controls, while negative coefficients indicate the opposite.* ***(B)*** *Bar plot showing significantly enriched GO terms (Fisher test with clustered terms, adj. P-value ≤ 0.05), for transcripts with a negative polysome interaction coefficient (blue) or positive polysome interaction coefficient (red) in the fat body.* ***(C)*** *Box plot showing the log-fold-changes of proteins (dilp$^{2\text{-}3,5}$ vs. w^{Dah}) that correspond to transcripts with a significantly up-regulated (red) or down-regulated (blue) polysome interaction coefficient. Below each box is the number of mapped proteins.*

In the gut, 7 transcripts had significant polysome interaction coefficients (Fig 35A). Two of these, *bic* and *Nacalpha*, encode members of the nascent polypeptide associated complex (NAC), a cytosolic ribosome-bound chaperone and regulator of proteostasis that couples protein folding status to translation (Kirstein-Miles et al. 2013). NAC can respond to proteotoxic stress by sequestration from the ribosome, leading to a decrease in translational capacity, and does so during ageing (Kirstein-Miles et al. 2013). Another, *RpL37a*, is a member of the 60S ribosomal subunit. All three transcripts were depleted in the *dilp$^{2\text{-}3,5}$* polysome fraction, but the total RNA transcript expression of *bic* and *Nacalpha* and the protein expression of all three genes were significantly increased in long-lived flies. As reported before (Tain et al. 2017), we observed an improvement of proteostasis in the guts of long-lived flies. A targeted binding assay may be able to determine if NAC ribosome association is involved in mediating this improvement.

In the fat body, we detected 191 transcripts with significant polysome interaction coefficients (Fig 35A). After splitting the response into genes whose ratios were either increased or decreased, we performed GO enrichment analysis using SETHRO (appendix). Genes showing a decrease in polysome ratio were much more likely to be involved in the humoral immune response, the wound response, or associated to the cytosolic ribosome, than expected by chance (Fig 35B). We detected a protein product for 27 of these transcripts, which strongly down-regulated on average (Fig 35C). Transcripts that showed an increase in their polysome ratio were highly enriched for serine-type endopeptidase activity (Fig 35). The 54 proteins corresponding to these transcripts showed no coordinated regulation (Fig 35C).

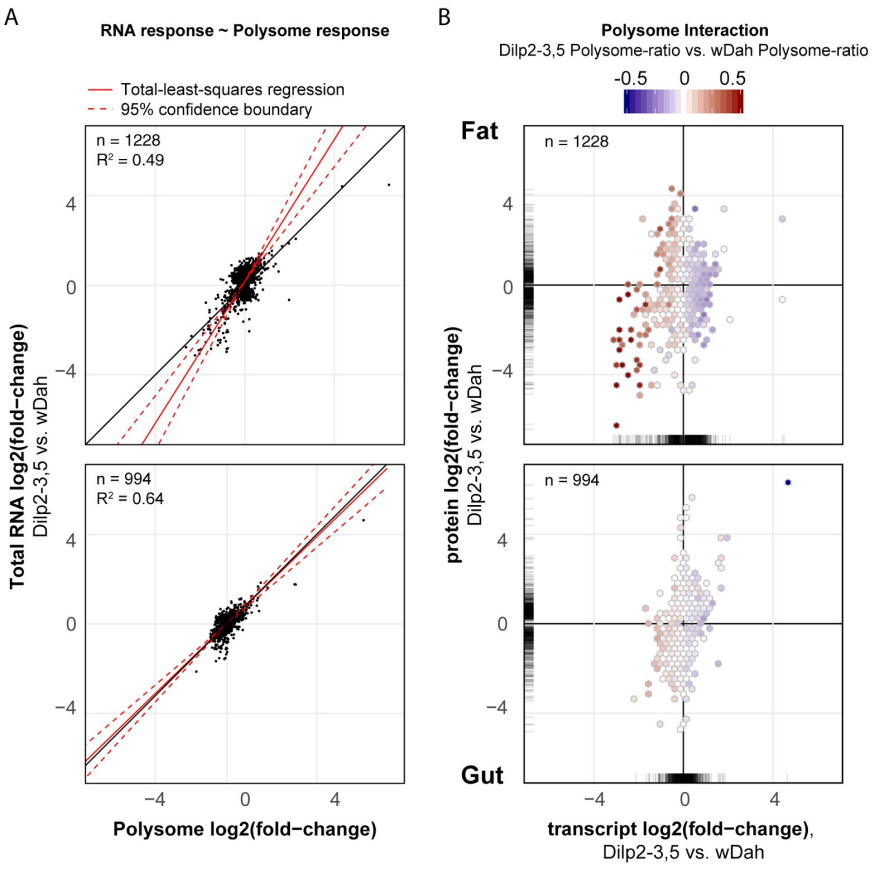


Fig 36***: Global bias of the polysome response to reduced IIS dampens the transcriptional response.***
(A) *Scatterplot showing the results of total-least-squares (TLS) regression analysis of the dilp$^{2\text{-}3,5}$ vs. w^{Dah} total RNA and polysome fraction responses, in fat and gut tissues. Only transcripts whose protein/transcript pair combined adj. P-value was ≤0.05 are shown. Black line shows the diagonal. Red line shows maximum likelihood TLS regression slope. Dotted red lines indicate the lower and upper bounds of the 95% confidence interval of the slope. R^2 values based on Pearson correlation are shown for each plot.* ***(B)*** *Stacked hexagonal scatterplots, showing transcript and protein log-fold-changes (dilp$^{2\text{-}3,5}$ vs. w^{Dah}) of significant protein/transcript pairs (combined adj. P-Value ≤0.05). Colours show the mean polysome interaction coefficient of points falling into each hexagonal bin.*

To determine if there was a global relationship between polysome level changes and total RNA changes between *dilp*$^{2\text{-}3,5}$ and *w*^{Dah}, we performed regression analysis of the total RNA change against the polysome RNA change upon reduced IIS. For this, we used total-least-squares (TLS) regression as implemented in the deming R package. Ordinary least squares regression assumes no error in the independent variable, leading to different estimates if independent and dependent variable are exchanged. In contrast, TLS explicitly models error terms in both dependent and independent variables (Deming 1943). In the gut, the observed TLS slope was not significantly different from the diagonal (Fig 36A). However, the TLS slope observed in the fat body was significantly higher than one (95% confidence interval = [1.28,1.89], Fig 36A), indicating that the polysome level response to reduced IIS was dampened compared to the total RNA response. This contradicts findings in yeast that translational regulation predominantly amplifies transcriptional changes (Albert et al. 2014; Muzzey et al. 2014). However, others have suggested that ribosome occupancy is more conserved than mRNA abundance between diverging yeast strains (McManus et al. 2014; Artieri & Fraser 2014; Wang et al. 2015). Consistent with this, a study in human cell lines from different individuals found that RNA-level variability between individuals is higher on average than variability in ribosome occupancy (Cenik et al. 2015). Translational adaptation may therefore be protective, dampening the influence of transcriptional perturbations.

3.3.4.6 Reduction of ER-specific translation in the fat body is a robust response to reduced IIS

To examine the mechanisms underlying the fat body-specific reduction in translation of IIS mutants we further interrogated our bioinformatic analysis. Another known element of the post-transcriptional landscape is the differential translation of cytosolic proteins versus those processed in the ER, i.e. secreted or membrane proteins (Reid & Nicchitta 2012). We observed an enrichment of ER-associated and secreted proteins among protein/transcript pairs that were post-transcriptionally down-regulated on the protein level (Fig 30). Proper ER function is required for intercellular signalling, and ER stress is known to be associated with a wide range of diseases such as diabetes and Alzheimer's disease (Marhfour et al. 2012; Engin et al. 2013; Santos & Ferreira 2017).

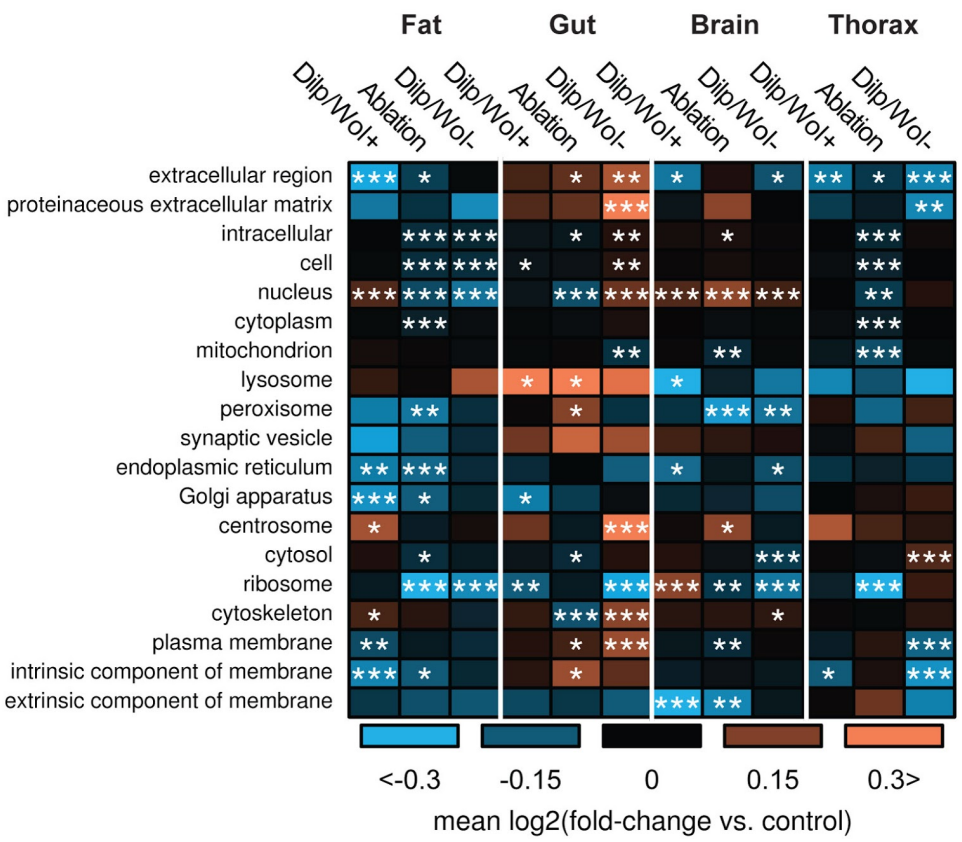


Fig 37**: Compartment-specific biases in the insulin response.**
Heatmap showing the average log-fold-changes of proteins annotated with a standard selection of GO compartment terms. Asterisks indicate significance of a two-sided t-test comparing the log-fold-changes against zero.

To determine if this response to reduced IIS was specific to ER-targeting we calculated average log-fold-changes between *dilp*$^{2\text{-}3,5}$ mutants and *w*^{Dah} controls for several cellular compartment terms, and compared them to ER-specific changes (Fig 37). To ensure that we detected robust responses to reduces IIS, we also calculated log-fold-changes between proteins of long-lived mNSC-ablated flies and their controls. If the response to reduced IIS was specific to ER-targeting we would expect regulation of proteins associated to the ER and Golgi body, along with proteins that undergo co-translational localization in general, i.e. secretory and membrane associated proteins. Indeed, ER and Golgi associated proteins were consistently down-regulated in the fat body of long-lived *dilp*$^{2\text{-}3,5}$ flies, and only in the presence of *Wolbachia* (Fig 37). This was also true for ER-targeted proteins such as intrinsic membrane proteins and extracellular space proteins (Fig 37).

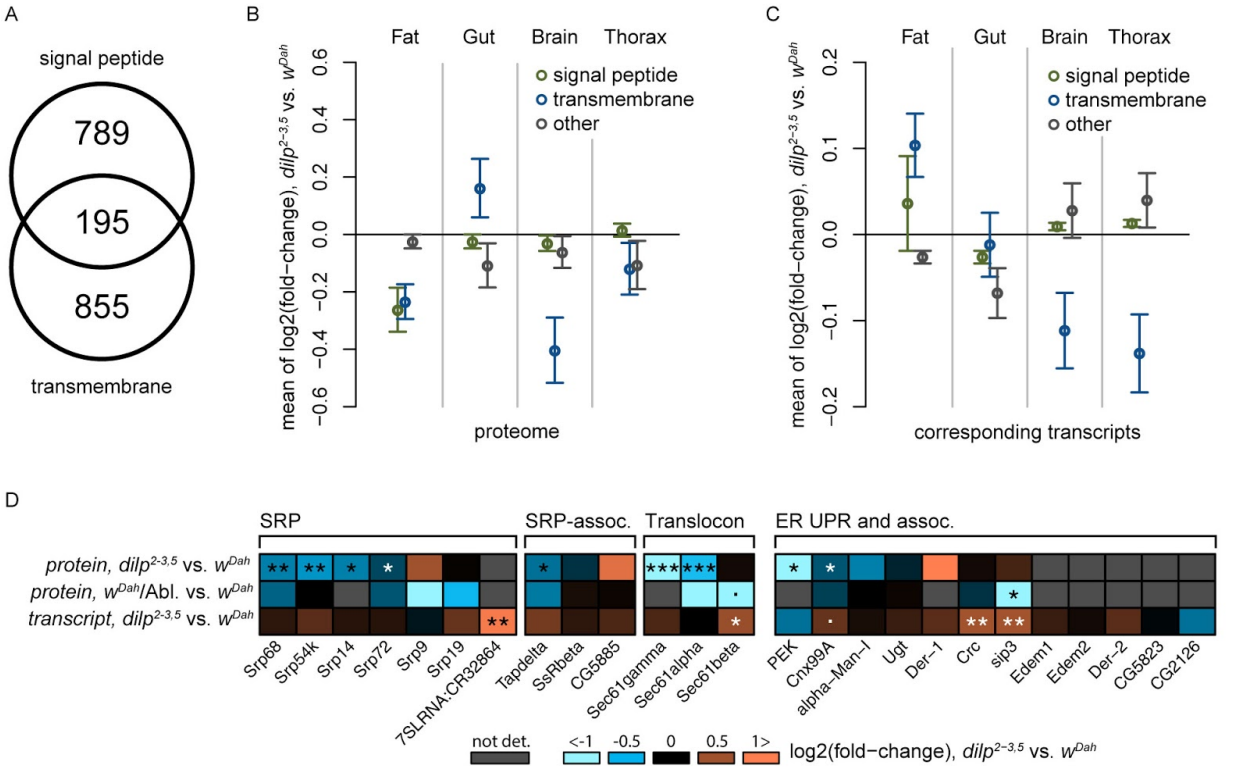


Fig 38***: Tissue-specific regulation of ER targeting mechanisms in response to reduced insulin.***
(A) *Venn diagram showing the number of proteins detected in the dilp$^{2-3,5}$ model proteomics annotated as possessing a signal peptide, transmembrane proteins, and their overlap (annotations from UniProt (The UniProt Consortium 2017)).* ***(B)*** *Average log-fold-changes (dilp$^{2-3,5}$ vs. w^{Dah}) of signal peptide containing, transmembrane, and all other proteins, in fat, gut, brain, and thorax. Error bars represent 95% confidence intervals.* ***(C)*** *Average log-fold-changes (dilp$^{2-3,5}$ vs. w^{Dah}) of transcripts corresponding to signal peptide containing, transmembrane, and all other proteins, in fat, gut, brain, and thorax. Error bars represent 95% confidence intervals.* ***(D)*** *Heatmap showing the regulation of SRP subunits, SRP-associates proteins, translocon proteins, and proteins implicated in the ER unfolded protein response (dilp$^{2-3,5}$ vs. w^{Dah} or w^{Dah}/Abl vs w^{Dah}), as well as the regulation of matching transcripts (dilp$^{2-3,5}$ vs. w^{Dah}). Asterisks indicate BH-corrected significance of the limma moderated t-test (proteomics) or DeSEQ2 Wald test (transcriptomics)(.p ≤ 0.1, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001).*

More general classes of proteins processed by the ER, i.e. transmembrane proteins and proteins possessing a signal peptide, were also consistently down-regulated in both models of reduced IIS (Fig 38A-C). Signal peptides tag nascent proteins for translocation to the ER (Saraogi & Shan 2011). Interestingly, among IIS-responding proteins that showed a significant statistical interaction between genotype and *Wolbachia* presence (*Wolbachia*-dependent), 72 (30%) were known to contain signal peptides, a highly significant overlap (hypergeometric test, p-value = 3.72e-10)(Fig 39). The overlap with transmembrane proteins (27 proteins) was not significant. This suggests that *Wolbachia* presence affects regulation of secreted proteins processed by the ER, but not transmembrane proteins.

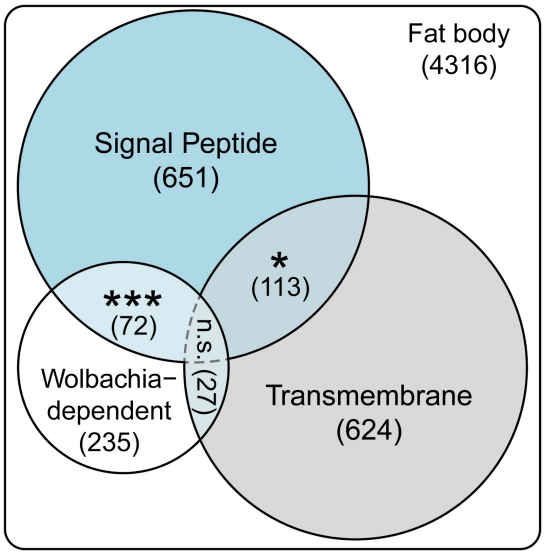


Fig 39***: Wolbachia-dependent proteins in the fat body are highly enriched for signal peptides.***
*Euler diagram showing the overlap of Wolbachia dependency, proteins with signal peptide, and transmembrane proteins in the fat body. Asterisks indicate significance of the overlap (*P-Value ≤ 0.05, ***P-Value ≤ 1e-9, n.s. = not significant, hypergeometric test).*

These results prompted us to investigate whether there were any changes in the ER translocation machinery upon reduced IIS. Nascent polypeptides with signal peptides are recognized by the signal recognition particle (SRP) in the cytosol and transported to ER translocon channels (Saraogi & Shan 2011). There, translation is resumed by the ER bound ribosomes, in a process called co-translational localization (Saraogi & Shan 2011). Regulation of the SRP thus determines ER import capacity. Another function of the SRP is to preserve polypeptides in their nascent state through elongation arrest, until the ribosome-polypeptide complex has been targeted to the translocon (Walter & Blobel 1981; Keenan et al. 2003). Therefore, maintaining SRP protein levels ensures correct ER-protein trafficking and the rate of import. Increased SRP activity can suggest increased co-translational elongation-arrest, reducing ER-import in response to reduced ER capacity (Lakkaraju et al. 2008; Walter & Blobel 1981). We found that four of the seven SRP subunits (*Srp68, Srp54k, Srp14,* and *Srp72)* were significantly reduced (12%-25%) in long-lived $dilp^{2\text{-}3,5}$ mutants (Fig 38D). However, we also detected 2-fold up-regulation of the SRP RNA, 7SL, in the fat body of long-lived flies (Fig 38D). 7SL abundance is an indicator for SRP assembly and is required for the formation of the SRP elongation arrest domain (Lakkaraju et al. 2008). Increased 7SL levels could suggest an increase in SRP activity despite reduced levels of SRP proteins; this could both increase protein import rate to the ER or decrease it through elongation arrest (Lakkaraju et al. 2008). We therefore turned our attention to the translocon complex, which forms the physical channel through which polypeptides enter the ER (Johnson & van Waes 1999; McCormick et al. 2003). Down-regulation of translocon

complex members could indicate a reduction in ER capacity in response to reduced IIS. Indeed, in response to reduced IIS, *Sec61alpha, Sec61beta*, and *Sec61gamma*,subunits of the core translocon channel, were down-regulated in one or both models of reduced IIS (Fig 38D). Together, these results suggest that reduced IIS leads to a reduction in co-translational protein import, and that this reduction may be associated to longevity.

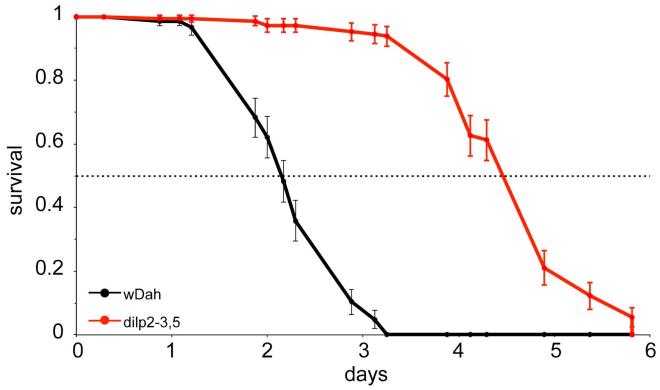


Fig 40*: **Survival of aged dilp$^{2\text{-}3,5}$ and w^{Dah} flies during exposure to ER stress.***
Results and figure by Luke Tain. Plot showing survival of flies exposed to 15uM tunicamycin, starting at 45d of age. Dilp$^{2\text{-}3,5}$ mutants are significantly more resistant to tunicamycin treatment compared to controls ($P \leq 0.0001$, log-rank survivorship test) .

3.3.4.6 The *dilp$^{2\text{-}3,5}$* fat body is resistant against ER stress

It has previously been reported that, as a result of altered ER proteostasis, an unfolded protein response (UPR) may be be activated and reduce translation (Boyce et al. 2005; Scheuner et al. 2005). Here, we found that in response to reduced IIS the translocation machinery, transporting mRNAs to the ER to be translated, is down-regulated, and may in turn activate the UPR, and thus reduce translation. The UPR functions by recognizing and re-localizing misfolded proteins into the cytosol, where they are degraded. To determine if changes to proteins associated to mechanisms of translation, such as ribosome biogenesis, translational rate modulation, and ER translocation, may influence the ability of cells to protect against ER stress, we examined the resistance of IIS mutant flies to tunicamycin. Tunicamycin disrupts the ability of the chaperone calnexin to bind to misfolded glycoproteins, leading to their accumulation in the ER and inducing ER stress (Parodi 2000). IIS mutant flies (*dilp$^{2\text{-}3,5}$*) were 90% more resistant to tunicamycin compared to control flies (*w^{Dah}*)(Fig 40). To determine if this related to reduced protein import into the ER, or an increased UPR, we investigated known markers of the endoplasmic UPR in our data (Fig 38D). Most proteins entering the ER undergo extensive folding and post-translational modification. Folding of glycoproteins, which make up the majority of secreted proteins, is facilitated by cycling polypeptide chains between the chaperones *Cnx99A* (calnexin) and *Crc* (calreticulin). We found both transcripts to be up-regulated in *dilp$^{2\text{-}3,5}$* flies,

but *Cnx99A* protein was down-regulated (Fig 38D). After de-mannosylation by *alpha-Man-I*, misfolded proteins are recognized by *Ugt* and targeted to the ERAD pathway, which re-translocates them into the cytoplasm to be degraded by the proteasome (Schröder & Kaufman 2005). While we found 7 out of 13 proteasome core complex members up-regulated specifically in *dilp*$^{2\text{-}3,5}$ mutants, we have previously found that activity of the proteasome in the fat body of *dilp*$^{2\text{-}3,5}$ flies is not increased upon reduced IIS (Tain et al. 2017)(Figure 2). We further failed to detect regulation of the ERAD pathway, or of *Ugt* or *alpha-Man-I* (Fig 38D). Only *sip3* transcript, but not its product, was up-regulated significantly in *dilp*$^{2\text{-}3,5}$ flies (Fig 38D), but down-regulated in mNSC Ablation flies. Furthermore, the ER transmembrane kinase *PEK*, which decreases translation through dephosphorylation of *eIF-2alpha* in response to ER stress (Sood et al. 2000; Harding et al. 2000), was down-regulated in response to reduced IIS (Fig 38D). Together, our results suggest that the down-regulation of ER proteins is not the consequence of chronic ER stress; however, the reduction in translation may alleviate proteostatic stress.

3.3.4.7 Regulation of miRNA upon reduced IIS

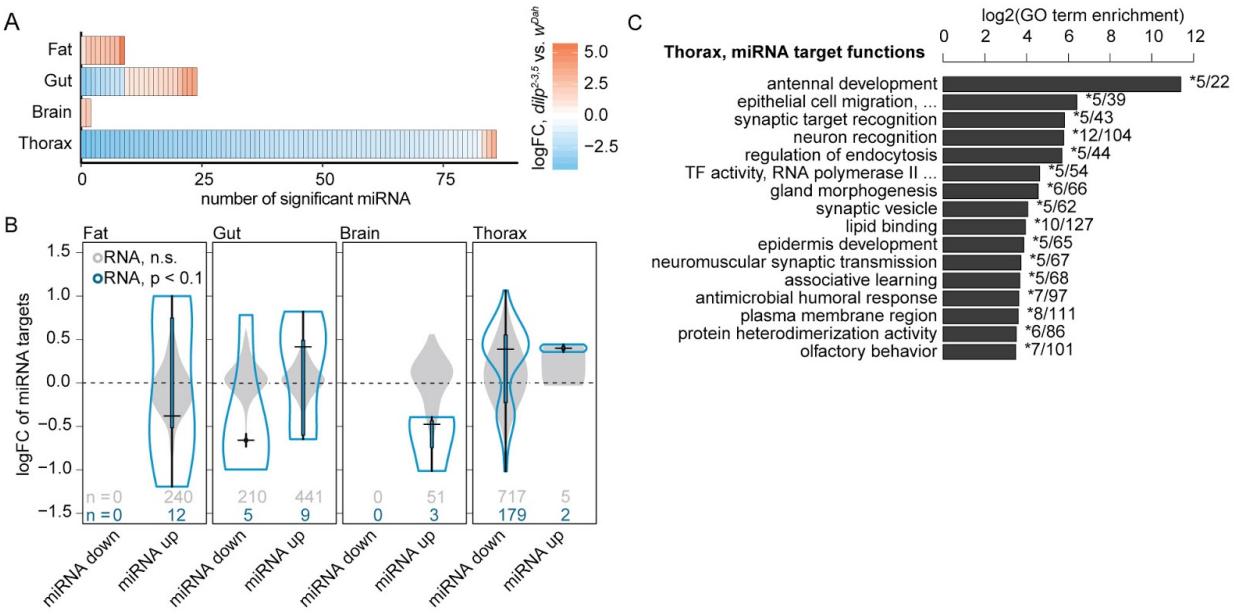


Fig 41***: Tissue-specific regulation of miRNA upon dilp$^{2\text{-}3,5}$ knockout.***

(A) *Venn diagram showing the number of detected miRNA per tissue and their overlap (above a median normalized count threshold of 5). Barplot shows the number of miRNA detected in at least n ([1,4]) tissues.* ***(B)*** *PCA projection of dilp$^{2\text{-}3,5}$ and w^{Dah} miRNA samples from all tissues.* ***(C)*** *Cell plot showing the number of significantly regulated miRNA (adj. P-Value $\leq$0.1). Cell colours represent log-fold-changes between dilp$^{2\text{-}3,5}$ and w^{Dah} in all tissues.* ***(D)*** *Violin plots showing the log-fold-change distributions (dilp$^{2\text{-}3,5}$ vs. w^{Dah}) of putative target genes of significantly up- or down-regulated miRNA. Grey plots show distributions of non-significant putative target genes, blue outlines show distributions of significantly changed (adj. P-Value $\leq$0.1) putative target genes.* ***(D)*** *Bar plot showing log-enrichment of GO terms in the set of significantly changed putative target genes (adj. P-Value $\leq$0.1) of regulated miRNA in the*

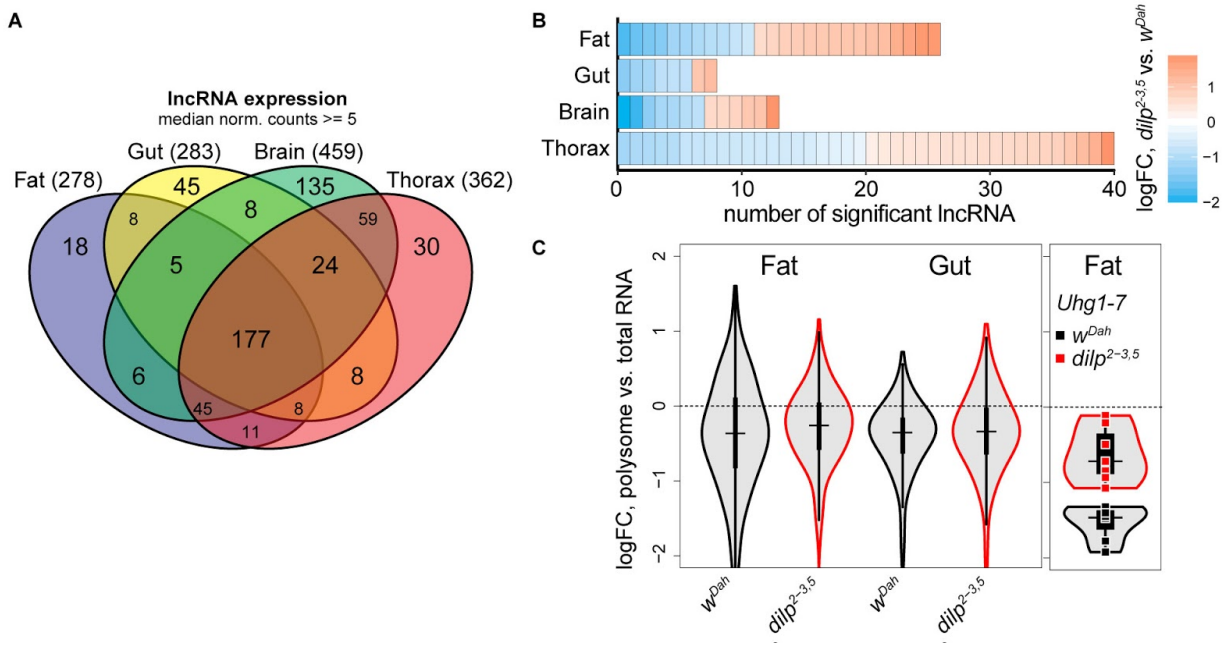
thorax. Numbers at the end of bars show the number of significant genes / total number of annotated genes detected in the thorax.

Control of gene expression through miRNA-mediated interference is a conserved mechanism of post-transcriptional regulation (Lee et al. 1993). To determine the extent to which miRNA-mediated interference influences gene expression in response to reduced IIS we quantified miRNAs in all four tissues of *dilp*$^{2\text{-}3,5}$ and controls flies. To ensure robustness against tissue-wide changes in miRNA regulation, we normalized all miRNA counts to 2S-rRNA expression, a commonly used control in *Drosophila* miRNA quantification (Ameres et al. 2010; Tang et al. 2010; Marques et al. 2010). In total 113 miRNAs were regulated in response to reduced IIS (Fig 41C). Of these, two were regulated in the brain, nine in the fat, 24 in the gut, and 86 in the thorax (Fig 41C). All but three significant miRNA in the thorax were down-regulated in *dilp*$^{2\text{-}3,5}$ mutants (Fig 41C). We detected down-regulation of *Dcr-2* and *R2D2* transcripts, necessary for siRNA processing (Ha & Kim 2014), in response to reduced IIS. However, we did not detect reduction of the miRNA processing machinery including *drosha*, *Dcr-1* and *AGO1-3* (File 1). This suggests that miRNA expression in the thorax may be down-regulated transcriptionally, not post-transcriptionally. To estimate the effects of miRNA on gene expression, we combined predictions from three miRNA target prediction tools—targetScan, mirna.org, and microT_CDS (Ruby et al. 2007; Enright et al. 2003; Reczko et al. 2012)—and generated a list of putative miRNA targets (File 5). This led to 1166 predicted targets of regulated miRNA, 202 of which were differentially regulated in our transcriptomics data. In the thorax, most significantly changed miRNA were down-regulated in response to reduced IIS (Fig 41C), which led us to expect an overall increase in the expression of their putative targets in long-lived mutants. Indeed, we found that 129 out of 179 transcriptionally regulated target genes were significantly increased in the thorax of *dilp*$^{2\text{-}3,5}$ flies, suggesting de-repression due to down-regulation of the corresponding miRNA (Fig 41D).

GO enrichment analysis of regulated targets in the thorax revealed terms related to development, transcription factor activity, immune response, and neuromuscular functions (Fig 41E). The most prolific miRNAs in the thorax are miR-310/miR-311/miR-312, which target 18% of regulated mRNA that were identified as putative miRNA targets in that tissue (File 5). This miRNA family regulates metabolic status and nutrient sensitivity (Çiçek et al. 2016) and, in part, functions as negative regulators of neurotransmitter release at neuromuscular junctions (Tsurudome et al. 2010). Another prolific regulated miRNA, miR-316, is known to be specific to visceral muscle (Aboobaker et al. 2005). The expression levels of several *Drosophila* miRNA can affect muscle function directly (Sokol & Ambros 2005; J.-F. Chen et al. 2006; Katti et al. 2017). Post-transcriptional regulation through miRNA de-repression might therefore contribute to the observed improvement in age-related muscle function (Fig 17). In the brain, miR-375 (both -3p

and -5p products), which regulates insulin secretion in humans (El Ouaamari et al. 2008), was up-regulated (File 5), whereas its putative targets in the were significantly repressed (Fig 41B). No global pattern of regulation by miRNA could be observed in fat and gut (Fig 41B). Taken together, our results suggests that although PTR occurs in the gut and the fat, in these tissues other mechanisms of PTR play a greater role than miRNA in response to reduced IIS.

3.3.4.8 Regulation of lncRNA upon reduced IIS



Long non-coding RNA (lncRNA) play important roles in development (B. Chen et al. 2016) and can regulate metabolism in response to insulin signalling (Ellis et al. 2014). Importantly, lncRNA can also mediate post-transcriptional regulation through a variety of mechanisms including fine-tuning of miRNA binding (Cesana et al. 2011), controlling translation-competency of opposite strand protein coding RNA (Carrieri et al. 2012), and as part of protein complexes (Gundelfinger et al. 1984; G Hendrickson et al. 2016). They can also serve as precursors for small nucleolar RNA's (snoRNA) that act to regulate ribosomal biogenesis (Hughes et al. 2012; Kiss 2002). In total, we quantified the expression of 587 annotated lncRNA's with our our transcriptomics approach in at least one of the tissues, and found 177 lncRNAs shared in all tissues (Figure 3D). Only 70 lncRNAs were differentially regulated in response to reduced IIS (Figure 3E). As yet very few lncRNAs have been functionally annotated in *Drosophila*. LncRNA may associate with ribonucleoproteins (RNPs), though whether they are targeted to ribosomes for degradation or fulfill a regulatory role is an unresolved question (Carlevaro-Fita et al. 2016). Here, we quantified the relative change of lncRNA abundance between the total RNA fraction and polysome fraction in fat and gut. The population of lncRNA bound to ribosomes in the gut

and fat was significantly less than that found in the total RNA sample, suggesting the majority of lncRNAs are not directly bound to ribosomes (Figure 3F). However, those lncRNAs that change their binding behavior in response to reduced IIS might be actively involved in the regulation of translation. To detect these lncRNAs, we determined the statistical interaction between their enrichment in the polysome fraction and their change in response to *dilp*$^{2\text{-}3,5}$ knock-out. We identified four such differentially enriched lncRNA's in the fat body. Three of those, CR44263, Uhg4, and Uhg2, were diminished by up to 74% in the polysome fraction of *w*^{Dah} flies, but significantly less (44%) in long-lived flies. This tendency was present in the expression of all seven detected U snoRNA host genes (Uhg's)(Figure 3G). While this implies a possible layer of PTR in the fat body that could regulate translation through snoRNAs, additional studies will be needed to characterize this interaction.

3.3.5 Dynamics of insulin signalling with age

3.3.5.1 Section introduction and overview

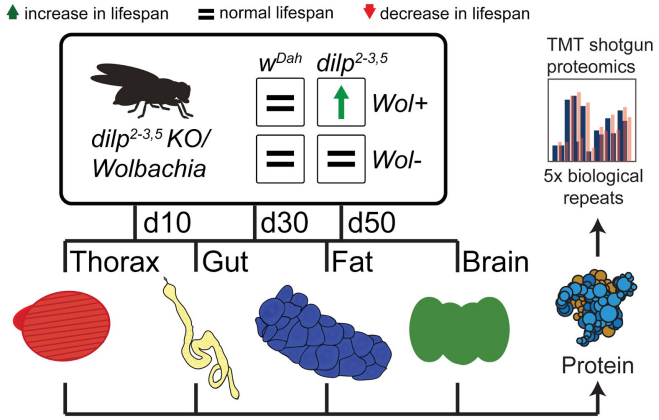


Fig 42: **Experimental design of the *dilp*$^{2\text{-}3,5}$ proteomics time series.**
Fly tissue samples from thorax, gut, fat, and brain were collected at days 10, 30, and 50 from four genotypes: w^{Dah} wild types and dilp$^{2\text{-}3,5}$ mutants in the presence or absence (T) of Wolbachia (note that this section describes preliminary results from only fat and thorax). Proteins were quantified by collaborators (Ilian Atanassov) using TMT shotgun proteomics.

Previous sections have focused on the analysis of the fly insulin response in *dilp*$^{2\text{-}3,5}$ and mNSC-ablation models in 10-day old flies. While many differences between insulin mutants and wild type flies are already apparent at this age and were confirmed to be associated to lifespan (see previous chapters), it does not follow that they reverse age-related changes. Other beneficial effects of insulin signalling may only be detectable at later ages. To test these possibilities, we quantified the tissue-specific proteomes of *dilp*$^{2\text{-}3,5}$ and w^{Dah} flies, with and without the endosymbiont *Wolbachia* present, at three time points: 10 day, 30 day, and 50 day old (Fig 42). This extended our ability to detect longevity associated changes upon reduced IIS in two ways: firstly, we may identify changes that occur due to reduced IIS, but change in the opposite direction during normal ageing. Secondly, we can observe changes that are not detectable in 10d old flies, but manifest as mutants and their controls grow older.

Here, we present a preliminary analysis of these *dilp*$^{2\text{-}3,5}$ model proteomics time series data, encompassing two tissues: the fat body and the thorax. We show that the insulin response may be stable across the age, but changes at a higher initial rate in the fat body than the thorax. We identified proteins that are not regulated between the long-lived flies and their controls at an early age, but diverged in older flies. Among these late-life divergent proteins, we further delineated those that differed in their expression pattern depending on presence or absence of *Wolbachia,* which is required for lifespan extension in *dilp*$^{2\text{-}3,5}$ flies. We found key processes to be regulated in this highly specific manner: intracellular transport in the fat body, as well as the

proteasome in the thorax. Our preliminary results therefore point the way for future study and experimental validation.

3.3.5.2 TMT proteomics enable extensive comparison across genotypes and age

We took advantage of tandem mass tag (TMT) multiplexing (Thompson et al. 2003; Dayon et al. 2008; Ting et al. 2011) to analyse all four genotypes of our model system at days 10 and 50, as well as two genotypes (dilp$^{2\text{-}3,5}$ and w^{Dah}) at day 30, with 5 replicates each. The primary benefit of TMT multiplexing is the concurrent handling and processing of samples from up to 10 samples. For each of our replicates, we therefore processed all 10 experimental conditions at once, eliminating one source of technical variability entirely (Table 1). We furthermore took care to rotate the chemical labels used for multiplexing across replicates (Table 1), to avoid any systemic bias that might arise from their use. The remaining technical bias—that between replicate runs—was corrected by incorporating the TMT replicate batch structure as an additional factor in the linear models used in the downstream analysis.

Table 1*: Technical batch structure of TMT proteomics quantification.*
Genotypes: D (dilp$^{2\text{-}3,5}$) or W (w^{Dah}), Wolbachia present (+) or absent (-). Numbers represent chemical labels.

Genotype:	**W-**	**W+**	**D-**	**D+**	**W+**	**D+**	**W-**	**W+**	**D-**	**D+**
Time point:	**10d**				**30d**		**50d**			
Replicate 1	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Replicate 2	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*1*	*2*
Replicate 3	*5*	*6*	*7*	*8*	*9*	*10*	*1*	*2*	*3*	*4*
Replicate 4	*7*	*8*	*9*	*10*	*1*	*2*	*3*	*4*	*5*	*6*
Replicate 5	*9*	*10*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*

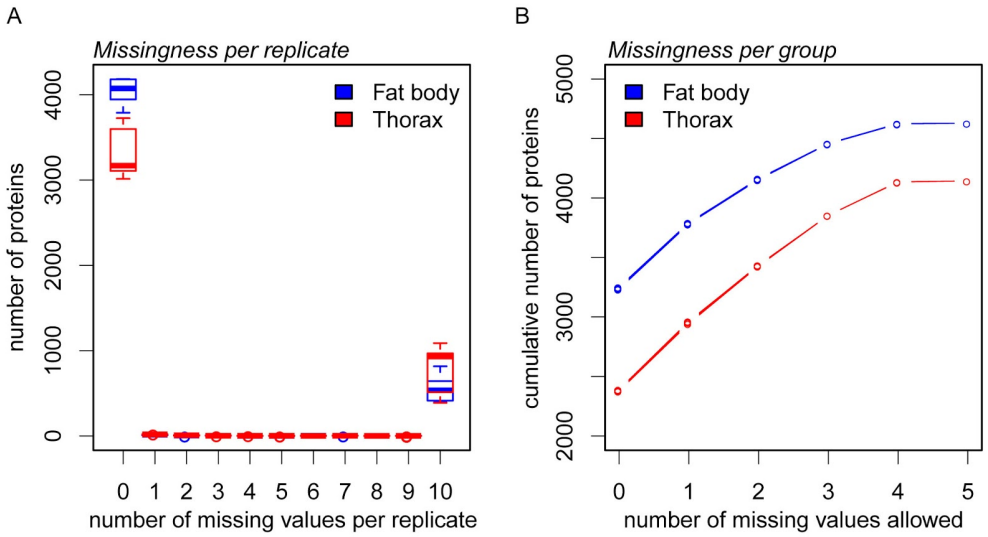


Fig 43***: Missingness in dilp$^{2\text{-}3,5}$ model TMT shotgun proteomics.***
(A) *Boxplots showing the number of proteins with exactly n missing values in each of the five TMT replicate runs (each containing all 10 experimental conditions), in fat body and thorax.* ***(B)*** *Line chart showing cumulative numbers of proteins, depending on the number of missing values allowed in each experimental condition (thus 10 lines are plotted for each tissue, but these are overlapping due to high consistency).*

After mass spectrometry measurements and quantification in MaxQuant by our collaborators (Ilian Atanassov) we carried out imputation of missing values. In LFQ shotgun proteomics experiments, missing values commonly arise from left-censoring as a result of the limited dynamic range (Karpievitch et al. 2012). As a consequence, more abundant proteins are more likely to be detected (Karpievitch et al. 2012). The resulting missingness pattern therefore requires a specialized imputation technique that samples from a left-shifted normal distribution fitted to the truncated observed distribution of LFQ values (Deeb et al. 2012). Though TMT proteomics share the dynamic range limitation of LFQ proteomics, we observed a different missingness pattern in the TMT data: most proteins were detected either in all or no experimental conditions in a given TMT replicate run (Fig 43A). Furthermore, we checked whether missingness varied based on experimental condition; this was not the case (Fig 43B). The missingness pattern was therefore highly regular, possibly due to simultaneous processing combined with information sharing between samples carried out by MaxQuant (REF). We concluded that most missing values in our TMT proteomics data can be treated as missing-at-random rather than left-censored. We first filtered out proteins that had more than three missing replicate values in any experimental condition. We then used k-nearest neighbor imputation to impute these values (*impute* R package)(Hastie et al. 2001).

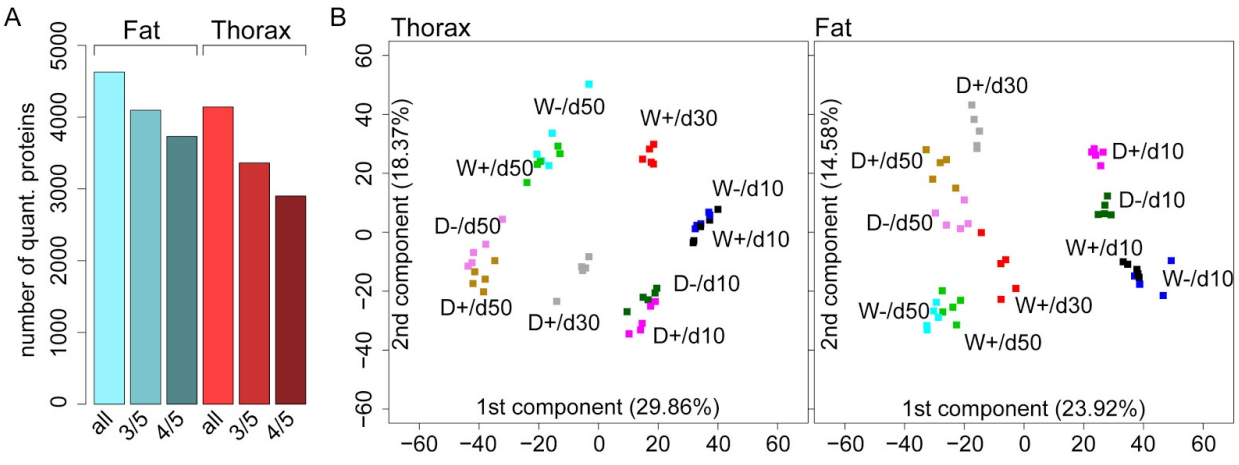


Fig 44*: **Quantification and quality control of fat and thorax proteomics.***
***(A)** Number of quantified proteins in each tissue at different filtering thresholds: quantified in at least one sample (all), or quantified in at least three out of five or four out of five samples (3/5, 4/5), from every genotype. **(B)** PCA projections of all sample replicate quantifications in fat and thorax. The first component mainly captures the time variable (d10, d30, d50), whereas the second component separates the genotypes, D (dilp$^{2\text{-}3,5}$) and W (w^{Dah}) as well as Wolbachia presence (+) or absence (-).*

The number of identified proteins matched the previous state-of-the-art quantification using LFQ proteomics in the fat body, and exceeded them in the thorax (Fig 44, Fig 3, Fig 13). Notably, PCA analysis of both fat and thorax showed a clear separation by age (1st component) and genotype (2nd component)(Fig 44).

3.3.5.3 Age-related changes in w^{Dah} flies

We first identified age-related changes in w^{Dah} flies, using a linear polynomial model to model the trend protein expression E with time t (Eq 10). Here, t is an ordinal categorical variable centered at zero (day 30). The coefficient β_0 is the intercept and takes on the mean expression of the respective protein across all time points. We incorporated the TMT replicate batch structure as a factor with five levels (coefficient β_3). The coefficient β_1 describes the magnitude of the monotonous change with t (linear trend), whereas β_2 captures the magnitude of the non-monotonous change (quadratic trend). The linear trend will be most significant for proteins whose expression levels increase consistently from day 10 to day 30, and then from day 30 to day 50. The quadratic trend can identify cases where protein expression is most extreme at the middle time point, e.g. if it increases between day 10 and day 30, but then decreases from day 30 to day 50 (or the opposite case). For our preliminary analysis, we decided to focus on the linear case.

$$E_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 Batch \quad (10)$$

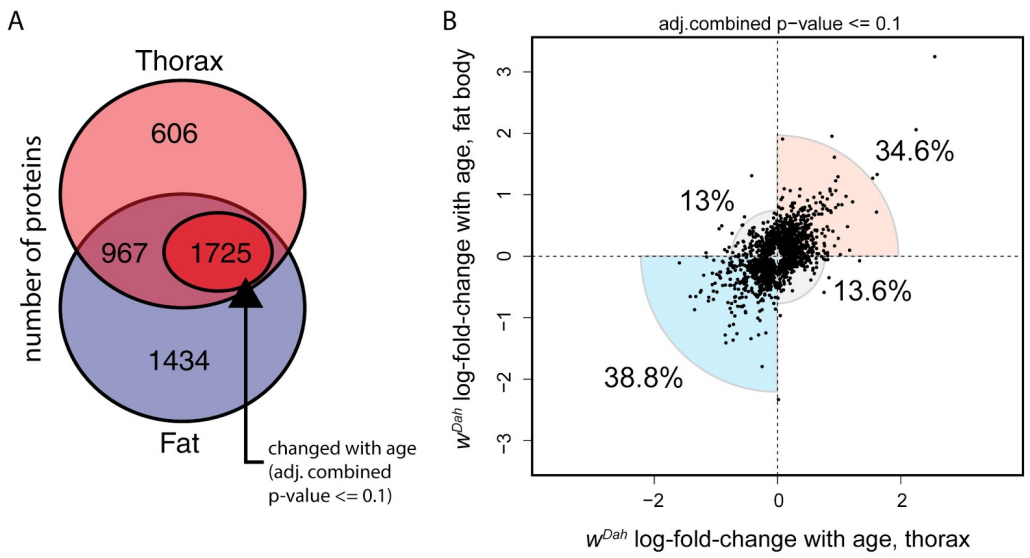


Fig 45***: Global similarity of age-associated changes in fat and thorax.***
(A) *Venn diagram showing the number of detected proteins (after thresholding) in fat and thorax. The intersection represents proteins common in both tissues. Of those, proteins with a combined adj. p-value of ≤ 0.1 (linear association with age in w^{Dah}) are a strict subset.* ***(B)*** *Scatterplot of proteins significantly associated with age (combined p-value ≤ 0.1), showing their log-fold-changes in the fat body (y-axis) and thorax (x-axis). Circle segments and percentages represent the relative proportion of points in each quadrant.*

At 95% confidence, 1336 proteins in the fat body and 1416 proteins in the thorax were observed to change monotonously with age in w^{Dah} flies (Fig 45A). To determine how similar these age-related changes were between the fat and the thorax, we first combined their p-values using the Stouffer method (Stouffer et al. 1949) and adjusted them for multiplicity using the BH method. Out of 2692 proteins quantified in both tissues, 1725 were significantly changed with age at a combined adjusted p-value of ≤ 0.1 (Fig 45A). 73.4% of significant protein log-fold-changes were regulated in the same direction in both tissues, resulting in a high positive rank correlation (rho = 0.54)(Fig 45B). This suggests that a large fraction of systemic

age-associated changes are common to both fat and thorax.

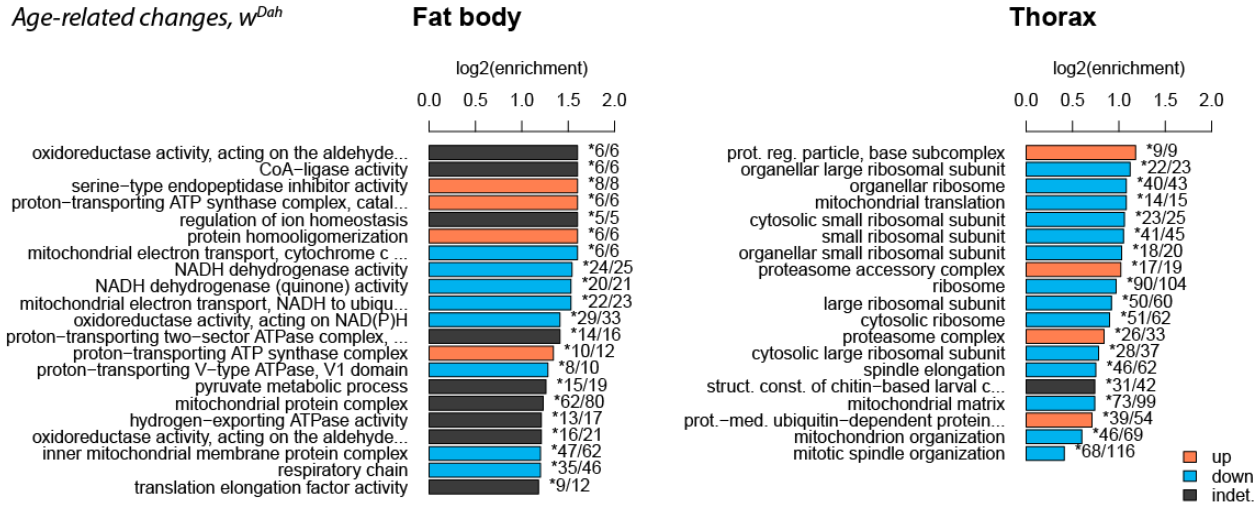


Fig 46***: Age-related changes in w^Dah flies.***

Bar charts displaying GO categories enriched for proteins that are showing a significant monotonous increase or decrease with age, from day 10 to day 30 to day 50. After each bar, the number of significant proteins and the total number of annotated and quantified proteins for the respective term are shown. Colored bars indicate a significant bias towards up-regulation (orange) or down-regulation (blue) of that term with age, as determined by two one-sided t-tests against zero carried out on significant log-fold-changes (p-value ≤ 0.05).

We then carried out GO enrichment analysis on significantly age-associated proteins (adj. P-value ≤ 0.05) in each tissue separately. This revealed a coordinated down-regulation of the four respiratory chain complexes (ETC) and mitochondria with age (Fig 46). It is well known that mitochondrial function declines in older organisms (Bratic & Larsson 2013). With age, the magnitude of the membrane potential declines as proton leakage increases, potentially due to oxidative damage (Shigenaga et al. 1994; Hagen et al. 1997; Harper et al. 1998). Additionally, the structure of the mitochondrial inner membrane becomes more disordered, leading to the dissociation of ATP synthase complexes (Daum et al. 2013). However, we observed an up-regulation of the proton-transporting ATP synthase complex with age in the *w^Dah* fat body (Fig 46), which suggests a possible compensatory role of this complex. Additionally, we recorded a mixed regulation of translation elongation factor activity (Fig 46). Detected components of elongation factors *eEF1* and *eEF2* showed moderate but significant down-regulation with age, suggesting that a slowing down of translation rate may occur. However, several mitochondrial elongation factors (*mEFTs, mRRF2, mEFTu1, mEFTu2*) were consistently up-regulated with age.

In the thorax, mitochondrial translation, ribosomal subunits, and the proteasome accessory complex were most clearly affected (Fig 46). Mitochondrial translation was consistently down-regulated with age. Mitochondrial ribosomal proteins were most consistently and strongly affected, but also elongation factors such as mEFG1 and *mEFTu2* (but not *mEFTu1*). Interestingly, both proteins of the proteasome complex and the proteasome accessory complex

were on average significantly up-regulated with age. While proteasome function also crucially depends on complex assembly which we did not measure, the observed increase in protein expression may be a response to increased proteotoxic stress with age, as has been reported before (Cohen & Dillin 2008). In support of this, the age-associated heat shock factor Hsp22 (Yang & Tower 2009) was highly up-regulated in both measured tissues of aged w^{Dah} flies (thorax: 5.6-fold, fat: 9.5-fold).

3.3.5.4 The global insulin response is stable throughout life

We next asked if and how the response to $dilp^{2\text{-}3,5}$ knock-out varies across age and tissue. For this, we fitted a separate linear model to our protein expression values to allow the subsequent fitting of contrasts between individual experimental conditions (Eq 11). These conditions are encoded in the 10-level factor *GenotypeAndTime*. We again included the TMT replicate batch levels as an additional factor to eliminate batch biases between replicates.

$$Expression = 0 + GenotypeAndTime + Batch \qquad (11)$$

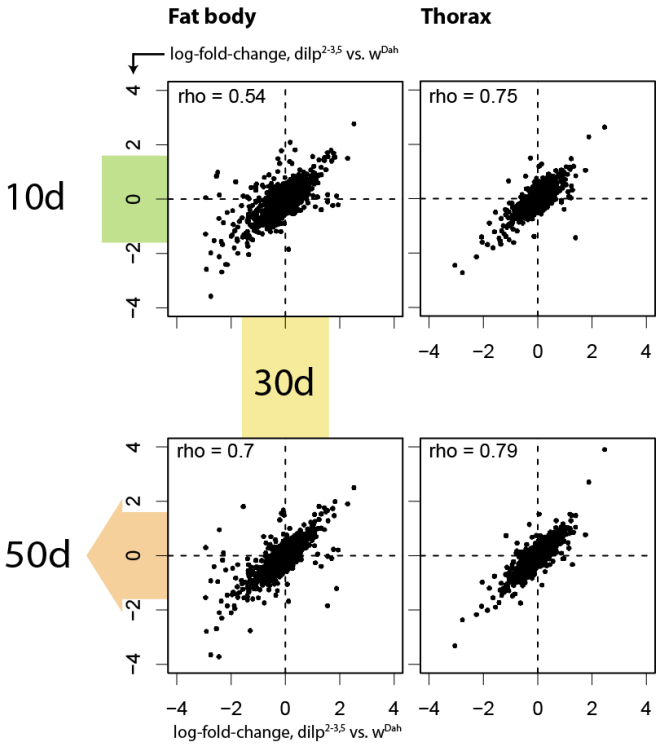


Fig 47*: **Correlation of the insulin response between subsequent time points.***
Scatterplots comparing log-fold-changes between $dilp^{2\text{-}3,5}$ and w^{Dah} in fat and thorax at subsequent time points: the x-axis shows changes at day 30, the y-axis of the first row changes at day 10, and the y-axis of the second row changes at day 50. The corresponding Spearman correlation coefficient rho is noted in the top left corner of each scatterplot.

We then fitted contrasts to calculate the differential expression between *dilp*$^{2\text{-}3,5}$ and *w*^{Dah} at day 10, 30, and 50 (File 3). To determine how coherent the insulin response is at subsequent time points, we estimated their pairwise rank correlations. In the fat body, the insulin responses at day 10 and 30 showed a moderate correlation (0.54), while day 30 and 50 were more closely correlated (0.7)(Fig 47). In the thorax, day 10 and day 30 were as closely correlated with each other (0.75) as day 30 and 50 (0.79)(Fig 47). This suggests that the insulin response is generally stable throughout life, but the expression pattern of the fat body changes more strongly in early to mid life.

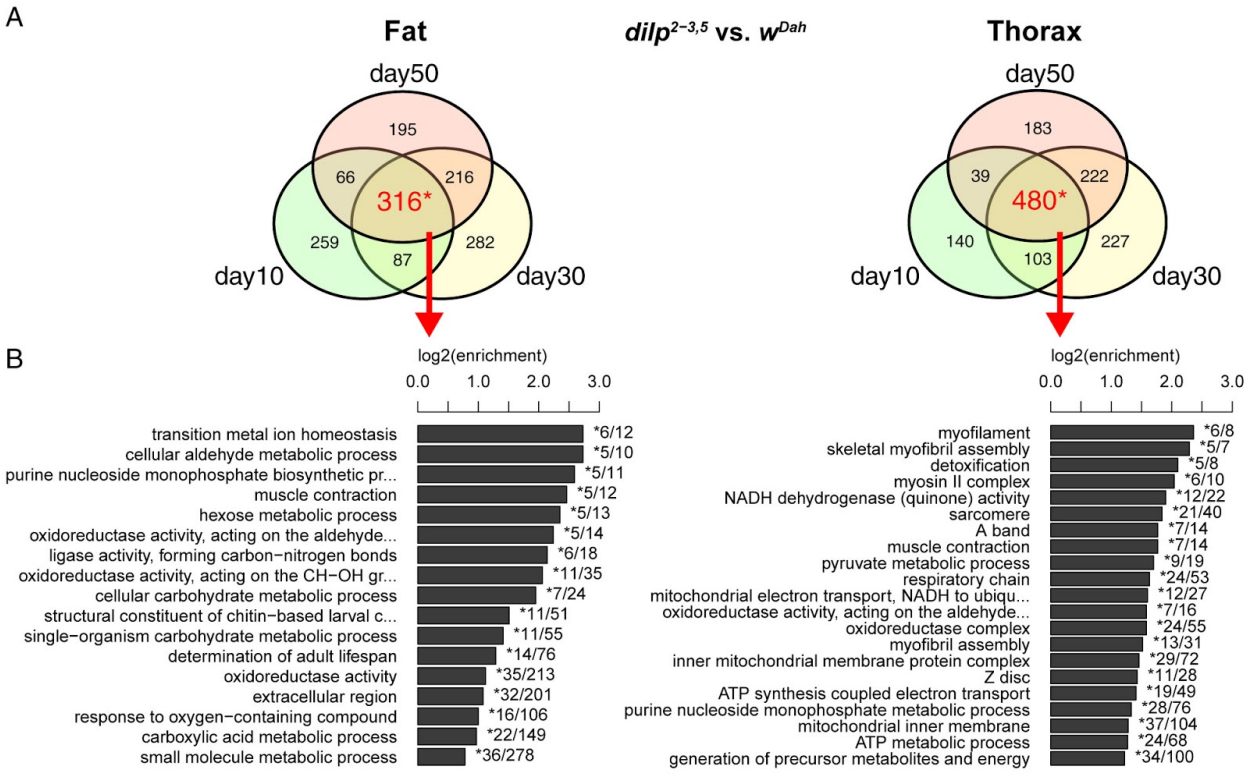


Fig 48*: **Functional enrichment of the core insulin response.***
***(A)** Venn diagrams showing the number of proteins significantly changed between dilp*$^{2\text{-}3,5}$ *and w*^{Dah} *in each tissue at day 10, 30, and 50. Proteins changed at all time points in a given tissue are designated the core insulin response (numbers in red).* ***(B)** Bar charts showing GO terms enriched in proteins of the core insulin response.*

To determine the processes most consistently affected by insulin signalling at every age, we selected in each tissue all proteins that were changed in response to reduced IIS at every time point (Fig 48A) and performed GO enrichment analysis on these subsets (Fig 48B). Top terms in the fat body included transition metal ion homeostasis and carbohydrate metabolism related terms, as well proteins known to regulate adult lifespan (Fig 48B). In the thorax, we saw muscle and respiratory chain related terms most strongly affected. This is consistent with our earlier observations that muscle function declines with age, but is improved in *dilp*$^{2\text{-}3,5}$ mutants (Fig 17).

3.3.5.5 Divergence between the *dilp*$^{2\text{-}3,5}$ and *w*^{Dah} insulin response with age

To determine which IIS induced changes were most likely associated to an increase in lifespan, we identified proteins whose response to IIS with age differed between the long-lived *dilp*$^{2\text{-}3,5}$ mutants and *w*^{Dah} controls, testing the statistical interaction between genotype and age (File 3). For this, we fitted an interaction term (Eq 12) to our time-dependent linear model (Eq 10). Here, *L(dilp*$^{2\text{-}3,5}$*)* and *L(w*^{Dah}*)* represent the coefficients of the age linear trend, or simply the change of protein levels with age in long-lived mutants and their controls.

$$AgeInteraction = L(dilp^{2\text{-}3,5}) - L(w^{Dah}) \qquad (12)$$

Proteins with significantly regulated (adj. P-Value ≤ 0.1) *AgeInteraction* coefficients were then subjected to further filtering to determine proteins whose expression pattern diverged with age between long-lived flies and controls (here called late-life divergent). Finally, we identified proteins that varied in their response depending on the presence of the intracellular symbiont *Wolbachia*, which is required for longevity of *dilp*$^{2\text{-}3,5}$ flies.

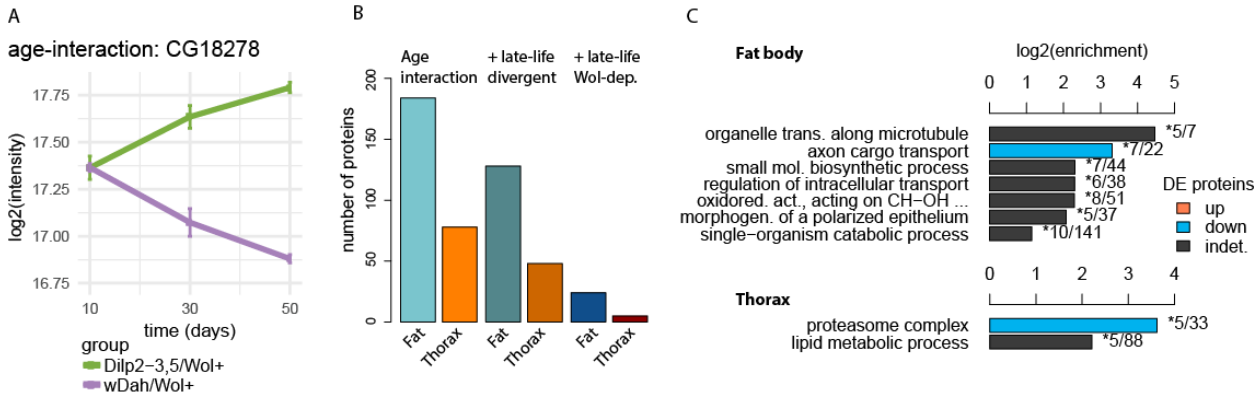


Fig 49***: Identifying the insulin-dependent ageing phenotype.***
(A) *Example time course of a protein with a significant interaction between age and genotype, whose expression levels diverge between dilp*$^{2\text{-}3,5}$ *and w*^{Dah} *with age.* ***(B)*** *Bar charts showing the number of proteins in fat and thorax whose expression levels show a significant interaction between age and genotype (age interaction), and in addition to that show a greater log-fold-change at day 50 than at day 10 (late-life divergent) and may depend for their response at day 50 on the presence of Wolbachia (late-life Wolbachia dependent).* ***(C)*** *Bar charts showing GO terms enriched in late-life divergent proteins, in fat and thorax. Colored bars indicate a significant bias towards up-regulation (orange) or down-regulation (blue) of that term with age, as determined by two one-sided t-tests against zero carried out on significant log-fold-changes (p-value ≤ 0.05).*

One example of a protein fulfilling all three conditions is CG18278, which increases with age in the fat body of *dilp*$^{2\text{-}3,5}$ flies, but decreases in wild types (Fig 49A). CG18278 is a sparsely

annotated gene involved in the metabolism of glycosaminoglycan (FlyBase Curators et al., 2004-), a diverse class of proteins that participate in extracellular matrix formation and developmental signalling (Esko & Selleck 2002; Lin 2004). Overall, we identified 245 unique proteins and isoforms that differed in their insulin response with age (age-interacting)(Fig 49B). Of these, 67% were late-life divergent, and 11% were *Wolbachia* dependent at age 50d as well as late-life divergent (Fig 49B). To determine their putative functions, we carried out GO term enrichment analysis of late-life divergent proteins using SETHRO.

In the fat body, late-life divergent proteins were most strongly enriched for GO terms related to intracellular transport. These included the ubiquitous Khc (Kinesin heavy chain)(Table 2) and Dhc64C (Dynein heavy chain) proteins (File 3). Intracellular transport can have multiple roles in metabolism: it has been shown that translocation of glucose transporter GLUT4 to the membrane depends on trafficking along microtubules, thereby modulating glucose uptake in adipocytes (Fletcher et al. 2000). In particular, we observed *dilp*$^{2\text{-}3,5}$ down-regulation of lipid homeostasis associated proteins Lsd-1/Lsd-2 (Miura et al. 2002) in a late-life divergent manner (Table 2, File 3). Lsd-2 is responsible for the transport of lipid droplets along microtubules, thereby connecting intracellular transport to lipid homeostasis (Welte et al. 2005); flies lacking Lsd-2 are leaner than their wild type counterparts (Grönke et al. 2003). Interestingly, Bmm, whose reduction increases fat storage in flies by preventing mobilization of fat stores (Grönke et al. 2005; Grönke et al. 2007), was also down-regulated and late-life-divergent (File 3).

Table 2*: Fat body, top 30 most significant age-divergent proteins and their functions.*

Protein	w^{Dah}	*dilp*$^{2\text{-}3,5}$	Interaction	adj. P-Value	Notes
CG13315	2.1	0.84	-1.2	1.60E-05	
CG18278	-0.34	0.3	0.65	3.60E-05	glycosaminoglycan metabolic process, **Wolbachia-dependent (d50)**
Rtnl1	0.87	0.098	-0.77	3.90E-05	endoplasmic reticulum organization
Pepck	0.12	-0.54	-0.66	5.00E-05	gluconeogenesis
CG1882	0.63	-0.23	-0.86	8.40E-05	carboxylic ester hydrolase activity
CG14630	0.11	0.83	0.72	0.00029	oxidoreductase activity, **Wolbachia-dependent (d50)**
AGBE	-0.22	0.3	0.51	0.00037	glycogen synthesis, determination of lifespan, **Wolbachia-dependent (d50)**
Cyp12b2	0.65	0.15	-0.5	0.00044	heme binding, iron ion binding, **Wolbachia-dependent (d50)**
CG7675	1.1	0.12	-0.98	0.00044	oxidoreductase activity
CG3609	-0.67	6.00E-04	0.67	5.00E-04	oxidoreductase activity, **Wolbachia-dependent (d50)**
slmo	0.85	0.29	-0.56	0.00075	mitochondrion, locomotory behavior
shg	-0.39	-0.0022	0.39	0.00083	DE-cadherin homolog, formation of epithelia, [...]
RhoGAP68F	0.36	-0.21	-0.56	0.00092	recycling endosome, complex with Rab4
Gl	0.12	-0.21	-0.33	0.00099	eye morphogenesis
GlyS	-0.21	0.28	0.49	0.0014	response to starvation
Hsp22	3.3	1.1	-2.1	0.0018	unfolded protein response, determination of lifespan

CG3699	-1.3	-0.24	1	0.0018	fatty acid beta-oxidation, **Wolbachia-dependent (d50)**
CG9119	-0.84	-0.21	0.63	0.0021	zinc ion binding, **Wolbachia-dependent (d50)**
Wsck	0.56	0.14	-0.42	0.0021	immune response, differentiation
Dlic	0.21	-0.15	-0.37	0.0024	dynein heavy chain binding, axo-dendritic transport
CG4080	0.7	-0.26	-0.95	0.0027	protein polyubiquitination
Lsd-1	0.62	-0.21	-0.82	0.0029	lipid storage, lipid transport
Atg17	0.26	-0.15	-0.41	0.0036	response to starvation, glycophagy, neuromuscular function
mRpS5	0.31	-0.1	-0.41	0.0037	mitochondrial translation
CG5167	0.66	0.21	-0.45	0.0045	oxidoreductase activity
CG3091	0.51	0.1	-0.4	0.0045	
CG4199	0.4	0.047	-0.35	0.0046	oxidoreductase activity
CG15117	0.4	0.84	0.45	0.0049	carbohydrate metabolic process, extracellular space, **Wolbachia-dependent (d50)**
Akap200	0.094	-0.39	-0.48	0.0051	response to ethanol, neg. reg. of Ras signalling, **Wolbachia-dependent (d50)**
Khc	0.058	-0.22	-0.28	0.0051	microtubule motor activity

In the thorax, proteasome complex annotated proteins were significantly enriched, as well as proteins involved in lipid metabolism (Fig 49). The proteasome is crucial for turn-over of short-lived, damaged, and misfolded proteins, and declines with age (Vilchez et al. 2014). Previously, we showed that proteasome function is improved upon reduced IIS in the gut of 10d old long-lived flies, but we did not observe a difference in proteasome function in the thorax (Tain et al. 2017). Indeed, no differences in thorax age-divergent proteasome proteins were detected at day 10 in our time series proteomics data. However, at day 50, we observed the coordinated down-regulation upon reduced IIS of *Prosα 2/3/6*, structural elements of the 20S core particle, as well as members of the 19S regulatory particle (*Rpn13, Rpt5*)(File 3). Examining their expression change with age, we found that most of these proteins did not change with age in long-lived mutants, but were highly significantly up-regulated in w^{Dah} controls (File 3). This could be due to a lower prevalence of unfolded or misfolded proteins in $dilp^{2\text{-}3,5}$ flies, e.g. by increased availability of heat shock proteins (HSP's) that act as chaperones. Of seven detected, four were up-regulated with age in both genotypes (*Hsp22/26/60/83)*(File 3). *Hsp10 and Hsp23* were down-regulated with age, but the decline of *Hsp23* was significantly attenuated in long-lived flies. This suggests that reduction of IIS might not confer a higher proteasome activity in the thorax, but may act to relieve proteotoxic stress in the thoraces of older animals.

Table 3*: Thorax, top 30 most significant age-divergent proteins.*

protein	wDah	dilp2.3.5	interaction	adj.P.Val	Notes
Txl	0.49	-0.0086	-0.5	0.00042	protein-disulfide reductase, response to oxidative stress, **Wolbachia-dependent (d50)**
CG2233	0.99	1.8	0.81	0.00042	
Pepck	0.17	-0.3	-0.47	0.00079	gluconeogenesis
CG14630	0.52	1.1	0.59	0.0029	oxidoreductase activity
smp-30	-0.79	-0.21	0.58	0.0048	cell proliferation, cold acclimation
CG3699	-0.65	0.2	0.85	0.0061	fatty acid beta-oxidation, **Wolbachia-dependent (d50)**
Treh	-0.45	-0.089	0.36	0.0089	trehalase, stem cell population maintenance
CG33307	0.78	0.16	-0.61	0.0097	**Wolbachia-dependent (d50)**
CG18067	1.1	0.29	-0.83	0.0097	extracellular space
CG5791	0.99	0.35	-0.65	0.0098	extracellular space
Lsd-2	0.3	-0.074	-0.38	0.0098	lipid storage, lipid transport
Prosalpha3	0.28	0.008	-0.27	0.0098	proteasome core complex
Lsd-1	0.23	-0.18	-0.4	0.011	lipid storage, lipid transport
Got2	0.088	0.31	0.22	0.013	glutamate metabolism, **Wolbachia-dependent (d50)**
mRpS30	-0.2	0.1	0.3	0.013	mitochondrial translation
CG13315	2.3	1.4	-0.81	0.013	
Rpn13	0.58	0.29	-0.29	0.014	proteasome assembly, regulatory particle
CG6767	0.38	0.83	0.45	0.015	nucleoside metabolic process
Amy-d	-0.68	0.087	0.76	0.019	alpha-amylase activity, digestive enzyme
CG10512	0.11	-0.16	-0.27	0.02	oxidoreductase activity
CG9914	-0.79	-0.28	0.51	0.026	fatty acid metabolic process
Ms	0.29	-0.12	-0.41	0.026	negative regulation of smooth muscle contraction
CG30382	0.58	0.27	-0.31	0.026	proteasome core complex
CG8209	0.59	0.36	-0.23	0.026	nucleic acid binding, **Wolbachia-dependent (d50)**
Gclc	-0.053	0.18	0.23	0.036	glutathione metabolic process
Mal-B1	0.091	1.2	1.1	0.042	carbohydrate metabolic process
TER94	0.25	0.045	-0.21	0.043	autophagosome maturation
Fmo-2	-0.53	-0.14	0.39	0.043	oxidation-reduction process
Rpt5	0.27	0.093	-0.18	0.044	proteasome regulatory particle
Hsp23	-0.29	-0.74	-0.45	0.047	chaperone-mediated protein folding

3.3.5.6 The consistent late-life divergent insulin response highlights key metabolic enzymes

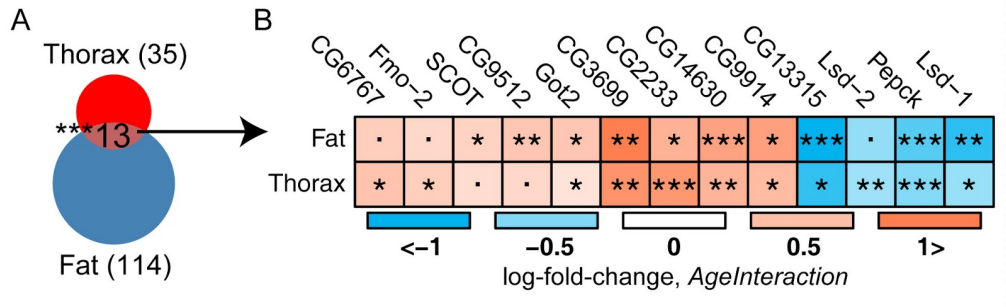


Fig 50*: **The consistent late-life divergent insulin response.***
***(A)** Euler diagram showing the number of late-life divergent proteins in the thorax, in the fat, and their overlap. Asterisks indicate significance of the overlap (hypergeometric test, P-Value ≤ 1e-8). **(B)** Heatmap showing the wild-type-normalized change of expression in dilp$^{2\text{-}3,5}$ flies with age (age interaction coefficient), for overlapping late-life divergent proteins in fat and thorax. Asterisks indicate significance of the limma moderated t-test (.p ≤ 0.1, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001).*

Lastly, we asked which proteins were classified as late-life divergent in both fat body and thorax. We identified 13 proteins that fit these stringent criteria, a highly significant overlap (hypergeometric test, P-Value ≤ 1e-8)(Fig 50A). We then checked the agreement of their age interaction coefficients, i.e. their change with age when normalizing for wild type expression. We found that all overlapping late-life divergent proteins were regulated in the same direction in fat and gut (Fig 50B). This included several important regulators of metabolism: the lipostasis proteins Lsd-1 and Lsd-2 (Miura et al. 2002; Grönke et al. 2003; Welte et al. 2005); Pepck, which catalyses a key step in gluconeogenesis and is under the control of insulin (O'Brien et al. 1990); Got2, an enzyme crucial for glutamate metabolism and therefore neurotransmitter function (Chase & Kankel 1987); and SCOT (Succinyl-CoA:3-ketoacid CoA transferase), which plays a key role in ketone body metabolism (Stern et al. 1956; Zhang et al. 2013). Additionally, we found Fmo-2 to be up-regulated with age compared to w^{Dah} flies; over-expression of this detoxification enzyme extends lifespan in *C.elegans (Leiser et al. 2015)*. Two proteins were found to be late-life divergent and *Wolbachia*-dependent at age 50d in both tissues: Got2 and CG3699. The seven proteins represented only by gene annotation IDs (CGs) are largely uncharacterized and present promising targets for further investigation.

3.4 Discussion

Reduced activity of the evolutionarily conserved insulin/IGF-like signalling network extends longevity and reduces age-associated pathologies, extending health and vitality (Fontana et al. 2010). Using a combination of tissue-specific proteomics screens in two independent models of insulin signalling, *dilp*$^{2\text{-}3,5}$ and mNSC-ablated flies, as well as transcriptomics screens of the *dilp*$^{2\text{-}3,5}$ model, we were able to add significantly to the previous understanding of the insulin response in *Drosophila*. We found that up to 39% of each tissue's proteome of mNSC-ablated flies or *dilp*$^{2\text{-}3,5}$ flies, and up to 22% of the tissue-specific *dilp*$^{2\text{-}3,5}$ transcriptome responded to reduced IIS. Using this multi-profiling approach we have compiled a resource with tissue-specific profiles of 6085 proteins (mNSC ablation)(Fig 3) and 7234 proteins (*dilp*$^{2\text{-}3,5}$)(Fig 13), over 11000 transcripts (*dilp*$^{2\text{-}3,5}$)(Fig 25) and 268 miRNAs (Fig 25). Separately, we have collected proteomics data of 5552 proteins in fat and thorax of *dilp*$^{2\text{-}3,5}$ flies at three time points (10d, 30d, 50d old flies)(Fig 44). We have published the first of these data sets in (Tain et al. 2017). Overall, these data represent a comprehensive resource of gene expression upon reduced insulin signalling in *Drosophila* and will be made available to the scientific community upon publication.

Our analysis thus integrated data from multiple connected omics data from two independent models of reduced insulin signalling. An overview of findings discussed in this thesis is provided in Table 4 (chapter 3.4.1, first subsection of this discussion).

Firstly, we leveraged the known dependence of IIS-mediated lifespan extension on dFOXO to identify and focus on three tissue-specific mechanisms of longevity assurance: Reduced translation in the fat body, increased proteasome assembly and proteostasis in the gut, as well as increased mitochondrial biogenesis and respiration in the fat body (Fig 9).

Secondly, we analysed the proteome of *dilp*$^{2\text{-}3,5}$ mutant flies and delineated which elements of the insulin response depended on the presence of the endosymbiont *Wolbachia*, which is required for lifespan extension in this background. We could confirm several conserved responses to reduced insulin signalling (Table 4). However, we also identified a *dilp*$^{2\text{-}3,5}$ specific response that protects against DNA damage in the fat body and prevents proliferation of transposons with age (Fig 22, Table 4).

Thirdly, we integrated matching proteomics and transcriptomics data of the *dilp*$^{2\text{-}3,5}$ model to screen the insulin response for mechanisms regulated on a posttranscriptional level. The translation machinery was affected by reduced IIS in all tissues except the brain, but through distinct mechanisms (Fig 32, Fig 33). Only in the fat body, both rate of translation and translation potential was decreased upon reduced IIS (Fig 8, Fig 31, Fig 34). We observed that ER-processed proteins in particular were down-regulated (Fig 37, Fig 38), implicating a reduction in co-translational protein import into the ER as an important mechanism of this reduced translation.

Lastly, we quantified the tissue-specific proteomes of 10d old as well as 30d and 50d old *dilp*$^{2\text{-}3,5}$ flies, taking advantage of the increased throughput and fidelity of TMT multiplexing proteomics (Thompson et al. 2003; Dayon et al. 2008). This enabled us to identify differences between long-lived mutants and wild type flies that arose during ageing, but are not manifested in young adult flies.

3.4.1 Summary: mechanisms of the tissue-specific insulin response

Table 4: Summary of tissue-specific responses to insulin signalling described in this thesis.

Process	Tissue	Response to reduced IIS	References
Splicing	Brain	Increase of protein expression, *dfoxo-dependent*	Fig 5, Fig 7
Mit. respiration	Fat	Improved function	(Tain et al. 2017)
		Increased proteins, *dfoxo-dependent*	Fig 5, Fig 9
	Gut	Reduced function	(Tain et al. 2017)
		Regulated proteins, *dfoxo-independent*	Fig 5, Fig 9
	Thorax	Increased mRNA expression	Fig 30
Muscle function	Thorax	Improved function	Fig 17
		Regulated gene expression	Fig 16, Fig 30, Fig 48
		Decrease of miRNA targeting muscle function	Fig 41
Intracellular trafficking	Brain	Regulation of actin filament polymerization, longevity-specific, Wolbachia-dependent	Fig 21
	Fat	Down-regulation of intracellular transport (lipid met.), age-divergent	Fig 49
Proteostasis	Gut	Increase of proteasome activity, dfoxo-dependent	(Tain et al. 2017)
		Regulated proteins, dfoxo-dependent, *Wolbachia-dependent*	Fig 10, Fig 21
		Increased proteins (alpha-mannosidases & lysosome), *part. Wolbachia-dependent*	Fig 16, Fig 20, Fig 30
		Decreased proteins, cytosolic protein folding	Fig 16
	Brain	Decreased proteins, cytosolic protein folding	Fig 16
	Thorax	Ameliorated increase of proteasome proteins with age	Fig 49
Epithelial integrity	Gut	Increase, *dfoxo-dependent*	(Tain et al. 2017)
		Regulated gene expression, dfoxo-dependent, *Wolbachia-dependent*	Fig 21
	Brain	Decrease of cuticle proteins, PTR	Fig 16, Fig 30
		Decrease of glial blood-brain-barrier proteins	Fig 16
Translation rate	Fat	Reduction, dfoxo-dependent, *Wolbachia-dependent*	Fig 34

		Regulation of initiation factors, *Wolbachia-dependent*	Fig 33
	Gut	Regulation of initiation factors	Fig 33
Translation potential	Fat	Reduction	Fig 31
		Reduced ribosomal proteins	Fig 5, Fig 16, Fig 30, Fig 32
		Global dampening of transcriptional IIS response	Fig 36
	Gut	Reduction, *part. Wolbachia-dependent*	Fig 31
ER-specific protein import & translation	Fat	Reduced ER/signal peptide/membrane proteins, *Wolbachia-dependent, PTR (mRNA opposite)*	Fig 16, Fig 30
Secretion	Fat	Reduction of extracellular proteins	Fig 16, Fig 30
DNA damage response	Fat	Increase, lowered transposon burden, $dilp^{2\text{-}3,5}$-specific, *Wolbachia-dependent, PTR*	Fig 22

3.4.2 Establishing the tissue-specific proteome of Drosophila

In recent years, the technology for proteomics has begun to reach the accessibility, sensitivity and reproducibility of genomic studies (Mann et al. 2013). Taking advantage of advancing technologies, we have performed an in-depth, tissue-specific proteomic screen in Drosophila, detecting over 6000 proteins in each of our two models of reduced insulin signalling (mNSC-ablated and $dilp^{2\text{-}3,5}$ mutant flies), over 40% of the predicted proteome. This dataset emphasizes the importance of tissue-specific profiling and provides a unique resource to guide future expression studies in Drosophila.

Our proteomic analysis of the mNSC ablation model proteome revealed that ~15% of the predicted fly proteome is remodelled in response to reduced IIS, with 2372 proteins showing significant up- or down-regulation in at least one tissue (Fig 15). In the more severe $dilp^{2\text{-}3,5}$ phenotype, this number rose to 4006 proteins (Fig 15). The RNA transcripts encoding over 60% of these proteins have not previously been identified as differentially regulated by reduced IIS in whole fly profiles. This may be due to the highly tissue-specific nature of the changes, with only two proteins (mNSC ablated flies) or 22 proteins ($dilp^{2\text{-}3,5}$) being regulated in all tissues.

Confirming the quality of this dataset, proteins associated with several IIS-mediated processes, including development, growth, sleep, lipid metabolism, translation and adult lifespan, were detected and differentially regulated in mNSC-ablated flies (Fig 5). We were able to pinpoint these responses to specific tissues (Fig 5). For example, global reduction in translation can extend lifespan in C. elegans and Drosophila (Hansen et al. 2007; Pan et al. 2007; Wang et al. 2014) and is an evolutionarily conserved response to reduced IIS (McElwee et al. 2007; Stout et al. 2013; Essers et al. 2016), and our study showed this response to be specific to the fat body.

Our dataset also identified possible novel mediators of responses to reduced IIS. For example, our proteomic analysis suggested a gut-specific regulation of proteasomal function in response to reduced IIS, which was confirmed by finding a corresponding gut-specific proteasomal phenotype (Tain et al. 2017). We have also characterized dfoxo-dependent changes in protein expression in IIS mutant flies, separating those changes associated with IIS-mediated longevity from those changes associated to other IIS-mediated phenotypes. Furthermore, we determined which IIS responsive protein-coding genes contain predicted dFOXO-binding motifs, identifying possible direct and indirect targets of dFOXO (Tain et al. 2017). Additionally, we examined transcriptional changes in several regulated candidate genes associated with mitochondria and the proteasome (Tain et al. 2017). Some candidate genes were regulated in a dfoxo-dependent manner, consistent with a direct regulation by dFOXO; however, many did not reflect the changes seen at the protein level, suggesting indirect regulation by dFOXO, possibly through post-transcriptional and/or post-translational regulation (Tain et al. 2017). Importantly, we have thus identified longevity-associated changes in the proteome of IIS mutant flies and the processes that they regulate, and we have experimentally demonstrated the role of some of these processes in extension of lifespan.

3.4.3 Tissue-specific regulation of mitochondrial number and respiration is a regulator of the IIS longevity response

Mitochondrial respiration and the production of reactive oxygen species (ROS) have been proposed as a mechanism of ageing (Bratic & Larsson 2013). According to this theory, ROS are produced as a by-product of mitochondrial respiration and on interaction with cellular macromolecules they induce damage (Harman 1956). However, the link between mitochondria and ageing is not simple, and the validity of this theory has been questioned (Bratic & Larsson 2013). Seemingly in agreement with reduced ROS being beneficial, reducing mitochondrial respiration extends lifespan in several organisms (Dillin et al. 2002; Rea et al. 2007; Copeland et al. 2009). In contrast, many pro-longevity interventions are associated with increased activation of mitochondrial respiration (Evans & Scarpulla 1990; Virbasius et al. 1993; Ristow & Schmeisser 2011), including models of dietary/caloric restriction (Mootha et al. 2003; Baker et al. 2006) and IIS, as found here and in *C. elegans* (Zarse et al. 2012). We and others found that declining mitochondrial number or function is associated with ageing (Fig 46)(Yen et al. 1989; Shigenaga et al. 1994; Hebert et al. 2015). Thus, increasing or maintaining mitochondrial biogenesis with age may be beneficial.

We found increased, fat body-specific respiration that was causal in the longevity of IIS mutants (Tain et al. 2017). Furthermore, we and others (Rera et al. 2011) have shown that increasing mitochondrial biogenesis extends lifespan. Although we found the effect to be confined to the fat body, while Rera et al. found an effect in both the fat body and the gut, the discrepancy is

possibly due to the use of different fly strains. We also found that the increased fat body-specific respiration and longevity of IIS mutant flies are dependent on PGC1-α in the fat body, indicating that PGC1-α is a downstream mediator of IIS, not only for growth (Tiefenböck et al. 2010; Mukherjee & Duttaroy 2013), but also, as found here, longevity.

How increased respiration can, directly or indirectly, extend lifespan remains unclear. Growing evidence suggests that ROS, produced by mitochondrial respiration, is increased by pro-longevity interventions and may act as signalling molecules to activate the expression of pro-longevity genes (Ristow & Schmeisser 2011). On the other hand, inhibition of target of rapamycin (TOR) in yeast increases respiration, but reduces ROS production, due to increased electron transport chain subunit expression and increased efficiency of electron transport, preventing a protracted electron transit time and ROS formation (Bonawitz et al. 2007). Furthermore, in flies, increased mitochondrial biogenesis protects against high fat diets, preventing lipid accumulation and heart dysfunction (Diop et al. 2015). To determine the exact mechanisms involved, analysis of tissue-specific determination of ROS production and electron transport chain complex status under impaired IIS and other pro-longevity interventions will be needed.

Interestingly, although not associated with longevity, we also found that respiration in the gut is decreased under reduced IIS, independently of *dfoxo* activity. Our analysis suggested that this response was associated with down-regulation of Complex III protein subunits (Tain et al. 2017). When Complex III subunits are ubiquitously reduced throughout development and adulthood, flies become considerably longer lived and produce fewer offspring (Copeland et al. 2009). However, adult-specific expression fails to extend lifespan, yet flies remain less fecund (Copeland et al. 2009). Since the down-regulation of respiration we observed on reducing IIS in the gut was not associated with increased lifespan, and occurred independently of *dfoxo*, the reduction in gut respiration may underlie the reduced fecundity of IIS mutants.

3.4.4 Tissue-specific proteostasis is a regulator of the IIS longevity response

Gradual loss of proteostasis is considered a hallmark of ageing (López-Otín et al. 2013). We found that, under reduced IIS, *Drosophila* tissue specifically regulate proteostasis, an evolutionarily conserved response that correlates with longevity. In flies, the age-associated decline of proteostasis has been linked to declining proteasome function (Tsakiri et al. 2013), whilst maintaining proteasome function is associated with longevity (Chondrogianni et al. 2000; Pérez et al. 2009; Vilchez et al. 2012; Ungvari et al. 2013). In yeast, *C. elegans*, and *Drosophila*, ubiquitous over-expression of proteasomal subunits can extend lifespan, possibly through increased maintenance of the cellular proteome (Q. Chen et al. 2006; Tonoki et al. 2009; Kruegel

et al. 2011; Vilchez et al. 2012). Increasing the maintenance of the cellular proteome and extending life span can also be achieved by reducing translation (Hansen et al. 2007; Bonawitz et al. 2007; Wang et al. 2014). We found in *Drosophila* that increased proteasome activity correlated with longevity specifically in the gut. We were able to recapitulate increased proteasome activity by gut-specific over-expression of RPN6, which was sufficient to extend lifespan in wild-type flies. Importantly, we could also show that increased proteasomal activity is required for the longevity of IIS mutants. Understanding how manipulation of one, or relatively few, protein subunits affects the function of a large protein complex is important for translating functional studies towards amelioration of the effects of human ageing. While each subunit plays a role, only some subunits, that is the catalytic subunits, are essential for proteasome activity. However, Rpt3, Rpt6 and RPN6 play an equally vital role, those of opening the catalytic core and assembly maintaining structural integrity of the proteasome (Murata et al. 2009; Tian et al. 2011; Pathare et al. 2012; Sokolova et al. 2015). RPN6 over-expression also influences lifespan (Vilchez et al. 2012), and we suggest through increased 26S proteasomal assembly, which recapitulates the phenotype seen in IIS mutant flies. It will therefore be important to determine whether similar regulation occurs in the tissues of long-lived mammalian IIS models.

3.4.5 Tissue-specific epithelial barrier integrity is improved under reduced IIS

Loss of gut barrier integrity is associated with age and makes old flies more susceptible to invasion by pathogens (Rera et al. 2012). Improving gut integrity function is therefore beneficial to lifespan. Our proteomics analyses suggested that epithelial integrity in the guts was improved in response to reduced IIS: firstly, we observed the regulation of key proteins involved in the regulation of septate junction permeability, specifically down-regulation of Yurt and up-regulation of Lac (Laprise et al. 2006; Llimargas et al. 2004). Secondly, we saw a concerted up-regulation of chitin-based cuticle proteins that form part of the peritrophic matrix. We hypothesized that the improvement of proteostasis we observed in the gut (Tain et al. 2017) may help preserve the proper maintenance of these epithelial components with age, and thus mediate the beneficial effect of reduced IIS on gut integrity. Indeed, over-expression of the proteasome regulatory particle component Rpn6 was sufficient to increase gut integrity and extend lifespan (Tain et al. 2017).

3.4.6 Unexpected regulation of the insulin target Tobi in the long-lived gut

A previous study in whole flies suggested that dILP activity in mNSCs is necessary for induction of the insulin and Akh target Tobi (Buch et al. 2008). Surprisingly, we observed a *dfoxo*-dependent up-regulation Tobi in the gut of both our insulin models (Fig 14).We also observed a 8.3-fold reduction of Tobi in dfoxo-null mutants, confirming that loss of *dfoxo*

reduces tobi levels dramatically (Buch et al. 2008). Ablation of mNSC's in this background reduced Tobi by an additional 27%. The up-regulation of Tobi despite the loss of dILPs may be the result of a subtle difference in the mNSC-ablation model used by our study: Buch et. al used a dILP3 promoter as opposed to the dILP2 promoter employed in our model. Supporting the importance of these differences, Buch et al. do not observe the reduction in fecundity that occurs in both *dilp*$^{2\text{-}3,5}$ flies and our mNSC-ablation model (Broughton et al. 2005). This suggests that seemingly small differences between models can influence the conclusions drawn from these models. Cross-referencing between heterogeneous models at even larger scales may therefore be beneficial in assessing the robustness of elements of the insulin response.

3.4.7 Dissecting the role of post-transcriptional regulation upon reduced IIS

Metabolic signalling is known to be affected by post-transcriptional regulation, from translation (Afschar et al. 2016; Kaletsky et al. 2016; Selman et al. 2009) to RNA editing (Gan et al. 2006; Yang et al. 2010), regulation by miRNA (Poy et al. 2004; Trajkovski et al. 2011; Karolina et al. 2011), splicing (Mazin et al. 2013; Heintz et al. 2016), mRNA stability (Fernyhough et al. 1989; Goalstone & Draznin 1999; Hsu et al. 2011), protein stability (Rice et al. 1993; Haruta et al. 2000; Hong et al. 2003; Facchinetti et al. 2008), and post-translational modifications (Vosseller et al. 2002). Recent studies in *Drosophila* have investigated tissue-specific regulation upon IIS modulation separately on the level of the transcriptome (Teleman et al. 2008; Alic et al. 2011) and the proteome (Tain et al. 2017). Integrated analyses in other organisms have highlighted a rich landscape of tissue-specific post-transcriptional regulation with age (Ori et al. 2015), but no such studies exist for the regulation of IIS. Here, we quantify the transcriptomes and proteomes in each of four fly tissues, in long-lived dilp$^{2\text{-}3,5}$ knock-out flies and their controls. Taking advantage of advances in proteomics, we were able to identify proteins originating from 48% of protein-coding *Drosophila* gene models across all tissues. Our data were highly reproducible and tissues were distinguished by characteristic functions. Responses to reduced insulin signalling in *dilp*$^{2\text{-}3,5}$ mutants confirmed known elements of the Drosophila insulin response such as reduced translation (Essers et al. 2016), increased detoxification (Amador-Noguez et al. 2007), and modulation of proteostasis (Bai et al. 2013; Tawo et al. 2017). Integrating our transcriptomics and proteomics quantifications enabled us to delineate responses that are induced on either one or both levels of regulation, and identify post-transcriptionally regulated processes.

In addition, we quantified miRNA and analysed the expression of their predicted targets. We could quantify 316 miRNA in at least one tissue (73% of annotated miRNA) with high reproducibility. Of these, 36% were regulated upon reduced insulin signalling. Most miRNA were regulated in the thorax muscle tissue, where they were reduced in long-lived *dilp*$^{2\text{-}3,5}$ mutants. It has been demonstrated before that particular families of miRNA can regulate muscle

development (Sokol & Ambros 2005), performance (van Rooij et al. 2009), and neuromuscular signalling (Tsurudome et al. 2010). Here, we report a systemic decrease in miRNA expression in the muscle upon reduced IIS that correlates with improved muscle function in aged flies. However, we could detect no coordinated regulation in the other tissues. This may be due to lack of power of our miRNA assay, which could be overcome in future studies by increasing replication or improving the precision of the normalization reference. However, we conclude that while fat, brain, and gut may be regulated by miRNA to some degree, it is not a major form of regulation upon reduced IIS.

3.4.8 Tissue-specific regulation of translation in response to reduced IIS

Translation of mRNA into protein is a crucial entry point for post-transcriptional regulation of gene expression (Sonenberg & Hinnebusch 2009), and reduced translation is a conserved response to IIS that is associated with longevity (Afschar et al. 2016; Kaletsky et al. 2016; Selman et al. 2009; Essers et al. 2016). We observed regulation of the translation machinery in response to reduced IIS in gut, fat, and thorax. While the pool of translated mRNA was reduced in both fat body and gut, rate of translation was reduced only in the fat body of long-lived *dilp*$^{2\text{-}3,5}$ mutants. Leveraging our transcriptomic and proteomic data, we found that this decrease in translation rate post-transcriptionally reduced the expression levels of membrane and secreted proteins, i.e. proteins processed in the endoplasmic reticulum (Caro & Palade 1964; Walter & Johnson 1994; Rapoport 2007). We confirmed this pattern of regulation in an independent model of reduced insulin signalling, mNSC-ablated flies (Broughton et al. 2005). Translation of these proteins can be inhibited in response to ER stress through the ER unfolded protein response and phosphorylation of *eIF2 (Sood et al. 2000; Harding et al. 2000; Boyce et al. 2005)*. However, translation in the ER is also governed by the rate of co-translational protein import through targeting of nascent polypeptides to ER translocons by SRP's, and their subsequent import and folding (Walter & Blobel 1981; Keenan et al. 2003). We found no evidence of an increase of the UPR in the fat body of our long-lived flies, but did detect regulation of the ER import machinery indicative of a reduced ER capacity. Furthermore, we found a significant overlap of proteins targeted to the ER and those that were regulated in a *Wolbachia*-dependent manner, and those proteins were enriched for functions relating to the humoral immune response. The fat body is the primary immune tissue of the fly (Hoffmann 2003) and *Wolbachia* has been observed co-localizing with ER and Golgi membranes (Fischer et al. 2014; Voronin et al. 2004). This raises the question as to whether the endosymbiont may intervene directly in the processing of secreted proteins to suppress the host immune response, and whether this phenotype is exacerbated by the increase of *Wolbachia* titers under conditions of reduced IIS (Serbus et al. 2015)(Fig 18). In summary, we report that the fat body-specific reduction in translation in response to reduced IIS is mediated by a decreased capacity for

co-translational protein import into the ER, and thus specifically affects secretory and membrane proteins.

3.4.9 Post-transcriptional regulation of the fat body DNA damage response

We discovered that key elements of the DNA damage response were strongly induced in the fat body of long-lived *dilp*$^{2\text{-}3,5}$ flies—but not the independent mNSC-ablation model of insulin signalling—and that this response was post-transcriptionally regulated. We experimentally confirmed a reduction of known markers of DNA damage such as *H2Av* (Rogakou et al. 1998; Madigan et al. 2002) in old long-lived mutants. Genomic instability, which can result from DNA damage (Celeste et al. 2002), leads to deregulation of transcription and has been associated with ageing across a wide range of model organisms including *Drosophila* (Wood et al. 2010; Maxwell et al. 2011). Combined with the dysregulation of DNA surveillance mechanisms with age, this can lead to the deleterious mobilization of transposable elements (Maxwell et al. 2011; Dennis et al. 2012; De Cecco et al. 2013; Van Meter et al. 2014; H. Chen et al. 2016). However, it has been shown that transposon de-repression and loss of gene silencing with age are ameliorated by dietary restriction (Jiang et al. 2013). In agreement with these findings, we found that transposon burden did not change in the fat body of aged *dilp*$^{2\text{-}3,5}$ mutants compared to their young (10d) counterparts, but was increased in the wild type. Additionally, expression of *piwi*, a protein known to suppress transposon mobilization (Sarot et al. 2004; Ross et al. 2014), was dramatically increased in long-lived flies. A recent study in *Drosophila* revealed that *piwi* mutants are starvation sensitive and accumulate more DNA damage with age, in part due to transposon de-repression in the fat body (Jones et al. 2016). Our study shows that a direct intervention to reduce IIS in the *Drosophila* fat body by *dilp*$^{2\text{-}3,5}$ knockout post-transcriptionally affects DNA maintenance mechanisms to maintain transposon repression and potentially delay genomic instability in ageing flies.

3.4.10 Intersections of Wolbachia host manipulation and insulin signalling

Wolbachia is a ubiquitous endosymbiont that is present in a wide range of arthropod species, with phenotypes ranging from parasitism to commensalism and symbiosis (Werren & Windsor 2000). *Wolbachia* proliferation depends on the status of TOR signalling and IIS; lowered IIS activity is associated with increased proliferation (Ikeya et al. 2009; Serbus et al. 2015). However, its effect on lifespan and metabolism is varied (Min & Benzer 1997; Toivonen et al. 2007), as are its interactions with the host immune system (Teixeira et al. 2008; Pietri et al. 2016; Sicard et al. n.d.; Faherty & Maurelli 2008). Interestingly, *Wolbachia* surface protein (WSP) may prevent caspase-3 activation and thus block apoptosis, even in human cell lines (Faherty & Maurelli 2008; Bazzocchi et al. 2007). In *dilp*$^{2\text{-}3,5}$ mutants, the presence of the endosymbiont is necessary for the longevity phenotype (Grönke et al. 2010), but the causal mechanism is not

known. Here, we investigated insulin signalling tissue-specifically both in the presence and absence of *Wolbachia*, identifying parts of the insulin response that depend on the endosymbiont (Fig 19, Fig 20). We found several processes that were *Wolbachia*-dependently regulated upon reduced IIS: gut proteostasis and integrity, actin polymerization in the brain, or cell adhesion and myosin binding in the thorax (Fig 21); regulation of cuticle proteins in gut and brain, and DNA replication machinery in the fat body (Fig 20). In some instances, such as regulation of the actin cytoskeleton, mechanisms of direct manipulation by *Wolbachia* have been reported before (Newton et al. 2015; Sheehan et al. 2016). We could furthermore associate *Wolbachia*-dependent proteins to the reduced translation of ER-processed secreted proteins in the fat body (Fig 39), supporting reports that *Wolbachia* may interact with the ER of its host (Cho et al. 2011; Fischer et al. 2014; Voronin et al. 2004). Additionally, considering its synergistic effect on lifespan in $dilp^{2\text{-}3,5}$ mutants, *Wolbachia* may counteract negative side-effects of lower-than-optimal insulin signalling (Cohen & Dillin 2008). In particular, *Wolbachia's* ability to block apoptosis (Faherty & Maurelli 2008; Bazzocchi et al. 2007) may be beneficial under conditions of reduced IIS, supplanting the role of insulin in preventing apoptosis (Kang et al. 2003; Porras et al. 2003; Kooijman 2006). Unfortunately, we could not assess caspase-3 phosphorylation. Future studies may augment the data gathered in this project by performing phosphoproteomics measurements. Overall, it is not evident from our results whether *Wolbachia* intervenes directly in multiple downstream processes of IIS, or merely enables beneficial effects of reduced IIS through a far more limited intervention. Determining the precise physical mechanism of *Wolbachia's* interaction with its *Drosophila* host may highlight or confirm key steps of IIS as well as point out a way to manipulate them, making this a promising avenue for future research.

3.4.11 Age-divergent mechanisms of the longevity-associated insulin response

The majority of analyses presented in this thesis focuses on data from samples of 10d old, long-lived flies with reduced insulin signalling. This may, by association, point towards causal mechanisms of reduced IIS that mediate longevity (Tain et al. 2017). However, these proteomics and transcriptomics screen do not directly connect the reduced IIS response to processes changed during ageing. To address this, we carried out a large tissue-specific proteomics screen of young, middle-aged, and old $dilp^{2\text{-}3,5}$ flies and their controls (10d, 30d, and 50d). Our quantification by TMT multiplex proteomics was competitive in terms of identified proteins with our previous proteomics (Fig 44, Fig 3, Fig 13). Importantly, multiplexing enabled us to avoid technical variability between biological groups, allowing direct and precise comparisons of all phenotypes and ages (Table 1). Time series of whole fly transcriptomics have been used to profile gene expression changes during age, as well as additional conditions such as dietary restriction or upon dietary switch (Pletcher et al. 2002; Whitaker et al. 2014; Carlson et al. 2016). More recently, a proteomics time series of *Drosophila* heads highlighted age-related

proteome changes, such as increased expression variance and decreased proteostasis in older flies (Brown et al. 2018); however, this study only identified 1281 proteins. In the preliminary results presented in this thesis, we quantified 5552 proteins across two tissues (the fat and the thorax) with high replication and accuracy. Measurements from two additional tissues, the brain and the gut, were generated and are currently being analysed. While our study was limited to three time points, compared to seven in the study by Brown et al., we were thus able to significantly extend the scope of age-dependent and tissue-specific proteomics profiles.

Our preliminary results identified 1336 and 1416 proteins changed with age in the thoraces and fat bodies of wild type flies (File 3). The tissue profiles were correlated with each other (Fig 45), but nonetheless showed important differences in their enriched functional annotations and global ageing trajectories (Fig 46, Fig 47). Similarly, analysis of the insulin response across ages indicated a stable but tissue-specific response to reduced IIS (Fig 48). Importantly, we delineated the age-divergent insulin response, i.e. elements of the response to reduced IIS in both tissues that grew over time (Fig 49)(File 3). Functional analysis of the age-divergent insulin response pointed towards intracellular trafficking (in the fat) and proteostasis (in the thorax) as key divergent functions affected in an age-dependent manner. We furthermore confirmed the regulation of important regulators of lipid and carbohydrate metabolism in the most stringent subset of this analysis: proteins that were regulated upon reduced IIS in an age-divergent manner in both fat and thorax (Fig 50). Among these, we highlight several co-regulated but unannotated proteins for future study (Fig 50).

3.5 Methods

3.5.1 Differential expression analysis of proteomics data using limma

PCA analysis was carried out using the FactoMineR R package (Lê et al, 2008). The LIMMA package (Bioconductor suite, R) was used for differential expression analysis (Schwämmle et al, 2013). Multiplicity-corrected P-values were calculated using the Benjamini-Hochberg procedure and significance determined using an adjusted P-value cut-off of 0.1.

3.5.1.1 mNSC Ablation analysis

Experimental groups were specified as follows: (A) w^{Dah}, (B) $dfoxo^{\Delta 94}$ mutant, (C) InsP3-Gal4/UAS-rpr and (D) InsP3-Gal4/UAS-rpr, $dfoxo^{\Delta 94}$. A linear model was fitted for each tissue using the limma R package, modeling protein expression for each group of replicates. Hypotheses were tested as linear combinations against the null hypothesis of no change: ablation-induced changes in the wild-type background (C - A), ablation-induced changes in the $dfoxo^{\Delta 94}$ background (D vs. B) and their interaction term ((C–A) - (D–B)). The subset of proteins differentially expressed in C vs. A was further subdivided into two categories: proteins that changed in response to mNSC ablation in a *dfoxo*-dependent manner, and those that behaved similarly regardless of genetic background were *dfoxo*-independent. Foxo-dependent: Differential in C vs. A and the interaction term. If a protein was regulated in the same direction in C vs. A and D vs. B, it was required to show a stronger response in the first contrast. Foxo-independent: Differential in C vs. A. Protein expression levels were required to be equivalent with regard to differential expression in the interaction term. In any given tissue, a protein was identified as equivalent if the 95% CI of its log2 fold change fell within an interval ([−t; t]). The parameter t was determined as minimum log-fold change of any significantly differentially expressed protein detected in the interaction term of each respective tissue: brain (t = 0.072), fat body (t = 0.177), gut (t = 0.155) and thorax (t = 0.216). Additionally, this subset included proteins found to be significant in the interaction term if they showed a stronger response to mNSC ablation in equal direction to that in the dfoxo Δ94 background compared to the wild-type background.

3.5.1.2 $Dilp^{2\text{-}3,5}$ analysis

Experimental groups were w^{Dah} controls, long-lived $dilp^{2\text{-}3,5}$ mutants, and *Wolbachia*-negative controls w^{Dah}/Wol- and $dilp^{2\text{-}3,5}$/Wol-. A linear model was fitted for each tissue, modeling protein expression for each group of replicates. Hypotheses were tested as linear combinations against the null hypothesis of no change: $dilp^{2\text{-}3,5}$-induced changes ($dilp^{2\text{-}3,5}$ - w^{Dah}) and the interaction of *Wolbachia* status and $dilp^{2\text{-}3,5}$ (($dilp^{2\text{-}3,5}$ - w^{Dah}) - ($dilp^{2\text{-}3,5}$/Wol- - w^{Dah}/Wol-)). The subset of significant $dilp^{2\text{-}3,5}$ induced changes was further subdivided into two categories: firstly, proteins

that changed in response to *dilp*$^{2\text{-}3,5}$ knockout in a *Wolbachia*-dependent manner, i.e. were significant in the interaction term. Secondly, those that behaved similarly regardless of *Wolbachia* presence. Protein expression levels were required to be equivalent with regard to differential expression in the interaction term. In any given tissue, a protein was identified as equivalent if the 95% CI of its log2 fold change fell within an interval ([−t; t]). The parameter t was determined as the mean of median absolute log-fold-changes in the interaction term of each tissue (t = 0.085).

***Dilp*$^{2\text{-}3,5}$ time series analysis**

Experimental groups were *w*^{Dah} controls, long-lived *dilp*$^{2\text{-}3,5}$ mutants, and *Wolbachia*-negative controls *w*^{Dah}/Wol- and *dilp*$^{2\text{-}3,5}$/Wol- Samples were collected at three different time points *T*: 10d-old flies, 30d-old flies (only *Wolbachia*-positive), and 50d-old flies.

The first linear model modeled protein expression for each group of replicates. To account for sample processing batches, the batch was included as a confounding factor in the linear model. Hypotheses were tested as linear combinations against the null hypothesis of no change: *dilp*$^{2\text{-}3,5}$-induced changes (*dilp*$^{2\text{-}3,5}{}_{T}$ - *w*$^{Dah}{}_{T}$) at each age *T*, and the interaction of *Wolbachia* status and *dilp*$^{2\text{-}3,5}$ at day 50 ((*dilp*$^{2\text{-}3,5}{}_{50d}$ - *w*$^{Dah}{}_{50d}$) - (*dilp*$^{2\text{-}3,5}$/Wol-$_{50d}$ - *w*^{Dah}/Wol-$_{50d}$)). Proteins were classified as *Wolbachia*-dependent at day 50 if they were significantly *dilp*$^{2\text{-}3,5}$-induced and significant in the interaction term. Additionally, we fitted a second linear model to *Wolbachia*-positive samples across all time points. Here we fitted a polynomial contrast with linear (*L*) and quadratic trends (*Q*) to each genotype, describing its mean expression depending on time as an ordinal variable. We defined different groups of proteins depending on their linear change with age: proteins that changed monotonously with age in *w*^{Dah} flies (L_{wDah}) and proteins that changed monotonously with age in *dilp*$^{2\text{-}3,5}$ mutants (L_{dilp}). We then defined an interaction contrast L_{int} to identify proteins whose linear trend differed significantly between the two genotypes (L_{dilp} - L_{wDah}). Combining the results from both linear models, we classified proteins as late-live divergent if their ageing trajectories differed significantly, as tested by L_{int}, and the absolute value of *dilp*$^{2\text{-}3,5}{}_{50d}$ - *w*$^{Dah}{}_{50d}$ was greater than the absolute value of *dilp*$^{2\text{-}3,5}{}_{10d}$ - *w*$^{Dah}{}_{10d}$. Proteins were classified as late-life *Wolbachia*-dependent if they were both late-life divergent and significant for the interaction of *Wolbachia* status and *dilp*$^{2\text{-}3,5}$ genotype at day 50.

3.5.2 Prediction of dFOXO binding sites

We used FIMO, MEME suite (Bailey et al, 2009) to identify genes whose transcription start site was within 1,000 bp of a dFOXO binding motif, using the default P-value threshold (1e−4). Sequences corresponding to genes of measured proteins were extracted from the BDGP6 reference genome, while the dFOXO binding motif was taken from Fly Factor Survey (Zhu et al, 2011). We then mapped FIMO hits to proteins whose expression changed under reduced IIS conditions or those that were identified as dfoxo-dependent and used a hypergeometric test to

evaluate the significance of the overlap. Selection and background were limited to the set of genes corresponding to measured proteins.

3.5.3 Dilp2-3,5 model transcriptomics data analysis

Ribo-minus RNAseq was carried out on all experimental groups of the Dilp model, in triplicate. In addition, the polysome fractions of *dilp*$^{2\text{-}3,5}$*/Wol+* and *w*^{Dah}*/Wol+* in fat and gut were isolated and sequenced. Transcripts were then mapped to the BDGP6.28 reference genome using tophat2 and counted via summarizeOverlaps (part of the Bioconductor R package GenomicAlignments). Differential expression analysis was carried out tissue-wise using a two-factor linear model with interaction effect via the DeSeq2 R package.

3.5.4 Gene ontology enrichment analysis of mNSC-ablation proteomics

topGO was used for GO term enrichment analysis and gene annotation performed using the Bioconductor annotation package org.dm.e.g.db. To identify enriched GO terms, the one-sided elim Fisher procedure was used ($\alpha \leq 0.05$)(Alexa et al. 2006). The enrichment score of a GO term is defined as *log*2 (#*Detected significant genes*/#*Expected significant genes*). For individual tissue-specific contrasts, the subsets of differentially expressed proteins were tested for term enrichment against a tissue-specific reference background. For PCA map functional annotation, the top 5% and lowest 5% quantiles of proteins according to their contributions to each dimension were tested for GO term enrichment, against the background of all detected proteins. The plot shows the 10 most significant terms with a minimum of 10 significant proteins, for each direction of the two dimensions.

4 Regulatory networks of mouse stem cell pluripotency

4.1 Introduction

Mouse embryonic stem cells (mESCs) are isolated from the naive epiblast, a group of cells forming at day 4 to 4.5 of embryogenesis as part of the inner cell mass of the blastocyst (E4.5)(Fig 51)(Kalkan & Smith 2014; Bradley et al. 1984). The distinguishing characteristic of mESC's is their naive pluripotency, i.e. the ability to commit to any cell fate (Nichols & Smith 2009).

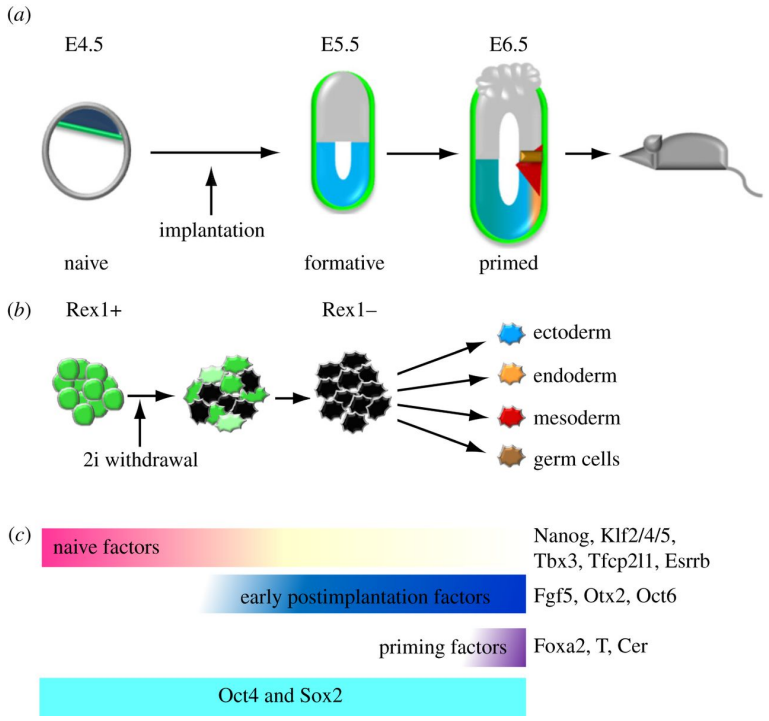


Fig 51 *(Kalkan & Smith 2014)*

The naive epiblast is poised to respond to differentiation signals from extraembryonic tissue, while being shielded from extraembryonic differentiation (Beddington & Robertson 1999). To maintain mESCs in culture, they have to similarly be shielded from differentiation cues, while activating pathways that promote self-renewal (Kalkan & Smith 2014). Until recently, fetal calf serum (FCL) derived medium supplemented with leukemia inhibitory factor (LIF)(Moreau et al. 1988) was commonly used to maintain mESCs in culture. LIF is an interleukine-6 cytokine that contributes to the ability of cells to self-renew and proliferate through at least three downstream branches: Firstly, by activation of the Jak-Stat3 pathway, which is sufficient to maintain self-renewal of mESCs (Niwa et al. 1998). Secondly, by activation of the MAPK pathway,

which is involved in mESC differentiation; inhibition of MAPK by MEKi promotes, but is not necessary for self-renewal (Matsuda et al. 1999; Burdon et al. 1999). Thirdly, by activation of the PI3K-Akt-mTOR pathways, which promote mESC proliferation (Welham et al. 2007). However, LIF alone is only effective for maintaining pluripotency in specific genetic backgrounds (Batlle-Morera et al. 2008). The development of 2i/3i medium (Ying et al. 2008) addressed this issue by adding small molecule inhibitors: firstly, PD184352, which inhibits Fgf/Erk signalling (Allen et al. 2003). It thereby prevents activation of *Erk1/2* by *Fgf4*, a developmental cue to exit self-renewal and commit to differentiation (Kunath et al. 2007). Secondly, the *Gsk3* inhibitor CHIR99021 (Gsk3i), which mimics Wnt signalling activation (Ying et al. 2008). *Wnt* activity blocks ubiquitination and therefore degradation of *β-catenin*, while Gsk3i directly blocks the phosphorylation of *β-catenin* that leads to its ubiquitination (Shy et al. 2013; Li et al. 2012). Consequently, the *β-catenin* target *Tcf7l1* (aka *Tcf3*) is repressed, and the repression *Tcf7l1* target genes including the naive pluripotency factors *Nanog* and *Esrrb* is relieved (Martello et al. 2012). Simultaneously, *β-catenin* activates a broad transcriptional programme by binding Tcf1/Lef (Cadigan & Waterman 2012). Gsk3i also modulates other pathways involved in pluripotency, possibly due to the wide range of *Gsk3* substrates (Wu et al. 2013; Doble & Woodgett 2003). A third inhibitor, SU5402, targets Fgf receptor (Fgfr1), and its effect is therefore largely redundant with PD184352 (Grand et al. 2004; Ying et al. 2008). When kept in 2i/LIF medium, mESC's retain their pluripotency even after extended passaging and are capable of self-renewal (Ying et al. 2008; Nichols & Smith 2012). Interestingly, the 2i inhibitors are sufficient to maintain undifferentiated mESCs even in the absence of LIF, though at the cost of some cell viability (Ying et al. 2008).

Differentiation of cultured mESCs can be triggered by the withdrawal of 2i/LIF or 2i and placement in a defined medium, such as N2B27 for neural fate commitment (Ying et al. 2003). Overall, their experimental malleability has positioned mESCs as an ideal model system for forward genetic screens, resulting in the discovery of the central pluripotency factor *Nanog*, as well as the development of induced pluripotent stem cells (Chambers et al. 2003; Takahashi & Yamanaka 2006). Since then, many markers for different stages of pluripotency have been discovered, resolving the circuitry of pluripotency with increasing depth (Fig 51)(Kalkan & Smith 2014; Young 2011; Bradley et al. 1984). However, the precise mechanisms that trigger the exit from naive pluripotency and initiate early differentiation are not yet fully understood.

The recent creation of haploid mESC's (Elling et al. 2011; Leeb & Wutz 2011) has greatly facilitated the discovery of new regulators of pluripotency and self-renewal by forward genetic screen (Leeb et al. 2014). While in diploid cells the phenotypes of random mutations can be obscured by a redundant second allele, haploid cells allow high-throughput mutagenic screens (Leeb et al. 2014). To screen for potential regulators of mESC self-renewal and pluripotency in an unbiased and genome-wide manner, our collaborators generated haploid mESC mutant cells using a piggyBac transposon system (Cadiñanos & Bradley 2007), which were kept in 2i medium.

Transfected cells were then placed in N2B27 medium and screened for differentiation defects. One way to assess the differentiation status of mESCs is by measuring the expression of the green fluorescent protein (GFP) tagged pluripotency marker Zfp42/Rex1 (Hosler et al. 1989; Rogers et al. 1991; Eiges et al. 2001; Kalkan & Smith 2014)(Fig 51). As *Zfp42* expression decreases rapidly in cells upon the start of differentiation (Kalkan & Smith 2014)(Fig 51), its presence in differentiation promoting conditions is an indicator of differentiation delay. Our collaborators used fluorescence-activated cell sorting (FACS) to identify Zfp-GFP positive cell lines after 24 hours of 2i withdrawal. In this way, they identified 230 genes that altered the pluripotency phenotype of cells in N2B27 (Garmhausen 2016). Following the functional analysis of these genes and their pathways, 130 candidates were selected for verification by targeted CRISPR knockout. An initial analysis of RNA sequencing data from 70 knockout conditions and wild type controls was was carried out by Marius Garmhausen (Garmhausen 2016).

Here, I describe preliminary results from an independent analysis of RNAseq data from 74 CRISPR-induced knockout mESC lines that affect differentiation (including the 70 previously analysed cell lines). We first the response to 2i withdrawal in wild type mESCs with high stringency and confidence, confirming many known elements of the early differentiation trajectory. Mutant cell lines were then characterised by both their impact on gene expression in the naive pluripotency state (2i medium), the N2B27 medium, and their divergence from the wild type differentiation trajectory. This enabled us to classify the knockout response as follows: constitutive knockout effects that occurred regardless of medium, and conditional knockout effects, i.e. those that showed a strong dependency on the culture condition. We dissected these responses further and identified groups of knockouts that regulate distinct functional gene clusters of the pluripotency network, suggesting that they affect differentiation by distinct but parallel mechanisms. Additionally, we integrated our results with the analyses of two supporting experiments: firstly, we characterized the LIF-specific response of wild type cells, showing that knockouts that mimic LIF-specific effects in 2i (“LIF-like” knockouts) are more likely to delay differentiation in N2B27. Secondly, we highlight genetic interactions of pluripotency that we identify from a combined knockout-/inhibitor-screen. We also developed mESCexplorer, a web-based application that makes our results available for interactive exploration and visualization.

4.2 Project aims and contributions

The ability of stem cells to differentiate into any cell type and self-renew has made them a primary focus for research in regenerative medicine (Tabar & Studer 2014; Mizuno et al. 2012), after the discovery of induced pluripotent stem cells first demonstrated that pluripotent stem cells could be obtained by reprogramming somatic cells using defined factors (Takahashi &

Yamanaka 2006). The directed differentiation of stem cells to grow intestinal tissue organoids *in-vitro* has been another landmark application of the growing understanding of differentiation factors (McCracken et al. 2011), further highlighting the importance of this field of research for future therapeutic applications. While the naive pluripotency state has been described in great detail (Young 2011), the circuitry that governs the exit from pluripotency and subsequent lineage specification is highly complex (Kalkan & Smith 2014; Leeb et al. 2014; Trott & Martinez Arias 2013) and still not fully understood.

The aim of this project is to advance the understanding of regulatory mechanisms that govern the exit from pluripotency. Due to the high interconnectedness of the pluripotency network, multiple partially redundant mechanisms are expected and known to exist. To disentangle them, I perform computational analysis of expression data obtained from a large mESC knockout screen, as well as the integration of these results with two auxiliary experiments. Through identifying groups of knockouts that impair the exit from pluripotency via distinct molecular pathways, this project contributes to the knowledge needed to more precisely engineer this exit. To facilitate this further, we present mESCexplorer, a web-based software that enables the interactive exploration and visualization of our results. This serves two important goals: firstly, it facilitates collaboration on this project, making it easier for our collaborators to apply their domain-specific knowledge. Secondly, it will increase the transparency of results upon publication and thus make them more accessible and useful for future research.

As this project is still being actively developed and extended, the results presented in this thesis are preliminary.

Contributions

I performed the analysis and integration of RNAseq data from the mESC stem cell knockout screen in 2i and N2B27 conditions, all computational analysis related to the supporting experiments, and developed the mESCexplorer web application. Design of the preceding forward genetic screen was supported by Marius Garmhausen, who also performed initial data processing and analysis (Garmhausen 2016). All wet lab experimental work and sequencing was carried out in the labs of Austin Smith (University of Cambridge), Francis Stewart (Dresden University of Technology), and our collaborator Martin Leeb (University of Vienna).

4.3 Results

4.3.1 Modelling transcriptomes from a large mESC perturbation screen

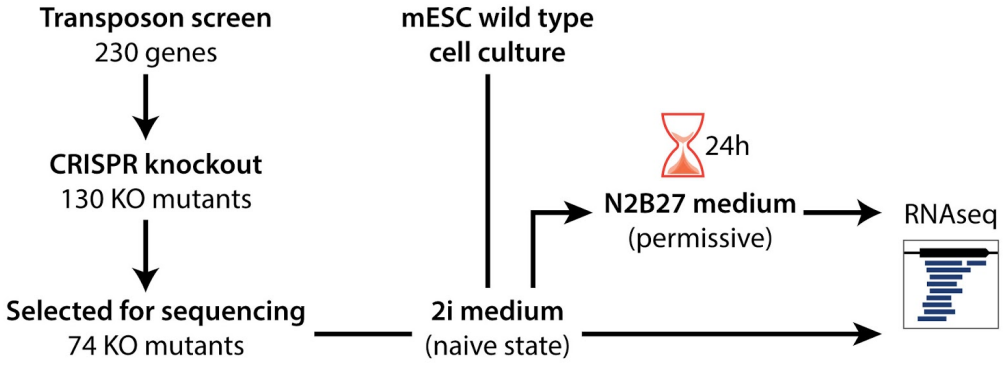


Fig 52***: Experimental setup of the mESC genetic perturbation screen.***
Schema showing the experimental setup of this project. We sequenced the transcriptomes of wild type mESC and 74 knockout mutants with an altered pluripotency phenotype. Samples were taken in two conditions: from 2i medium, which maintains pluripotency, and after 24h hours in the differentiation-permissive N2B27 medium.

To gain insight into which expression patterns might confer to mESC's an improved ability to self-renew in differentiating medium, we performed a forward genetic screen using the piggyBac transposon system (Garmhausen 2016), resulting in the identification of 230 genes that were associated with differentiation delay (Fig 52). Subsequently, 130 single gene knock-out mutants were generated using CRISPR (Fig 52). We obtained RNA sequencing data for 74 of these knockouts, with samples taken in two conditions: cells in 2i medium (at the 0h time point) and cells in N2B27 differentiation medium (after 24h)(Fig 52). This was done for two biological replicates per knock-out condition and two wild type (*RC9*) samples in every sample preparation batch (Fig 52). Alignment of reads to the mouse mm10 reference genome was performed using STAR (Dobin et al. 2013) as described in (Garmhausen 2016). Solely for the purposes of initial filtering, we then calculated FPKM values for every gene in every sample. FPKM stands for fragments per kilobase per million mapped fragments, and as such inherently corrects for sequencing bias arising from variable gene lengths ("per kilobase") and library size ("per million mapped fragments"). In single-end RNAseq experiments, as performed here, this is equivalent to RPKM, first introduced in (Mortazavi et al. 2008)(Eq 13).

$$RPKM_g = \frac{r_g \cdot 10^9}{length_g \cdot R} \qquad (13)$$

For all subsequent analysis, we used only genes that passed a threshold of RPKM ≥ 1 in at least one sample, thereby excluding 1020 genes in our reference annotation as too lowly expressed.

However, RPKM is not an accurate measure of absolute RNA concentration (Wagner et al. 2012). It is also not used in most tools for RNAseq differential expression analysis, as gene length normalization is unnecessary for comparing the same gene between conditions. For downstream differential expression analysis, we therefore relied on a different normalization procedure.

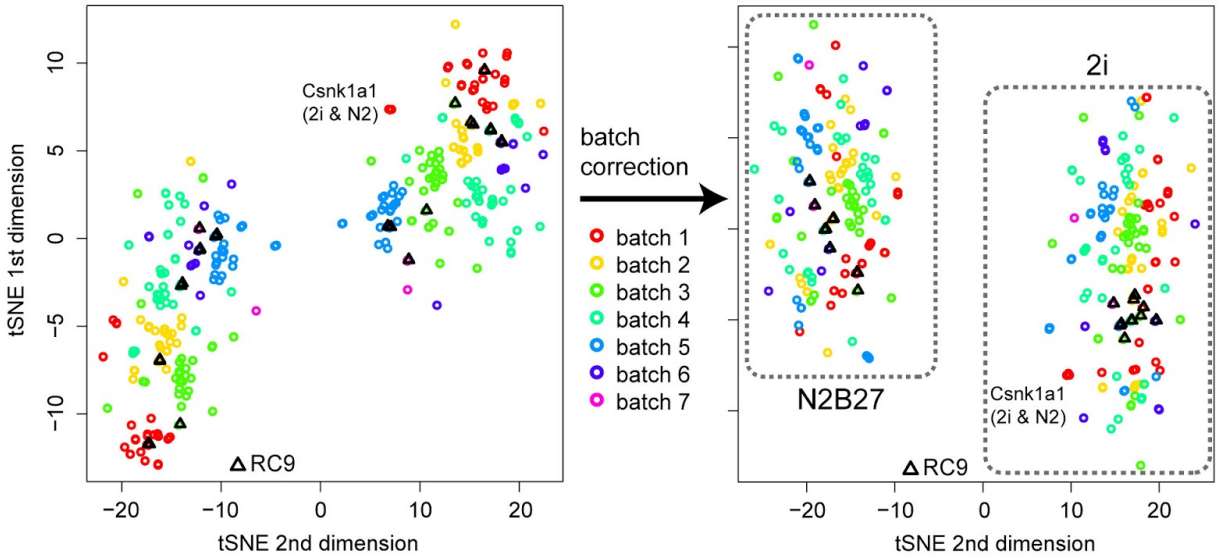


Fig 53***: Batch correction accounts for systemic biases from RNAseq sample preparation.***
Scatterplots showing t-SNE projections of all samples in 2i and N2B27 conditions, before (left) and after (right) batch correction using a linear model. Sequencing batches of samples are represented by colours. Black triangles mark RC9 control replicates. Dotted rectangles indicate samples from the N2B27 and 2i condition, respectively. Csnk1a1 samples are labeled as they are notable outliers.

To prepare our data for downstream analysis, we first corrected sample count values for differences in library size using edgeR with the trimmed mean of M-values method (Robinson & Oshlack 2010). To be able to make use of the flexible limma framework for linear modeling of expression data (Smyth 2005), we had to further transform our count data to correct for heteroskedasticity: the variance of count data generally increases with its mean, violating the assumption in linear modeling that variance is equal across the whole range of a variable (i.e. homoscedasticity). We therefore corrected our data using the voom method (Law et al. 2014). We then fitted a linear model to all expression values, with all unique combinations of genotype (74 *KO*, and the *RC9* control) and time point (N2B27 at 24h, 2i at 0h) encoded in one factor, *GenotypeAndTime* (Eq 14). To avoid biasing our analysis, we checked for batch effects, systemic biases known to occur from non-random sample treatment differences in large scale experiments (Leek et al. 2010). We found that the RNAseq sample preparation batch strongly

predicted clustering of our samples and incorporated it as a confounding variable (*SeqBatch*) in our linear model (Eq 14).

$$Expression = 0 + GenotypeAndTime + SeqBatch \quad (14)$$

For visualization purposes, we carried out batch correction separately to directly transform our matrix of expression values. The transformed matrix was however not used for downstream analysis, as using already corrected values as input for a linear model for differential expression would lead to an overestimation of the degrees of freedom in the model (and thus render the resulting statistics anti-conservative). Visualization was performed using t-SNE for dimensionality reduction before and after batch correction (Fig 53)(Maaten & Hinton 2008). A clear clustering by batch could be observed in the uncorrected data, but was largely resolved after correction (Fig 53). *RC9* controls are dispersed before correction, but cluster more closely together after (Fig 53). Overall, knockout samples clustered most strongly according to the cell culture medium, 2i or N2B27, forming two big clusters (Fig 53). One exception were *Csnk1a1* knockout samples, whose expression patterns were highly similar between N2B27 and 2i conditions. Overall, the cell culture medium was the dominant source of variability, and no mutant could be identified as belonging to an intermediate state based on global expression patterns alone.

4.3.2 Reorganization of the mESC transcriptome during normal differentiation

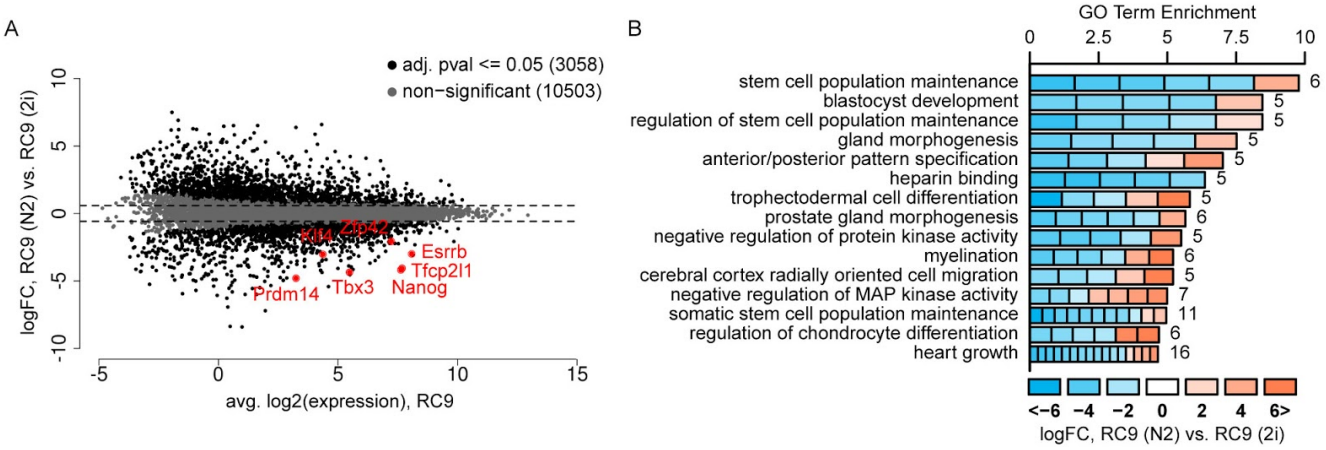


Fig 54***: The wild type differentiation response upon N2B27 placement.***
(A) *MA plot showing all genes by their log-fold-change in the $RC9_{N2}$ vs. $RC9_{2i}$ comparison, over their average expression. Genes changed significantly more than 50% (|logFC| > log2(1.5)) at a significance level of α = 0.05 are marked in black. Non-significant genes are marked in grey. Seven markers of naive pluripotency are highlighted in red.* ***(B)*** *Cell plot showing GO terms enriched in the top 10% most significantly changes genes. Numbers beside bars indicate the number of significant genes annotated with the respective term, while cell colours indicate the*

log-fold-changes of individual genes.

To identify changes that occur during normal differentiation, we first fitted a contrast to perform differential expression analysis, comparing the *RC9* wild type control expression in the N2B27 condition to those in the 2i condition (Eq 15).

$$NormalDifferentiation = RC9_{N2} - RC9_{2i} \qquad (15)$$

To ensure that we identified changes of biologically meaningful magnitude, we tested directly whether the observed change was significantly higher than 50%, using the TREAT method as implemented in limma (McCarthy & Smyth 2009). Thus, any given gene can only be found significant at the chosen confidence level if the entire corresponding confidence interval is above log2(1.5) or below -log2(1.5). This method is therefore more stringent than applying a flat fold-change cutoff, which favours noisy genes (genes with sufficiently high mean changes but high variances may pass the cutoff). P-values from the moderated t-test performed by TREAT were corrected for multiplicity using the BH-method. In total, we found 3058 genes to be differentially expressed between $RC9_{N2}$ and $RC9_{2i}$ at a significance level of $\alpha = 0.05$ (Fig 54A). In accordance with expectations, we could confirm that expression of key markers of naive pluripotency (*Zfp42, Prdm14, Nanog, Klf4, Tfcp2l1, Tbx3, Esrrb*)(Kalkan & Smith 2014) was strongly reduced in the N2B27 condition (Fig 54A). GO enrichment analysis of the top 10% most significantly changed genes confirmed a high enrichment in differentiation-associated processes, such as stem cell population maintenance and blastocyst development (Fig 54B). Thus, our controls behave as expected, reflecting the high quality of our experimental setup.

4.3.3 Knockout-induced changes in 2i and N2B27

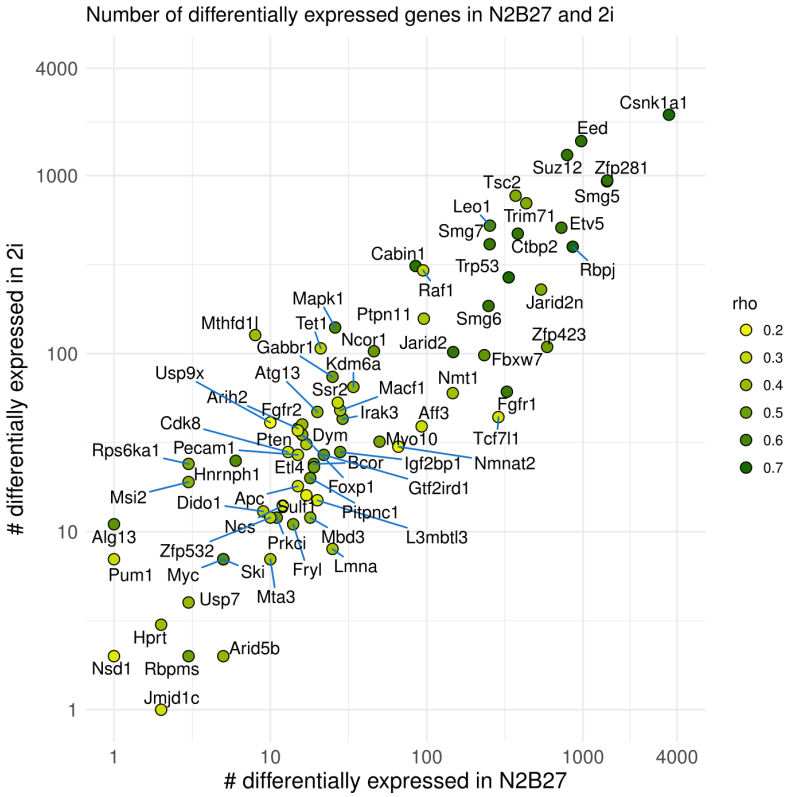


Fig 55: **Knockout-induced changes in 2i and N2B27.**
Scatterplot showing number of diff. expressed genes in each of the 74 knockouts vs. the RC9 wild type, in either 2i or N2B27 (adj. P-Value ≤ 0.05, |logFC| ≥ log2(1.5)). Colour indicates correlation between 2i and N2B27 changes.

We next asked which changes were induced by each knockout in either N2B27 or 2i condition, fitting a contrast for each knockout *k* and condition *t* (Eq 16).

$$KnockoutInduced_{k,t} = Knockout_{k,t} - RC9_{t} \qquad (16)$$

P-values from the TREAT moderated t-test were corrected for multiplicity using the BH-method. We found up to 2192 genes to be differentially regulated upon knock-out in 2i, and up to 3553 in N2B27 (Fig 55), at a significance level of $\alpha = 0.05$ and $|logFC| \geq log2(1.5)$. While the numbers of differentially expressed genes varied, they were highly correlated between samples from both conditions (Fig 55). Spearman rank correlations between corresponding knockout-induced changes in N2B27 and 2i showed moderate to high positive correlation (up to ρ=0.69)(Fig 55). This indicates that responses are partially dependent on the culture medium condition, but nonetheless reproducible.

4.3.4 Pluripotency marker expression in N2B27 correlates with a global delay of differentiation-associated changes

To characterize the pluripotency phenotype in our knockout conditions, we checked for known molecular signatures of pluripotency in the N2B27 condition. Preliminary classification of cells by our collaborators relied on the intensity of a *Zfp42*-GFP reporter construct. *Zfp42* presence is a reliable indicator of (but not required for) pluripotency (Rogers et al. 1991; Masui et al. 2008). To increase precision further we selected six additional pluripotency markers that are known to decline rapidly after onset of differentiation in wild type mESC's: *Nanog*, *Esrrb*, *Tbx3*, *Tfcp2l1*, *Klf4*, and *Prdm14* (Kalkan & Smith 2014).

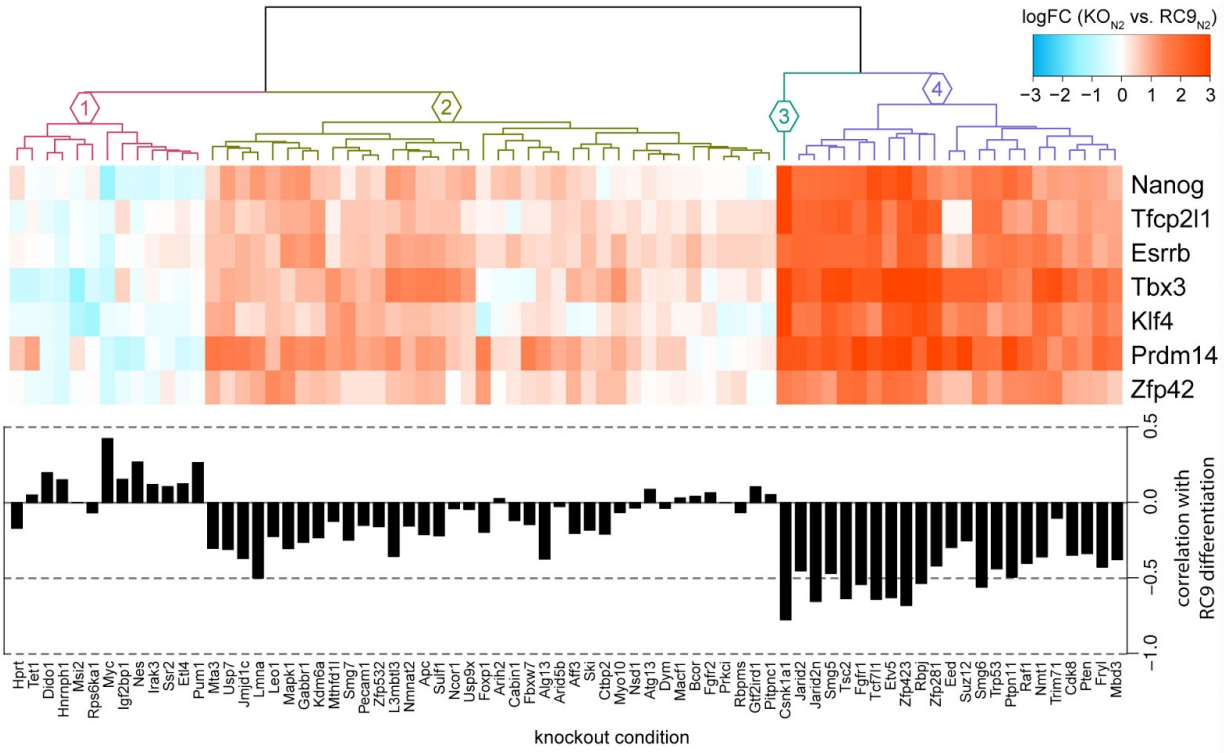


Fig 56***: Pluripotency marker expression reflects a global delay in the differentiation programme.***
Heatmap showing the log-fold-changes of seven core pluripotency markers in 74 knockout backgrounds compared to the wild type, upon withdrawal of 2i (i.e. N2B27 condition). Bar charts show the correlation of log-fold-changes induced by each knockouts with normal differentiation associated changes (RC9$_{N2}$ vs. RC9$_{2i}$).

We then clustered the log-fold-changes from each KO$_{N2}$ vs. *RC9*$_{N2}$ comparison. Besides *Csnk1a1*, which showed by far the highest up-regulation compared to the control and formed its own cluster, we could identify three major clusters (Fig 56). Cluster 4, consisting of 22 knockout conditions, showed a strong up-regulation of all pluripotency markers with minor variations (Fig 56). Notably, *Suz12* and *Eed* knockout resulted did not lead to the *Tfcp2l1* up-regulation that was observed in all other conditions in this cluster (Fig 56). Cluster 2 contained 38 knockout

conditions and was more heterogeneous, showing moderate to low up-regulation of markers (Fig 56). Cluster 1 contained 13 knockout conditions in which naive markers were not regulated or even down-regulated compared to *RC9* controls (Fig 56). This included *Myc* and *Nes* knockouts, which accelerate differentiation (prior experimental results, not shown) and were included as positive controls.

We next asked if increased naive marker expression reflects the inhibition of differentiation-associated remodeling of the transcriptome induced by the knockouts. We calculated the correlation between significant changes during normal differentiation in the wild type (*RC9*$_{N2}$ vs. *RC9*$_{2i}$), and changes observed in each KO$_{N2}$ vs. *RC9*$_{N2}$ comparison. If the transcriptional programme is delayed or abrogated by the mutants, this correlation should be negative, while a positive correlation might indicate an acceleration of differentiation. Indeed, changes in the positive differentiation controls *Myc* and *Nes* were concordant with normal differentiation changes, i.e. positively correlated (Fig 56). On the other hand, we observed negative correlations in knockout conditions with increased naive marker expression in N2B27 (Fig 56). This suggests that differentiation delay is reflected by the abrogation of transcriptional reprogramming on a global level.

4.3.5 Dependence of differentiation on naive marker control

We next wanted to identify genes that are closely associated with the core pluripotency machinery, i.e. the seven selected naive pluripotency markers (*Nanog*, *Esrrb*, *Tbx3*, *Tfcp2l1*, *Klf4*, and *Prdm14)*. This problem could be approached naively by averaging the normalized expression values of these markers and then calculating pairwise correlations with all remaining genes. However, while the naive marker expression patterns were highly similar overall, individual markers diverged markedly in some cases (e.g. Tfcp2l1 expression in *Suz12* and *Eed* knockouts)(Fig 56). A correlation analysis with averaged values would discard the information from these differences and thus underestimate the variance that a given gene shares with the whole set of genes. An alternative is to perform multiple regression analysis, modeling each gene's expression vector *G* as the linear combination of all *n* marker gene expressions (Eq 17).

$$G = \beta_0 + \sum_{1 \leq i \leq n} \beta_i M_i \qquad (17)$$

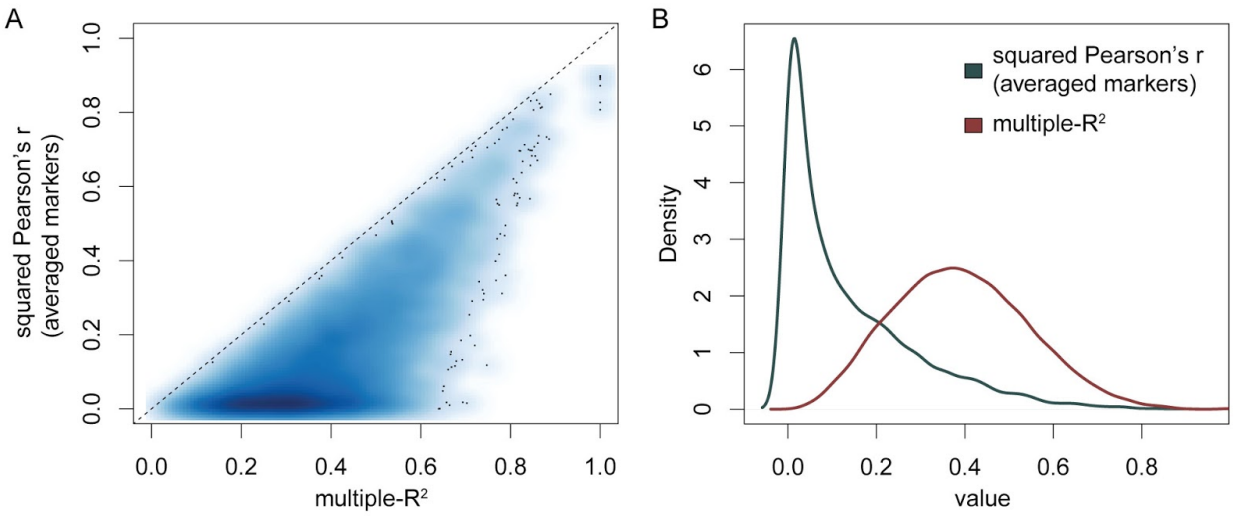


Fig 57***: Marker multiple regression compared to correlation with averaged marker expression.***
(A) Density scatter plot of squared Pearson's correlation values (genes ~ averaged marker expression) and multiple-R^2 values (genes ~ marker expression). (B) Line plot showing the corresponding density distributions.

We performed multiple regression analysis first. A high shared variance (R^2 value) indicates a high dependence on these central regulators of pluripotency, while a low R^2 value indicates relative independence from marker expression. We then compared this measure against the squared Pearson correlations of genes and averaged marker expressions. As expected, R^2 values from our multiple regression analysis were strictly higher that squared Pearson correlations (Fig 57A). While the two measures converged for higher ranked genes, the distribution of squared Pearson correlation was heavily skewed (Fig 57B), suggesting that the lower ranks were less informative of true marker-independence. We therefore decided to proceed with the R^2 estimates from our multiple regression analysis. We then used this measure to characterize the wild type differentiation response further.

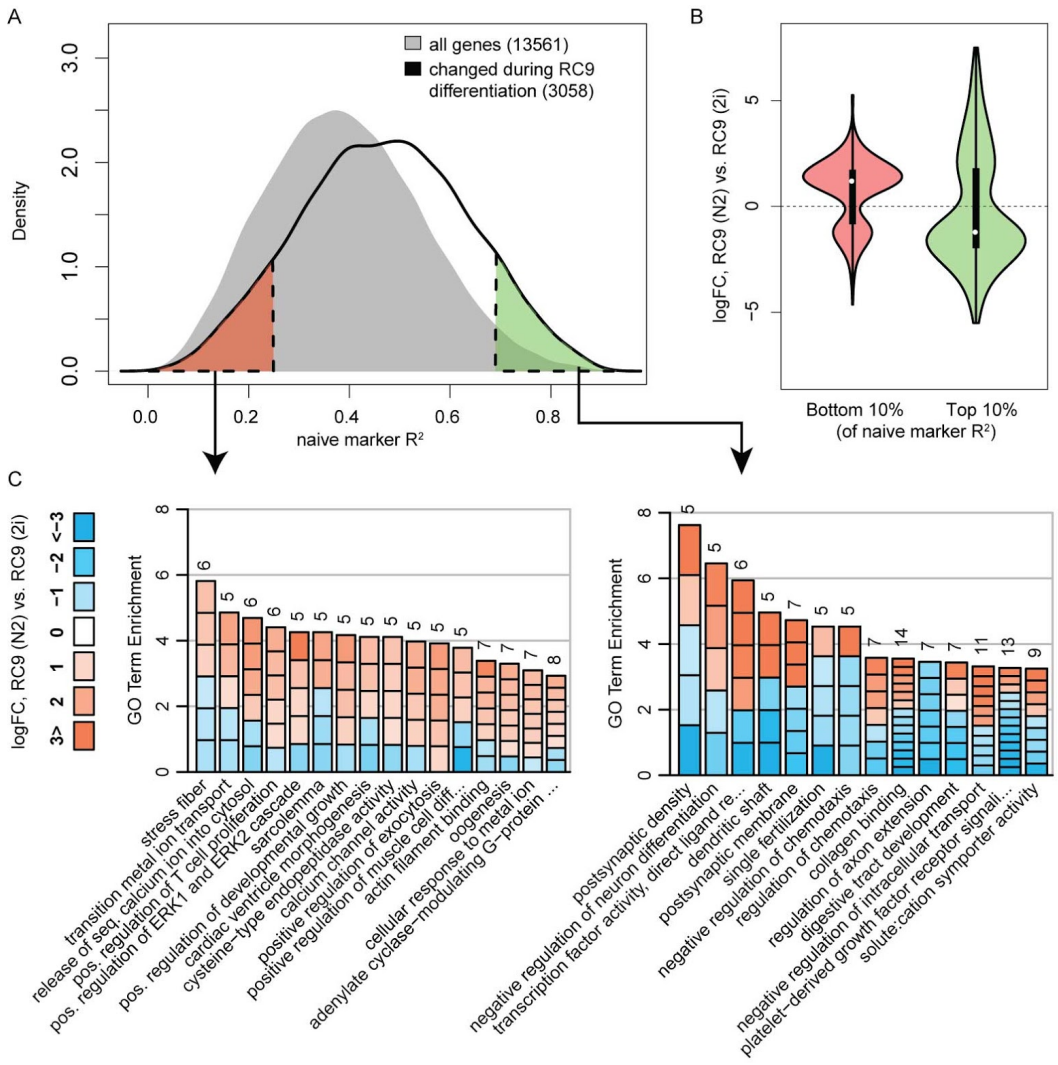


Fig 58**: Pluripotency marker dependent and independent elements of wild type differentiation.**
(A) Line chart showing the distribution of naive marker multiple-R^2 values, for all quantified genes (grey) or genes changed during wild type differentiation (black). Indicated in green and red are the top (marker-dependent) and bottom (marker-independent) 10% quantiles of R^2 values of differentiation associated genes, respectively. **(B)** Violin plots showing log-fold-change distributions ($RC9_{N2}$ vs. $RC9_{2i}$) of marker-independent (red) and -dependent (green) differentiation-associated genes. **(C)** Cell plots showing GO term enrichment and log-fold-changes of marker-independent (red) and -dependent (green) differentiation-associated genes.

From all differentiation-associated genes we selected the genes with the bottom and top 10% quantiles of marker R^2 values (Fig 58A). Comparing their log-fold-change distributions, we observed that marker-independent genes were much more likely to be up-regulated in N2B27 compared to 2i, whereas marker-dependent genes, just like the markers themselves, were overwhelmingly down-regulated (Fig 58B). We then performed GO enrichment analysis on each set of genes (Fig 58C). Marker-independently regulated differentiation-associated genes included many positive regulators of developmental processes; in particular ERK signalling,

which is inhibited in the naive state but crucial in differentiation (Fig 58C). Marker-dependent genes, on the other hand, were more likely associated with negative regulation of developmental terms and intracellular transport (Fig 58C), explaining their repression during differentiation (Fig 58B). These results confirms that reprogramming of the pluripotency network to trigger differentiation happens through both the dissolution of negative developmental cues, dominated by canonical pluripotency markers, and the institution of positive regulation of development. Importantly, we were able to distinguish between naive marker dependently and independently regulated genes through our regression analysis, adding valuable context to downstream analyses.

4.3.6 Entry points of regulation through the core pluripotency network

We next asked if we could determine the relative influence of individual pluripotency markers on each gene's expression in the N2B27 condition: is there one naive marker that best predicts the expression of the respective gene, or is this control shared? Are there sets of genes that change during differentiation in a manner specifically dependent on one of the markers?

We first interrogated our multiple regression analysis. Given sufficient independence of the predictor variables, coefficients determined by multiple regression should allow the ranking of predictors. However, if predictors are highly dependent, i.e. multicollinear, coefficients become unreliable (Mansfield & Helms 1982). We observed that N2B27 log-fold-changes of naive differentiation markers were highly correlated with each other (Fig 56). An additional commonly used measure to quantify multicollinearity in multiple regression analysis is the variance inflation factor (VIF). If a variable is linearly independent of all other predictors, the VIF equals one. In contrast, a high VIF indicates that the predictor is dependent on other predictors, resulting in unstable coefficients with inflated errors (Mansfield & Helms 1982). We calculated VIF's between 5.58 to 11.94 across marker expression profiles, confirming their high interdependency.

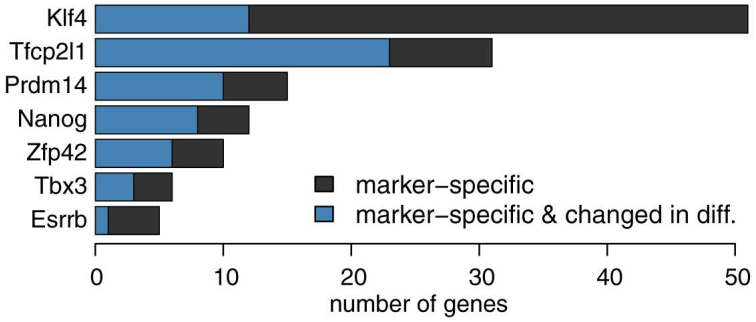


Fig 59*: **Marker-exclusively regulated genes.***
Barplot showing the number of genes that were marker-specific, i.e. highly correlated ($|r| \geq 0.7$) and highly exclusive to the respective marker (marker top ranked in over 90% of random sub-samples for that gene). Blue

portions of bars indicate the number of genes that was both marker-specific and strongly regulated in the normal course of differentiation (N2B27 vs. 2i wild types).

To guard against noisy coefficient estimates and obtain reliable rankings, we therefore employed an iterative sub-sampling method, stability selection (Meinshausen & Bühlmann 2010; Shah & Samworth 2013). For each gene we generated 1000 randomly chosen KO sample subsets, each consisting of half the available KO samples. We then selected, for each subset, the predictor that had the highest shared variance (R^2) with the respective gene. A high selection frequency indicates that out of the tested pluripotency markers, the selected marker consistently shows the strongest association to the gene. Thus, to find genes that are strongly and specifically associated with each marker, we selected genes whose expression was (a) highly correlated to that marker's expression across KO samples ($|r| \geq 0.7$), and (b) for whom the marker was ranked top in over 90% of stability selection subsamples. As an additional filter, we determined which of the genes thus selected were also changed during the normal course of differentiation. Between 5 and 51 genes were identified as specific for each individual marker (Fig 59).

While we failed to detect an overall enrichment of known functional categories (GO and reactome) in these highly specific gene sets, individual marker-gene associations have been reported before: The highest scoring *Klf4*-associated gene, *Cct2*, is involved in cell cycle control and has been previously identified as a putative *Klf4* target (Won et al. 1998; Whitney et al. 2006). The highest scoring *Nanog*-associated gene, *Alpl*, is a commonly used stemness marker that has been implicated in cancer (Liu et al. 2013). Interestingly, another high scoring *Nanog*-associated gene, *Slc2a3* interacts directly with the *Nanog* promoter in pluripotent stem cells (Apostolou et al. 2013). These results suggest that our analysis recovers biologically relevant associations of naive markers and other genes. We will further investigate promising candidates in the context of this project.

4.3.7 Disentangling the constitutive and conditional knockout responses

To discern which parts of the response to each knockout in either condition are dependent on the culture medium, we calculated the significance of the two-way interaction contrast between the medium and the knockout condition (*Knockout:RC9*), separately for each knockout *k* (Eq 18).

$$Knockout_k{:}RC9 = (KO_{k,N2} - RC9_{N2}) - (KO_{k,2i} - RC9_{2i}) \qquad (18)$$

We tested against a stringent threshold, requiring changes to be significantly higher than 1.5-fold. We could identify 2274 genes that were significantly changed in at least one *Knockout:RC9* interaction contrast (after BH-coeffection, $\alpha = 0.05$).

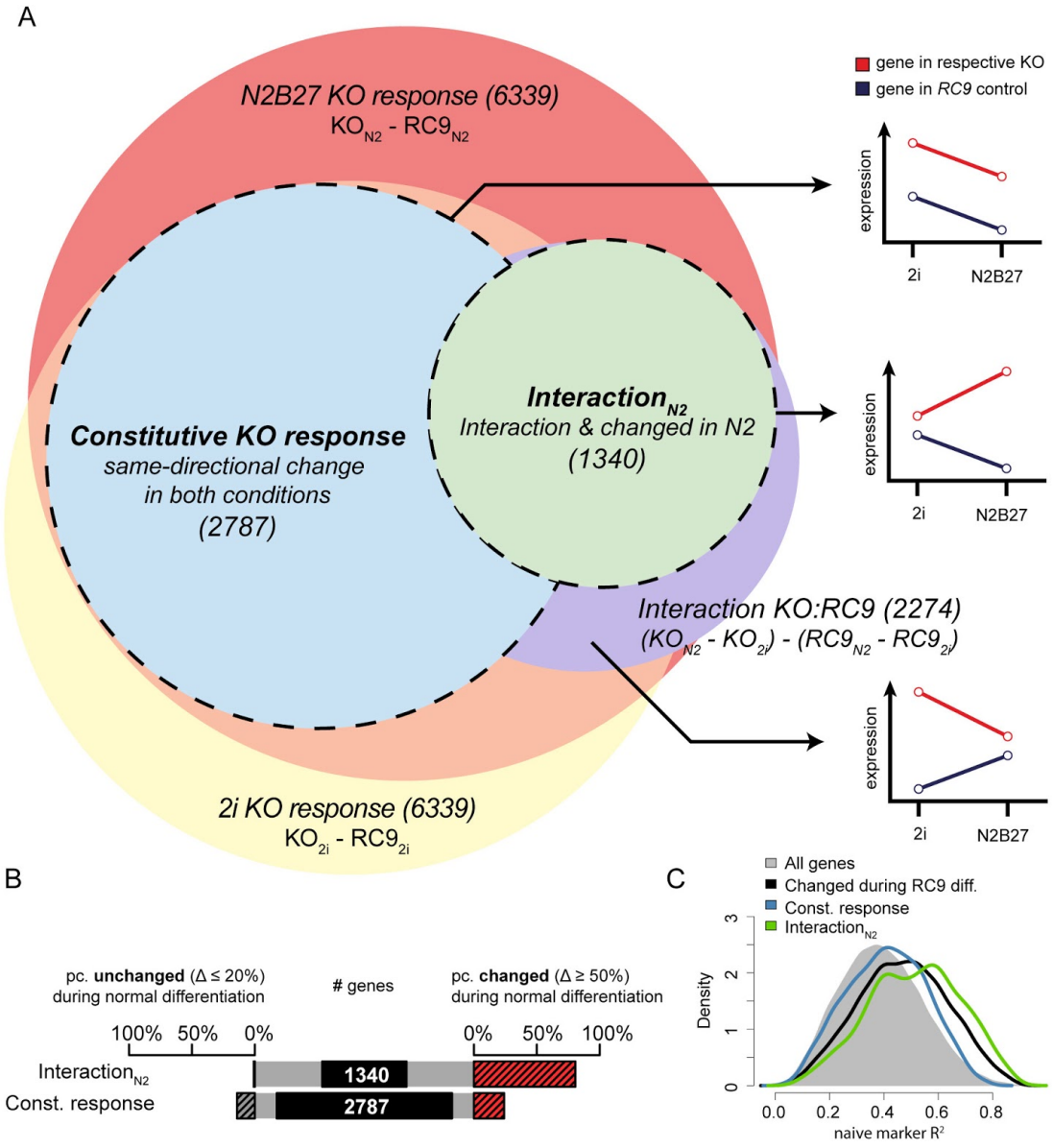


Fig 60***: Disentangling constitutive and conditional knockout responses.***

(A) *Euler diagram showing the three different contrasts (knockout effect in 2i or N2, interaction), their overlaps, and two specific subsets: the constitutive KO response and Interaction$_{N2}$. Areas are proportional to the number of genes in each set (numbers are shown in parentheses). Line charts show schematic representations of the selected expression patterns in the indicated subset. (B) Bar charts showing the total number of genes in the constitutive KO response and Interaction$_{N2}$ subsets (center column), as well as the percentages of these genes that are either significantly changed (logFC ≥ log2(1.5), adj. P-value ≤ 0.05) or unchanged (|logFC| ≤ log2(1.2), TOST adj. P-value ≤ 0.05) during normal differentiation (RC9$_{N2}$ vs. RC9$_{2i}$).* ***(C)*** *Line chart showing the distribution of naive marker multiple-R^2 value in N2B27 (chapter 3.3.5), for all genes (grey), genes changed during normal differentiation (black), the constitutive KO response (blue), and the Interaction$_{N2}$ subset (green).*

Based on the responses to the knockout conditions in both N2B27 and 2i and their interactions with the *RC9* control, we defined two main categories of responsive genes (Fig 60A):

(1) Constitutive knockout response: genes that were significantly changed in the same direction in both N2B27 and 2i, in at least one knockout condition. Genes from the $Interaction_{N2}$ category are excluded.

(2) Conditional N2B27 knockout response (= $Interaction_{N2}$): genes that were significantly changed in both N2B27 and the *knockout:RC9* interaction term, in at least one knockout condition. All seven selected key pluripotency markers (*Zfp42, Prdm14, Nanog, Klf4, Tfcp2l1, Tbx3, Esrrb*), were in this category.

Additionally, we determined how many genes in either subset were significantly changed or unchanged during normal differentiation ($RC9_{N2}$ *vs.* $RC9_{2i}$)(Fig 60B). Here, we defined unchanged genes as genes that changed by significantly less than 20% ($|t| = log2(1.2)$), as determined by a TOST-test (Westlake 1976) at significance level $\alpha = 0.05$ (after BH adjustment for multiplicity). To determine whether genes in the $Interaction_{N2}$ subset were overall more correlated with naive pluripotency markers, we checked the distribution of naive marker multiple-R^2 values (chapter 4.3.5) in the $Interaction_{N2}$ set compared to both the background distribution and genes classified as taking part in the constitutive KO response (Fig 60C). We found that this distribution was shifted towards higher R^2 values, suggesting that $Interaction_{N2}$ genes were more closely linked to the differentiation delay phenotype (Fig 60C). However, the bimodal distribution of R^2 values in this subset indicates that part of the $Interaction_{N2}$ subset is less dependent on global pluripotency marker expression and may therefore be specific to a sub-group of knockout conditions.

4.3.8 Functional clusters of the constitutive knockout response

The constitutive knockout response encompasses genes that respond to at least one knockout intervention in a consistent manner regardless of the medium. Differentiation causes a wide-reaching rewiring of transcription (Fig 54). Many of the observed changes in the N2B27 condition may therefore be secondary consequences of the altered state of pluripotency in our cell lines. By selecting responses that are already present in the 2i condition, we are less likely to select these secondary effects. However, a similar response in both 2i and N2B27 conditions might also imply a weaker overall association to changes in the pluripotency phenotype. Indeed, we found that only 24% of changes in this category were detected as significantly changed during the normal course of differentiation (Fig 60B). Conversely, 14% of genes were changed by significantly less than 20% during normal differentiation, as determined using a TOST test of equivalence (adj. P-Value ≤ 0.05)(Fig 60B).

To determine if distinct functional clusters of genes that may be activated in all, single, or sub-groups of knockouts, we clustered all genes based on their KO_{N2} vs. $RC9_{N2}$ log-fold-changes. For this, we considered several available clustering algorithms: k-means, HDBSCAN, full linkage hierarchical, single linkage hierarchical, and Ward's method. We determined empirically that Ward's method recovered the most distinctive clusters in our data. Similar to k-means clustering, it is based on the minimization of the within-cluster sums of squared errors, i.e. an analysis of variance approach (Ward 1963). Unlike k-means, Ward's method is an agglomerative clustering technique: it begins by treating all points (genes) as singlets. Clusters are then formed by iteratively linking previous clusters or singlets in a way that minimizes the total within-cluster variance *W*. Ward's method therefore recovers a hierarchical structure that can be conveniently visualized in a dendrogram. Distances between clusters in the hierarchical representation are proportional to the increase of within-cluster variance that would result by the fusion of these clusters. After thus learning the fusion hierarchy, the resulting tree needs to be cut to generate a number of clusters *k* that balances useful generalization against cluster coherence. As an objective criterion to aid this decision, we determined the total within-cluster variance *W* at different values of *k*, ranging from 3 to 70 clusters. We then plotted the within-cluster variances as well as their first-order differenced values against *k* (the latter was done for visual clarity). If there is a natural number of well defined clusters in the data that is recoverable by the clustering algorithm, an inflection point will appear in the resulting curve; this method of determining the optimal number of clusters is widely known as the "elbow method". In terms of our model, we sought to select *k* such that the model improved strongly by increasing the number of clusters up to that value, but only marginal gains result from increasing *k* further. Based on both visual inspection of the cluster dendrogram and the differenced within-cluster variances, we decided to proceed with a division of the constitutive KO response subset into 16 clusters (Fig 64). The sizes of these clusters ranged from 31 to 846 genes (Table 5).

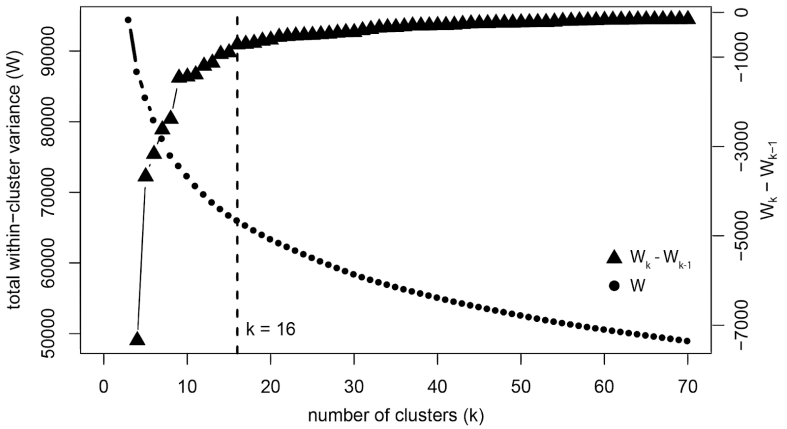


Fig 61***: Determining the optimal number of clusters of the constitutive knockout response.***
Line chart showing total within-cluster-variances (W, points) for different numbers of clusters k. The differenced values of W, i.e. the change of W from k-1 to k, are represented by triangles.

To summarize cluster expression, we calculated the mean of the KO_{N2} vs. $RC9_{N2}$ log-fold-changes in each cluster. We also calculated the cluster-wise z-values of each KO's mean expression in that cluster. KO's whose z-value in a cluster exceeded 1.96 (equivalent to 95% confidence) were deemed specific for that cluster. We then performed additional clustering of knockouts based on cluster mean expression values and plotted them alongside the changes observed in naive pluripotency markers in N2B27 knockout conditions (Fig 62). Constitutive cluster changes did not globally correlate with marker expression changes, as shown by the fragmentation of naive marker changes (Fig 62), indicating relative independence from the core pluripotency network.

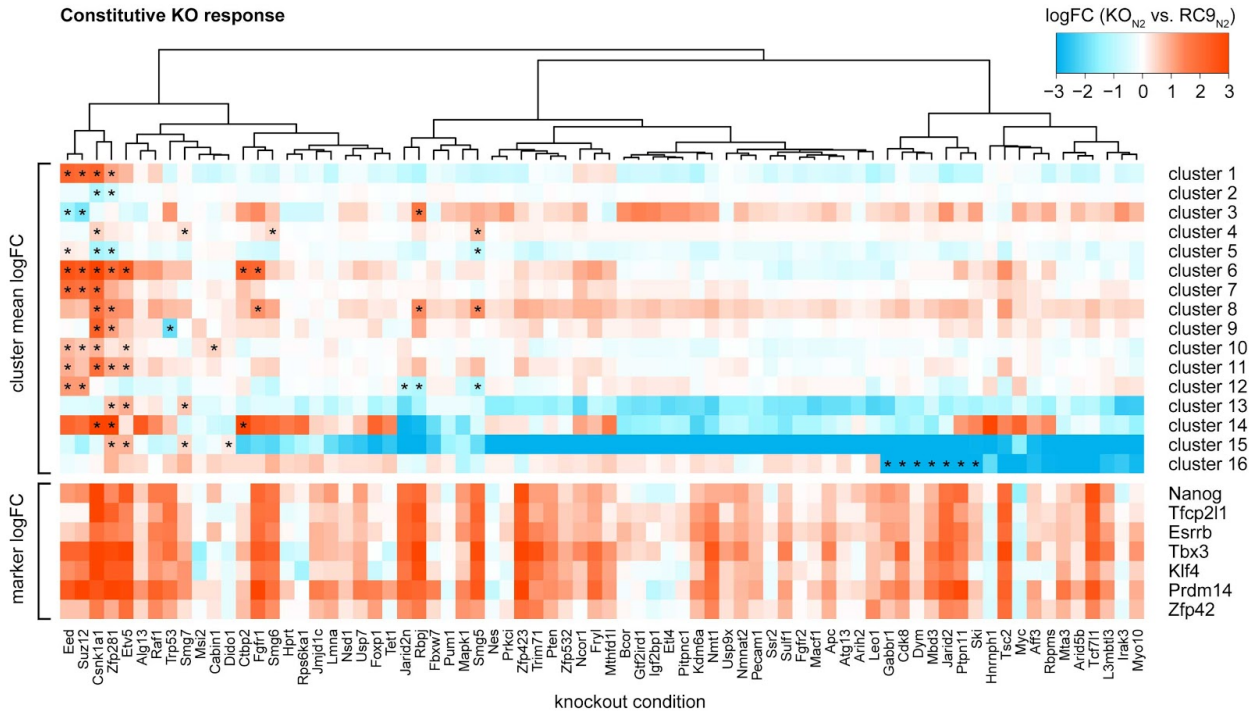


Fig 62: ***Clusters of the constitutive knockout response in N2B27.***
Heatmaps showing average log-fold-changes (KO_{N2} vs. $RC9_{N2}$) of constitutive response gene clusters. Asterisks indicate that the absolute row-wise z-value of the respective mean expression value was above 1.96 ($p \leq 0.05$). For reference, naive marker log-fold-changes from the same comparison are shown in a separate heatmap (column ordering derived only from constitutive response gene clusters).

To identify processes that are regulated in each cluster, and specifically by the cluster-specific KO interventions, we performed GO enrichment analysis on each cluster using SETHRO. In 14 of the 16 clusters, we could identify significant enrichment of at least one GO term that contained more than 5 genes of interest (Table 5). Interestingly, the first cluster, which was specifically up-regulated in the knockouts of *Csnk1a1, Eed, Suz12,* and *Zfp281*, contained multiple *Hox* genes (*Hoxa3,4,10, Hoxb2,3,7*) and associated genes, which are responsible for lineage determination and pattern specification (Seifert et al. 2015)(Table 5). In particular, *Hoxa* and *Hoxb* family transcription factors can drive specification towards a haematopoietic lineage (Ernst

et al. 2004; Mulgrew et al. 2014; Sugimura et al. 2017). Several developmental functional terms specific for different tissues were enriched in other gene clusters (such as ovarian follicle development, glial cell differentiation, skeletal system development, and spermatid development)(Table 5). This suggests that knockouts may tip the balance towards specific lineages in addition to their overall differentiation delay phenotype. However, more analysis is necessary to precisely determine the role of this as well as other identified gene clusters.

Table 5: Knockout specifically regulated clusters of the constitutive knockout response.

Cluster (genes)	**KO specific regulation**	**Enriched functions**
1 (70)	**Up:** Csnk1a1, Eed, Suz12, Zfp281	anterior/posterior pattern specification, embryonic skeletal system morphogenesis, endoderm formation, cell fate commitment, dorsal/ventral pattern formation, renal system development, proteinaceous extracellular matrix
2 (372)	**Down:** Csnk1a1, Zfp281	glycolytic process, ATP metabolic process, magnesium ion binding, myelin sheath, positive regulation of release of cytochrome c from mitochondria, organic acid biosynthetic process, isomerase activity, organellar large ribosomal subunit, DNA-dependent DNA replication, serine family amino acid metabolic process
3 (31)	**Up:** Rbpj **Down:** Eed, Suz12	response to endoplasmic reticulum stress
4 (846)	**Up:** Csnk1a1, Smg5, Smg6, Smg7	antiporter activity, lipoprotein metabolic process, synaptic vesicle priming, anion:anion antiporter activity
5 (197)	**Up:** Eed **Down:** Csnk1a1, Smg5, Zfp281	ovarian follicle development, response to toxic substance, sodium ion transmembrane transport, voltage-gated cation channel activity, BMP signaling pathway, glial cell differentiation, RNA polymerase II transcription cofactor activity, negative regulation of protein modification by small protein conjugation or removal, regulation of muscle system process, monosaccharide transport
6 (55)	**Up:** Csnk1a1, Ctbp2, Eed, Etv5, Fgfr1, Suz12, Zfp281	embryonic skeletal system development, sensory organ morphogenesis
7 (228)	**Up:** Csnk1a1, Eed, Suz12	transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding, regulation of ossification, labyrinthine layer morphogenesis, cell fate commitment, membrane assembly, intermediate filament, cardiac atrium morphogenesis, heart morphogenesis, embryonic skeletal system morphogenesis, synapse assembly
8 (130)	**Up:** Csnk1a1, Fgfr1, Rbpj, Smg5, Zfp281	positive regulation of blood circulation, cellular response to acid chemical, transforming growth factor beta receptor signaling pathway, integrin binding, myofibril, proteinaceous extracellular matrix, plasma membrane receptor complex, protein tyrosine kinase activity, positive regulation of cytosolic calcium ion concentration, cell junction assembly
9 (139)	**Up:** Csnk1a1, Zfp281 **Down:** Trp53	cell cycle arrest, negative regulation of cell growth, intrinsic apoptotic signaling pathway by p53 class mediator, negative regulation of cell migration, zymogen activation, detection of stimulus, DNA damage checkpoint, regulation of heart contraction, positive regulation of apoptotic signaling pathway, blood coagulation

10 (328)	**Up:** Cabin1, Csnk1a1, Eed, Etv5, Suz12	organic anion transmembrane transporter activity, receptor complex, filopodium, GTPase activator activity, L-amino acid transmembrane transporter activity
11 (162)	**Up:** Csnk1a1, Eed, Etv5, Zfp281	proteinaceous extracellular matrix, cellular response to interleukin-1, collagen trimer, basement membrane, metallopeptidase activity, sarcoplasm, carbohydrate derivative catabolic process, cellular response to tumor necrosis factor, endopeptidase inhibitor activity, sarcolemma
12 (109)	**Up:** Eed, Suz12 **Down:** Jarid2n, Rbpj, Smg5	negative regulation of cell morphogenesis involved in differentiation, heparin binding, regulation of synapse structure or activity, glycoprotein binding, apical plasma membrane, postsynaptic membrane, cell-cell junction organization, extracellular matrix component, cell maturation, axon guidance
13 (42)	**Up:** Etv5, Smg7, Zfp281	-
14 (60)	**Up:** Csnk1a1, Ctbp2, Zfp281	thiol-dependent ubiquitin-specific protease activity
15 (13)	**Up:** Dido1, Etv5, Smg7, Zfp281	spermatid development
16 (5)	**Down:** Cdk8, Dym, Gabbr1, Jarid2, Mbd3, Ptpn11, Ski	-

Additionally, we asked if any functions were specifically enriched in the part of the constitutive knockout response that was left unchanged during normal differentiation (Fig 60B) – 394 genes in total. These functions might point to secondary, highly specific knockout effects. We again clustered KO_{N2} vs. $RC9_{2i}$ log-fold-changes of these genes, identifying three distinct gene clusters, and calculated the mean of these values for each cluster and knockout-condition (Fig 63). Interestingly, only six knockout-conditions were responsible for this subset of the response: we saw a down-regulation of cluster 1 in *Zfp281* and *Csnk1a1* knockouts (Fig 63). Cluster 2 was up-regulated in these two knockouts, but down-regulated in the *Trp53* knockout (Fig 63). Cluster 3 was exclusively (up-) regulated in *Smg5, Smg6,* and *Smg7* knockouts (Fig 63). We inferred functional enrichment of GO terms within these clusters using SETHRO (Table 6). The first two clusters were enriched for ribosomal genes as well as components of the smoothened signalling pathway, which is activated by the proliferation-regulating Hedgehog pathway (Rimkus et al. 2016)(Table 6). Cluster 3 was highly specific for genes involved in DNA damage control (Table 6). While primarily involved in nonsense mediated decay of aberrant mRNA transcripts, *Smgs* are known to be involved in genome quality control and DNA repair (Isken & Maquat 2008).

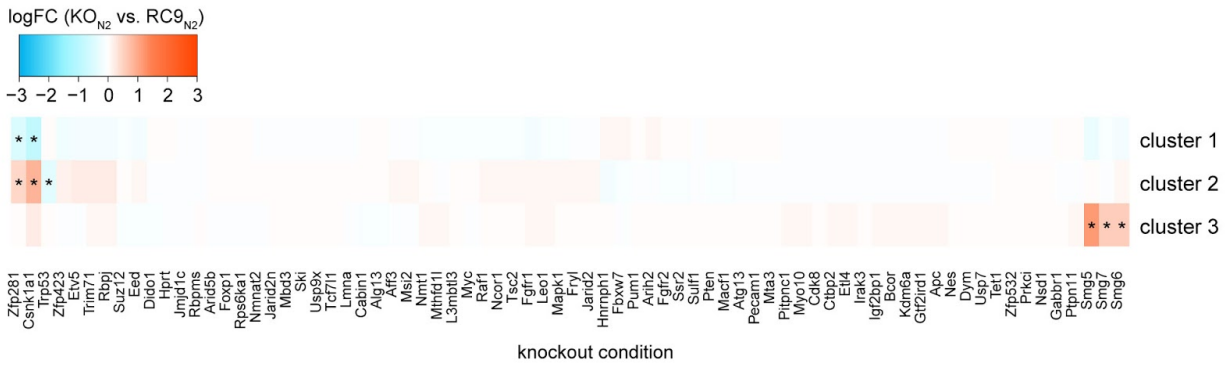


Fig 63: ***Clusters of the constitutive knockout response unchanged in normal differentiation.***
Heatmaps showing average log-fold-changes (KO_{N2} vs. $RC9_{N2}$) of the subset of the constitutive response that did not change during normal differentiation ($|logFC| \leq log2(1.2)$, $RC9_{N2}$ vs. $RC9_{2i}$). Asterisks indicate that the absolute row-wise z-value of the respective mean expression value was above 1.96 ($p \leq 0.05$). For reference, naive marker log-fold-changes from the same comparison are shown in a separate heatmap (column ordering derived only from constitutive response gene clusters).

Table 6: Knockout specifically regulated clusters of the constitutive knockout response.

Cluster (genes)	**KO specific regulation**	**Enriched functions**
1 (129)	**Down:** Csnk1a1, Zfp281	organellar ribosome, nucleoside phosphate catabolic process, protein folding, magnesium ion binding, purine nucleoside monophosphate metabolic process, isomerase activity, myelin sheath, glucose metabolic process, purine ribonucleotide biosynthetic process, lyase activity
2 (106)	**Up:** Csnk1a1, Zfp281 **Down:** Trp53	smoothened signaling pathway, response to UV, nuclear membrane, peptidyl-serine phosphorylation, positive regulation of protein serine/threonine kinase activity
3 (159)	**Up:** Smg5, Smg6, Smg7	G2 DNA damage checkpoint, nucleotide-excision repair, P-methyltransferase activity, protein lipidation, RNA modification, mRNA catabolic process, DNA conformation change, nuclease activity, glycosylation

4.3.9 Functional clusters of the conditional knockout response

In contrast, a gene shows a conditional knockout response if it responds to the medium change differently from the *RC9* control (i.e. significant interaction). In total, we found that 2274 genes responded in this manner in at least one KO condition (Fig 60A). We distinguished two types of interaction: firstly, genes that were also significantly changed only in the respective KO_{2i} vs. $RC9_{2i}$

comparison, but not changed in the N2B27 background. These genes were unlikely to be related to differences between KO and control in N2B27, and were therefore excluded from further consideration. Secondly, genes that were also significantly changed in at least one KO_{N2} vs. $RC9_{N2}$ comparison. We detected 1340 such genes, which we called the $Interaction_{N2}$ subset (Fig 60A).

We then checked how many of the $Interaction_{N2}$ genes were also significantly regulated during the normal course of differentiation (Fig 60B). We found a highly significant overlap of 80% of all $Interaction_{N2}$ genes (Fig 60B). To stringently evaluate whether any genes were unchanged during wild type differentiation, we performed an additional TOST equivalence test. We could only identify 12 genes of the $Interaction_{N2}$ subset that were changed by less than 20% during normal differentiation (abs. log-fold-change ≤ log2(1.2), adj. P-Value ≤ 0.05). This indicates that differentiation delay in knockouts is most likely mediated by the perturbation of the normal re-wiring of differentiation networks.

We then investigated whether it was possible to further subdivide the $Interaction_{N2}$ gene subset into distinct functional clusters of genes. As with the subset of constitutively changed genes, we used Ward's method to learn the gene cluster hierarchy based on minimization of within-cluster variance. Then, from visual inspection of the cluster dendrogram and the differenced within-cluster variances, we partitioned the $Interaction_{N2}$ subset into 12 clusters.

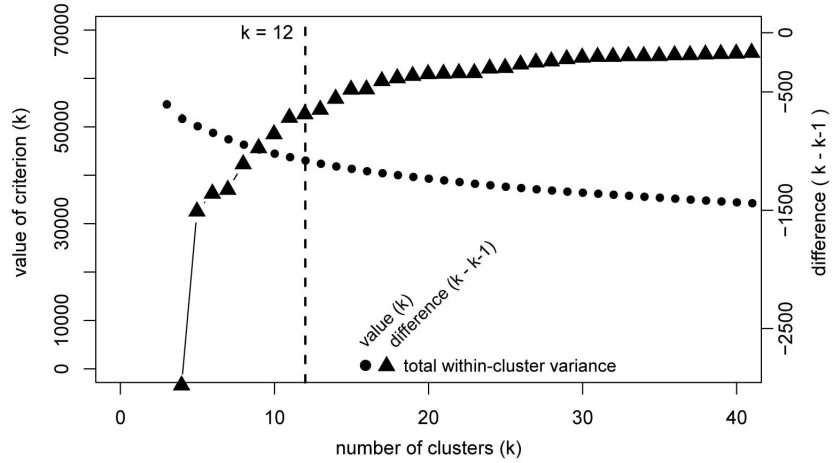


Fig 64***: Determining the optimal number of clusters of the $Interaction_{N2}$ knockout response.***
Line chart showing total within-cluster-variances (W, points) for different numbers of clusters k. The differenced values of W, i.e. the change of W from k-1 to k, are represented by triangles.

We expected three general categories of clusters: Firstly, clusters whose expression correlated or anti-correlated with the differentiation phenotype (as characterized by marker expression or anti-differentiation correlation). Genes classified this way may mirror results from our earlier regression analysis, which also screened for genes linearly dependent on pluripotency marker expression. Secondly, clusters that were entirely uncorrelated with the differentiation phenotype, i.e. likely to encompass coincidental effects of knockouts that are unrelated to the

differentiation phenotype. Lastly, clusters that partially correlate with the differentiation phenotype. This category therefore had the biggest potential to complement our previous analysis: clusters partially linked to the differentiation delay phenotype may indicate distinct mechanisms that affect differentiation in parallel and may thus be used to further characterize subgroups of knockouts.

To summarize cluster expression, we calculated the mean of the interaction log-fold-changes in each cluster; i.e. the differentiation response of each knockout normalized to the wild type response (Fig 65). We also calculated the cluster-wise z-values of each knockout's mean expression in that cluster (Fig 65). This allowed as to classify knockouts that had an unusually high impact on the respective cluster's expression, i.e. whose mean expression absolute z-value exceeded 1.96 (equivalent to 95% confidence)(Fig 65). The strongest knockout-specificity was seen in cluster 10, which was highly up-regulated specifically in $Nmnat_{KO}$ and $Aff3_{KO}$ (Fig 65). In others, the distribution was smoother, indicating a gradual change in the degree of regulation. This was the case in cluster 3, which is up-regulated to a moderate to high degree in most differentiation-delayed knockouts (Fig 65).

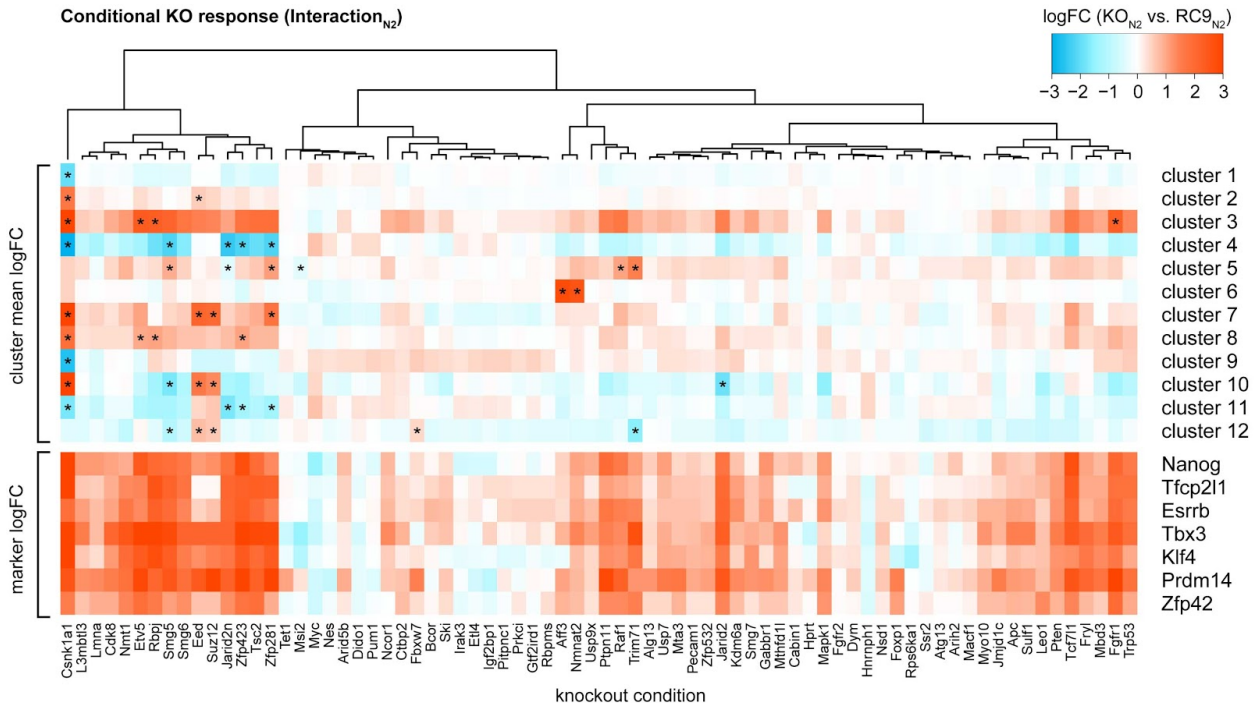


Fig 65: ***Clusters of the constitutive knockout response in N2B27.***

Heatmaps showing average log-fold-changes (KO_{N2} vs. $RC9_{N2}$) of constitutive response gene clusters. Asterisks indicate that the absolute row-wise z-value of the respective mean expression value was above 1.96 ($p \leq 0.05$). For reference, naive marker log-fold-changes from the same comparison are shown in a separate heatmap (column ordering derived only from constitutive response gene clusters).

For the identification of gene cluster-specific functional GO enrichment, we again made use of SETHRO, recovering informative annotations in every cluster except cluster 10 (Table 7), which consisted of only 26 genes. Cluster 3, which smoothly correlated with naive pluripotency marker changes in N2B27 (Fig 65), was enriched for response to growth factor stimulus and stem cell population maintenance and included the naive pluripotency markers themselves (Table 7). Several other pathways strongly connected to pluripotency were enriched in at least one of the clusters (BMP signalling, Wnt signalling, SMAD protein signal transduction, ERK signalling, MAPK activity)(Table 7). This confirms the ability of our approach to recover meaningful functional clusters of pluripotency, both on the gene-level and the knockout-level.

Table 7: Knockout specifically regulated clusters of the Interaction$_{N2}$ knockout response.

Cluster (genes)	**KO specific regulation**	**Enriched functions**
1 (280)	**Down:** Csnk1a1	TAP binding, glutamate receptor signaling pathway, calcium channel activity, nerve development, axon guidance, platelet-derived growth factor receptor signaling pathway, positive regulation of extrinsic apoptotic signaling pathway, skeletal system morphogenesis, protein tyrosine/serine/threonine phosphatase activity, phosphatidylinositol bisphosphate binding, ..., negative regulation of MAP kinase activity, regulation of ERK1 and ERK2 cascade
2 (135)	**Up:** Csnk1a1, Eed	SMAD protein signal transduction, growth factor receptor binding, cellular response to tumor necrosis factor, positive regulation of transmembrane receptor protein serine/threonine kinase signaling pathway, regulation of cell-substrate adhesion, sphingolipid metabolic process, negative regulation of epithelial cell proliferation, BMP signaling pathway, pallium development, organonitrogen compound catabolic process
3 (78)	**Up:** Csnk1a1, Etv5, Fgfr1, Rbpj	response to temperature stimulus, negative regulation of cellular response to growth factor stimulus, stem cell population maintenance, amino acid transport, skin development, morphogenesis of a branching epithelium
4 (96)	**Down:** Csnk1a1, Jarid2n, Smg5, Zfp281, Zfp423	glial cell differentiation, cytokine activity, anchored component of membrane, cell chemotaxis, B cell activation, synaptic membrane, striated muscle cell differentiation, regulation of lipid metabolic process
5 (67)	**Up:** Raf1, Smg5, Trim71, Zfp281 **Down:** Jarid2n, Msi2	cell chemotaxis, epithelial cell migration
6 (52)	**Up:** Aff3, Nmnat2	cytosolic ribosome, structural constituent of ribosome, cytosolic large ribosomal subunit, cytosolic small ribosomal subunit, ribosomal small subunit biogenesis, rRNA processing, rRNA binding
7 (99)	**Up:** Csnk1a1, Eed, Suz12, Zfp281	skeletal system morphogenesis, bone morphogenesis, cartilage development, mesenchymal cell development, heart morphogenesis, positive regulation of smooth muscle cell proliferation, embryonic skeletal system morphogenesis, biomineral tissue development, epithelial to mesenchymal transition, G-protein coupled receptor activity, ... , negative regulation of Wnt signaling pathway
8 (299)	**Up:** Csnk1a1, Etv5, Rbpj, Zfp423	heparin binding, embryonic pattern specification, regulation of Notch signaling pathway, stem cell population maintenance, secretory granule, response to steroid hormone, stem cell differentiation, neuroepithelial cell differentiation, somatic stem cell population maintenance, myosin complex

9 (58)	**Down:** Csnk1a1	negative regulation of secretion by cell
10 (26)	**Up:** Csnk1a1, Eed, Suz12 **Down:** Jarid2, Smg5	-
11 (88)	**Down:** Csnk1a1, Jarid2n, Zfp281, Zfp423	mesenchymal cell development, positive regulation of cell growth, negative regulation of neuron differentiation, cellular calcium ion homeostasis
12 (62)	**Up:** Eed, Fbxw7, Suz12 **Down:** Smg5, Trim71	second-messenger-mediated signaling

4.3.10 LIF-dependent regulation of pluripotency

LIF is an interleukine-6 cytokine that plays a crucial role in the maintenance of the naive pluripotency state. Supplementation of LIF contributes to the ability of cells to self-renew by several avenues: Firstly, by activation of the Jak-Stat3 pathway, which is sufficient to maintain self-renewal of mESCs (Niwa et al. 1998). Secondly, by activation of the MAPK pathway, which is involved in mESC differentiation; inhibition of MAPK by MEKi promotes, but is not necessary for self-renewal (Matsuda et al. 1999; Burdon et al. 1999). Thirdly, by activation of the PI3K-Akt-mTOR pathways, which promote mESC proliferation (Welham et al. 2007). In our knockout screen, we used 2i medium without LIF, which robustly maintains pluripotency in cell lines from diverse genetic backgrounds (Ying et al. 2008). However, supplementation of LIF in 2i medium showed additional benefits for self-renewal (Ying et al. 2008), indicating that LIF and 2i downstream processes are partially independent. Moreover, knockouts that affect the pluripotency state might do so in a manner that recapitulates, at least in part, the effects of LIF.

To quantify the baseline LIF response, we first determined the additional effect of LIF supplementation on our *RC9* control cell lines in 2i medium. For this, our collaborators carried out additional Quantseq sequencing (Moll et al. 2014) on biological duplicates of *RC9* wild type cell lines in both 2i and 2i/LIF conditions and provided us with the raw counts. Quantseq is restricted to sequencing a region close to the 3' end of polyadenylated transcripts and produces only one fragment per gene (Moll et al. 2014). As therefore no bias should arise from gene length we did not use RPKM for filtering lowly expressed genes as before. Instead, we excluded transcripts with a median count of ≤ 5, which resulted in 13432 remaining genes. We then processed the count values using voom and limma as described before. We used TREAT to test the difference between $RC9_{2i}$ and $RC9_{2iL}$ directly against an absolute log-fold-change threshold of log2(1.5). Resulting p-values from the moderated t-test were corrected for multiplicity using the

BH-method. We detected 260 genes whose expression was significantly increased or decreased in *RC9* cells upon addition of LIF (at a significance level of α = 0.1).

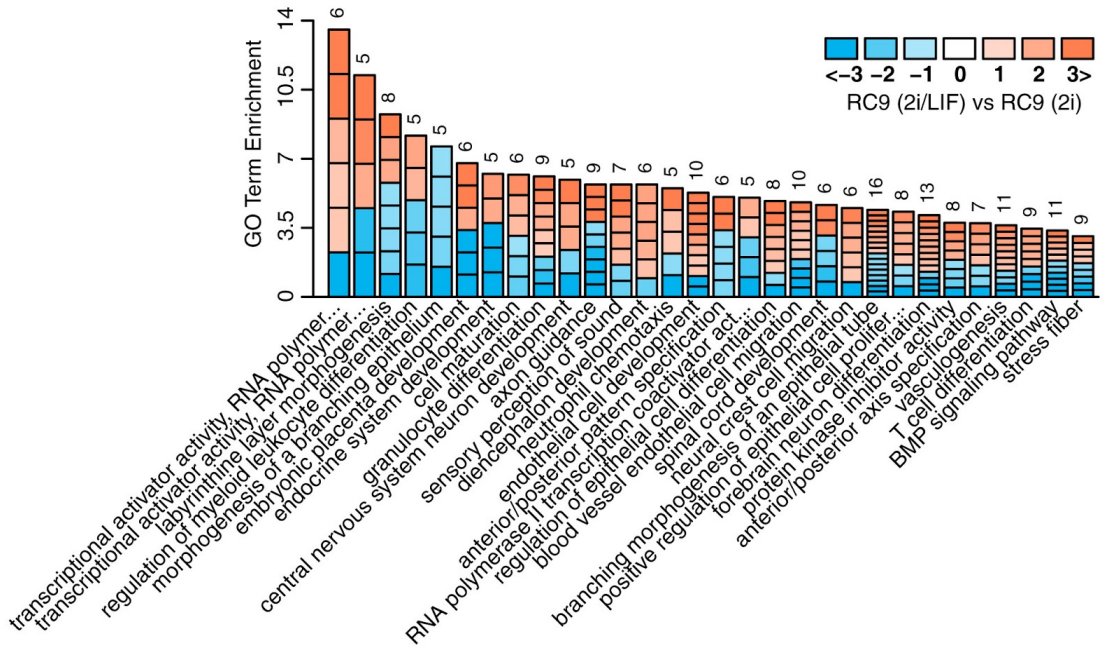


Fig 66***: Wild type response to LIF in 2i medium.***
Cell plot showing the 30 most enriched functional terms enriched in the RC9 LIF-response in 2i medium ($RC9_{2i/LIF}$ vs. $RC9_{2i}$). All terms were found significant after multiplicity correction (Fisher test, adj. P. Value ≤ 0.05).

We carried out GO enrichment analysis on the set of thus identified LIF-specific genes using SETHRO, confirming that LIF affected the expression of genes involved in differentiation and development (Fig 66). *Socs3,* a canonical downstream target induced by the Jak-Stat3 pathway (Starr et al. 1997), was increased by more than 128-fold (log2-fold-change = 7.2). While normally crucial for mESC self-renewal, Jak-Stat3 and *Socs3* induction is known to be dispensable in 2i-medium (Ying et al. 2008). Conversely, *Lefty1* expression was reduced 26-fold. *Lefty1* is a transforming growth factor β (TGF-β) inhibitor that disrupts BMP signalling and thereby promotes differentiation (Ulloa & Tabibzadeh 2001; Ying et al. 2003). *Lefty1* is present in both undifferentiated cells and differentiating cells (Ulloa & Tabibzadeh 2001) and is known to increase upon removal of LIF (D.-K. Kim et al. 2014). However, BMP is not necessary for self-renewal in 2i medium (Ying et al. 2008). These data suggest that our results are of high quality and in line with expectations from literature. Additional investigation is likely to uncover high confidence target genes not previously associated with LIF.

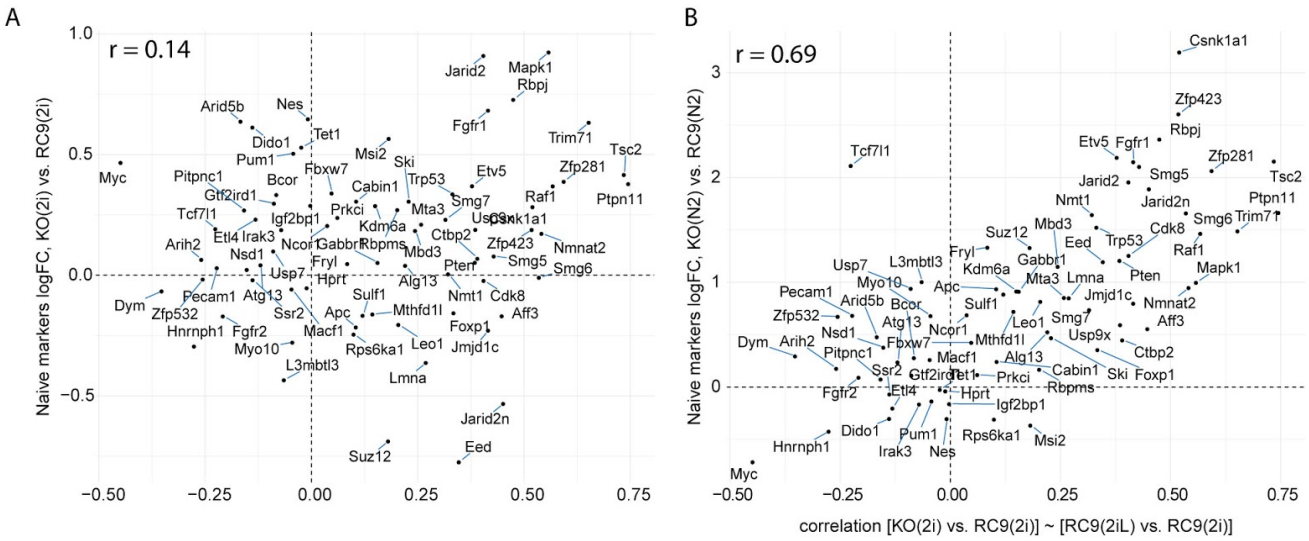


Fig 67***: LIFlikeness of knockouts is associated with delayed differentiation in N2B27.***
(A) *Scatter plot showing average change in naive marker expression induced by knockouts in* ***2i*** *(y-axis) and their LIF-likeness in 2i (x-axis, correlation of RC9 LIF response with knockout-induced response in 2i).* ***(B)*** *Scatter plot showing average change in naive marker expression induced by knockouts in* ***N2B27*** *(y-axis) and their LIF-likeness in 2i (x-axis, correlation of RC9 LIF response with knockout induced response in 2i).*

We next asked if we could identify knockout mutants that behaved in a LIF-like manner: does the respective knockout (in 2i medium) induce changes that are similar to the specific effect of LIF on wild type cells? For this, we selected the log-fold-changes from all 74 KO_{2i} *vs.* $RC9_{2i}$ comparisons for the 260 LIF-specific genes. We then correlated each knockout's expression pattern with the log-fold-changes induced by LIF in the wild type ($RC9_{2i/LIF}$ *vs.* $RC9_{2i}$), determining its "LIF-likeness". To determine whether LIF-likeness of a knockout could predict the strength of the differentiation delay phenotype in N2B27, we further correlated this correlation against the mean of naive marker log-fold-changes (KO_{N2} *vs.* $RC9_{N2}$). We found that mutants that induced LIF-like effects in 2i were indeed more likely to show abnormally high naive marker expression in N2B27 (r = 0.69)(Fig 67B). No such relationship was found with marker expression in 2i (r = 0.14)(Fig 67A).

One notable outlier was the *Tcf7l1* knockout cell line, which showed a strong differentiation delay phenotype as measured by marker expression, but whose effects in 2i were not correlated to LIF effects. *Tcf7l1* (aka *Tcf3*) is a repressor of several pluripotency marker genes including *Esrrb* and *Nanog (Martello et al. 2012)*, and can be inactivated through Wnt/β-catenin signalling. *Tcf7l1* suppression, just like *Wnt* activation (Ohtsuka et al. 2015), may therefore be independent of and complementary to LIF effects.

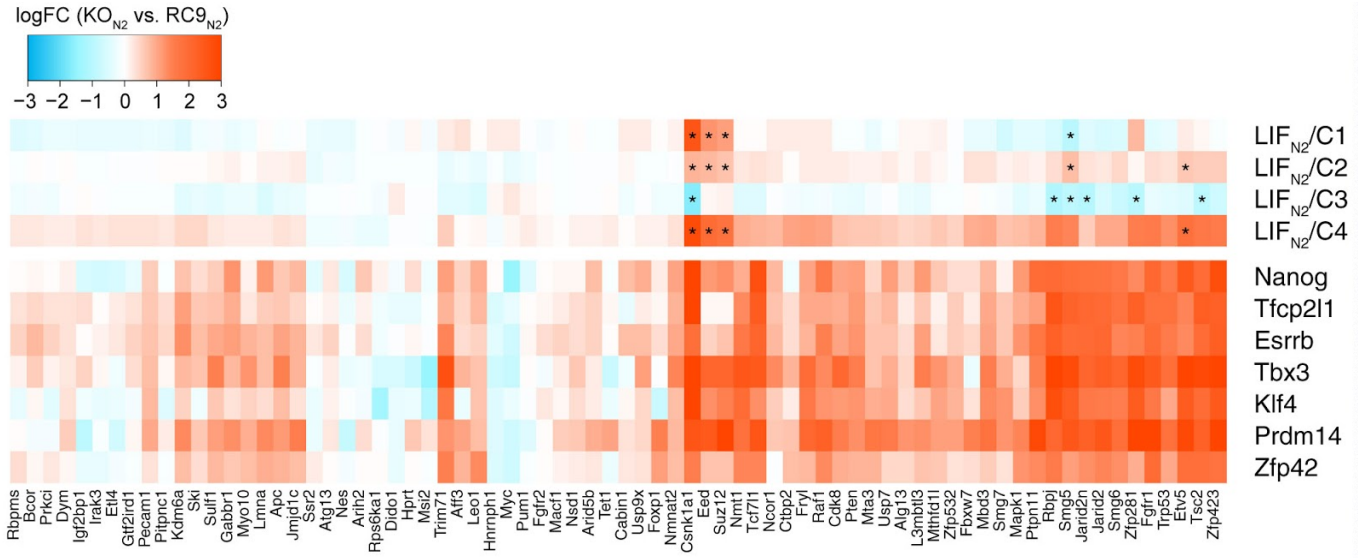


Fig 68***: Clusters of the knockout response of LIF-dependent genes in N2B27.***

Heatmaps showing average log-fold-changes (KO$_{N2}$ vs. RC9$_{N2}$) of LIF$_{N2}$ gene clusters. Asterisks indicate that the absolute row-wise z-value of the respective mean expression value was above 1.96 ($p \leq 0.05$). For reference, naive marker log-fold-changes from the same comparison are shown in a separate heatmap (column ordering derived only from LIF$_{N2}$ clusters).

To dissect the knockout response of LIF-dependent genes further, we first selected the 165 genes of the significant LIF response in 2i that were also significantly changed compared to their control in at least one knockout in N2B27. We then carried out cluster analysis using the Ward method as described in previous sections. We identified four clusters and calculated their mean log-fold-changes (*KO$_{N2}$ vs. RC9$_{N2}$*) as well as row-wise z-values (Fig 68). Clusters one through three were regulated in only a subset of knockouts, while cluster four regulation was positively correlated to marker expression in most knockouts (Fig 68). For each cluster, we carried out GO enrichment analysis using SETHRO (Table 8).

Table 8: Knockout-specifically regulated clusters of LIF-specific genes in N2B27.

Cluster (genes)	KO specific regulation	Enriched functions
LIF_{N2}/1 (29)	**Up:** Csnk1a1, Eed, Suz12 **Down:** Rbpj	-
LIF_{N2}/2 (40)	**Up:** Csnk1a1, Eed	embryonic placenta development
LIF_{N2}/3 (59)	**Down:** Csnk1a1, Rbpj, Smg5, Jarid2n, Zfp281, Tsc2	limb morphogenesis, glial cell differentiation
LIF_{N2}/4 (37)	**Up:** Csnk1a1, Eed, Suz12, Etv5	regulation of angiogenesis, regulation of vasculature development

4.3.11 Inhibitor interactions of pluripotency regulation

To screen for genetic interactions that may be complementary in producing differentiation delay in N2B27, we carried out an inhibitor screen using X of our knockout lines in the presence of one of five inhibitors. Firstly, the two components of 2i medium (Ying et al. 2008): CHIR99021 (CH), which mimics Wnt signalling activation by inhibiting *Gsk3* and preventing the degradation of β-catenin (Ying et al. 2008), and PD184352 (PD), an inhibitor of *Mek* (Allen et al. 2003). Secondly, LIF, which acts through activation of at least three pathways: Jak-Stat3 (Niwa et al. 1998), MAPK, (Matsuda et al. 1999; Burdon et al. 1999) and PI3K-Akt-mTOR (Welham et al. 2007). Thirdly, Cdk8i, which inhibits *Cdk8* expression, an oncogene and positive regulator of β-catenin and Myc (Firestein et al. 2008; Adler et al. 2012). Lastly, Gö, an inhibitor of atypical protein kinase C (PKC), specifically *Prkci (Leeb et al. 2014)*. Suppression of *Prkci* counteracts differentiation through modulation of the conserved Notch pathway (Dutta et al. 2011; Mah et al. 2015).

Our collaborators measured the percentage of undifferentiated cells (Zfp42-GFP positive) after 48h (two replicates) and 72h (one replicate) after placing cells from 2i/LIF into N2B27 medium. Additionally, all cell lines were placed in unmodified N2B27 medium as control. As a genetic control, *RC9* cell lines were included in all conditions. As the fractions of undifferentiated cells are by definition bounded between 0 and 1, they are expected to follow a beta distribution. To estimate the whether any interaction of knockout (*K*) and condition (*C*) resulted in a greater or lower fraction of undifferentiated cells than expected from the behaviour of their controls, we therefore fitted beta regression models to our data (Eq 19) using the betareg R package (Cribari-Neto & Zeileis 2010).

$$g(\mu_i) = \beta_0 + K_i C_i \beta_1 + K_i \beta_2 + C_i \beta_3 \qquad (19)$$

We then tested the significance of the interaction coefficients β_1 using a partial Wald test. A negative coefficient indicates a lower proportion of undifferentiated cells than expected, whereas a positive coefficient indicates an increase over expectation.

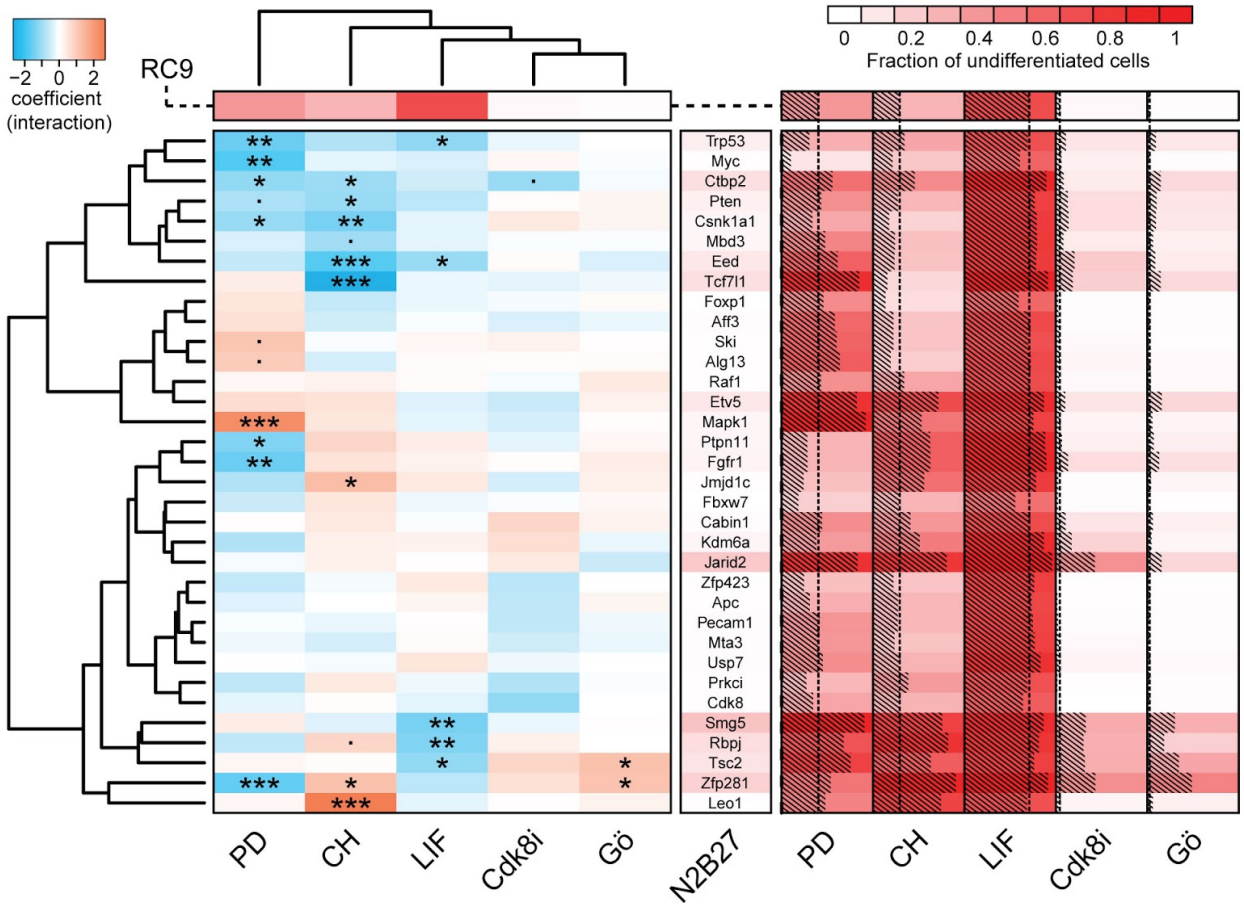


Fig 69***: Inhibitor interactions of pluripotency regulation.***
Left: *Interaction coefficients of the beta regression model interaction term, for all inhibitor-knockout combinations. Asterisks indicate significance of the coefficient (.p ≤ 0.1, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001). A significant positive value indicates that a higher fraction of undifferentiated cells was observed than expected (dependent on both wild type levels and N2B27 levels), while a significant negative value indicates the opposite.* ***Right:*** *Heatmap showing the average fraction of undifferentiated cells in each inhibitor-knockout combination as well as the RC9 wild type and N2B27 (no inhibitor) control medium. Shaded bars represent the fractions in each combination (redundant with colour information). Dotted lines indicate wild type levels for easier comparison.*

Under PD condition, most significant interaction coefficients (Wald test p-value ≤ 0.05) were negative: those of *Trp53, Myc, Ctbp2, Csnk1a1, Ptpn11, Fgfr1* and *Zfp281* knockouts (Fig 69). We investigated the *Zfp281* knockout more closely, as it has a strong control (N2B27) phenotype that is not exacerbated by PD (Fig 69). *Zfp281* is essential for stem cell priming and known to be

down-regulated by PD (Fidalgo et al. 2016). This suggests that PD and *Zfp281* knockout delay cell differentiation through overlapping mechanisms. Surprisingly, *Mapk1 (Erk2)* knockout cells showed a dramatic increase in the number of undifferentiated cells in the presence of only PD. As PD inhibits signal transduction via the Mapk/Erk pathway (Allen et al. 2003), we expected a redundant phenotype similar to *Zfp281* knockouts. This suggests that in the absence of a signal from *Mek*, *Mapk1* might be able to be activated through an alternative mechanism. Inhibiting both *Mek* and knocking out *Mapk1* might therefore more completely inhibit the downstream ERK pathway that positively regulates cell proliferation.

In the CH supplemented medium, three knockouts showed a positive interaction: *Leo1*, *Zfp281*, and *Jmjd1c*. The synergistic effect with *Zfp281* further supports the notion that *Zfp281* may be partially redundant with PD and therefore fulfill a similar function in tandem with CH. On the other hand, five knockouts were less responsive than expected: Tcf7l1, Eed, Csnk1a1, Pten, and Ctbp2.

In the LIF condition, we identified five knockouts that fell short of the expected increase in undifferentiated cells: *Tsc2*, *Rbpj*, *Smg5*, *Eed,* and *Trp53* (Fig 69). These knockouts also show differentiation-delay phenotypes in the N2B27 control condition, indicating that their effects may be partially redundant with LIF activity.

4.4 Discussion

Dissecting the precise mechanisms by which pluripotent stem cells exit their naive pluripotent state and choose their differentiation trajectory is of crucial importance for applications in regenerative medicine (Tabar & Studer 2014; Mizuno et al. 2012). A multitude of studies have focused both on the naive state itself (Young 2011), or on the conditions that trigger differentiation (Ying et al. 2003; Wang et al. 2012; Leeb et al. 2014). The generation of haploid mouse embryonic stem cells (mESC's) has greatly facilitated the search for the causal drivers of priming and differentiation (Elling et al. 2011; Leeb & Wutz 2011; Leeb et al. 2014). However, these screens are still restricted to identifying the responses to single knockouts. Complex molecular signalling networks are characterized by a high degree of redundancy (Gitter et al. 2009; Macneil & Walhout 2011), suggesting that multiple parallel or interlocking mechanisms may govern cell fate decisions (Leeb et al. 2014). Due to the exponential search space of possible experiments with two or more simultaneous knockouts, as well as the associated costs, a brute-force or random sampling strategy is infeasible. Therefore it is crucial to pursue a systems level approach and characterize the interdependencies of regulatory network components to maximize information from single knockout screens.

In this project, we carried out the analysis of RNAseq data from one such large single knockout screen. Clonal cultures of 74 knockout mutants were kept in 2i medium, thereby maintaining pluripotency, or placed in N2B27 differentiation medium. The knockouts were selected by our collaborators based on their expression of a canonical marker of naive pluripotency, Zfp42 (Rex1), in N2B27 culturing conditions, indicating different degrees of differentiation delay. As controls, knockouts that accelerate differentiation, KO_{Myc} and KO_{Nes}, were included. We then integrated these data with information from two supporting experiments: the identification of LIF-specific target genes, and the direct quantification of differentiation delay caused by a combination of mutants and one of five inhibitors that affect differentiation. These data will be made available to the research community, both directly and through an interactive web interface, providing a high quality and comprehensive resource.

At the time of writing, this project is actively being developed. However, through the analysis presented in this thesis, we have already achieved substantial progress towards several key goals.

4.4.1 High confidence characterization of early wild type differentiation

We sequenced a total of 14 wild type samples in each condition, N2B27 and 2i. Sample sizes in RNAseq experiments are often low, with experiments commonly carried out in triplicate. Though our high replication of wild type samples was motivated primarily by our experimental design, i.e. the need to include control condition samples in every sequencing batch to allow for

effective batch correction, it also considerably boosted our power to detect differences between N2B27 and 2i wild type samples (Liu et al. 2014). We could identify substantial changes during wild type early differentiation in 3058 genes at very high stringency (|logFC| > log2(1.5), adj. P. Value ≤ 0.05)(Fig 54A). We confirmed that the identified genes were highly enriched for known markers of pluripotency (Fig 54). However, nearly 20% (589) of the responding genes were not annotated for any GO terms with less than 500 members, and 2% (59) were unnamed genes or predicted genes. These genes responded to withdrawal of 2i by changes in mRNA expression by up to 87-fold and may therefore represent important, entirely novel components of the differentiation response. Further work on this project will allow the informed selection of promising but unannotated candidate genes. Our companion app mESCexplorer (appendix) allows the selection of associated genes based on a co-expression network calculated from our data.

4.4.2 Global regression analysis of differentiation delay

To identify the genes responsible for the observed differentiation delay across our 74 knockout mutant cell lines, we compared them to wild type expression levels in both conditions. Each knockout and condition was represented by only two biological replicates. Nonetheless, due to the high precision of our control measurements (14 replicates per condition), we achieved high power in our differential expression analysis, identifying up to 3553 genes per knockout that differed substantially between knockouts and wild type (|logFC| > log2(1.5), adj. P. Value ≤ 0.05)(Fig 55). Importantly, the observed phenotypes spanned a wide range from severe to moderate differentiation delay, to acceleration of differentiation-related changes (Fig 56). This rich variation allowed us to quantify the degree to which genes were co-expressed with naive pluripotency markers in N2B27 with high confidence. The preliminary functional analysis of differentiation-associated genes showed that marker-dependent genes were mostly down-regulated during normal differentiation and enriched for negative regulators of development (Fig 58). Conversely, positive regulators of differentiation were more likely expressed independently of naive markers and up-regulated (Fig 58). Interestingly, differentiation-associated genes with a high marker-association were better annotated: only 15% (out of 306 genes in the top 10% quantile) lacked any GO annotation with less than 500 members, compared to 27% of genes with low marker-association. It will be of interest to further characterize these poorly annotated genes and fill gaps in our understanding of the core pluripotency network.

4.4.3 Partitioning the anti-differentiation response

A primary goal of this project is the identification of parallel or interlocking mechanisms by which pluripotent cells can prevent differentiation and maintain self-renewal. We therefore classified the response observed in our knockout mutants in N2B27, with regard to wild type

and knockout expression patterns in both conditions. We identified 16 gene clusters, encompassing 2787 genes, that were up- or down-regulated constitutively by at least one knockout in both 2i and N2B27 (Fig 60). We additionally characterized 12 gene clusters, encompassing 1340 genes, that were conditionally activated or repressed by knockouts in N2B27 (the InteractionN2 subset)(Fig 60). Our preliminary functional analysis reveals that overall, many of the canonical pathways involved in maintenance of pluripotency or differentiation were enriched in at least one cluster (such as BMP signalling, Wnt signalling, SMAD protein signal transduction, ERK signalling, MAPK activity). However, we find that even within groups of knockouts with similar differentiation delay phenotypes, individual clusters may be highly heterogeneous. This suggests that, indeed, different parallel paths may lead to the same differentiation delay. We suggest that combinations of functionally complementary knockouts (or their key targets) as identified by our cluster analysis may guide the design of complete and precise differentiation defect mutants.

4.4.4 Identifying genetic interactions of pluripotency

To directly identify epistatic interactions leading to differentiation delay phenotypes, we carried out two auxiliary experiments:

First, we determined the precise effect of LIF addition to wild type mESC cultures in 2i medium. While LIF is not necessary for stem cell self renewal in 2i, it promotes pluripotency further and improves cell viability (Ohtsuka et al. 2015). We therefore hypothesized that LIF-specific genes, i.e. genes that change upon addition of LIF to wild type cultures kept in 2i medium, might be important to the differentiation delay observed in our mutants. Indeed, we found a strong positive association between the LIF-likeness of knockouts in 2i and their eventual differentiation delay phenotype in N2B27 (Fig 67B).

We sought to corroborate possible interactions with inhibitors more directly and carried out a differentiation screen, observing the effect (% undifferentiated cells) of five different inhibitors in a subset of 34 knockout mutant backgrounds. We identified 29 significant interactions of inhibitor and knockout effects – i.e. combinations that resulted in a significantly higher or lower proportion of undifferentiated cells compared to expectation (from both N2B27 medium and RC9 genetic controls)(Fig 69). Nine of these are positive interactions, indicating complementary effects, while the negative interactions suggest redundant effects (Fig 69). Ongoing work focuses on cross-referencing this information with functional changes observed in our knockout screen.

Curiously, the response to knockout of *Tcf7l1*, a negative regulator of pluripotency that is repressed by Wnt/β-catenin, was independent of LIF-specific changes in spite of its strong differentiation delay phenotype (Fig 67B). This makes it unique among the strong differentiation delay phenotypes we observed, which generally induced LIF-specific changes (Fig 67B).

Wnt/β-catenin acts synergistically with LIF (Ogawa et al. 2006) and affects transcription through at least two separate branches: firstly, binding of Tcf1/Lef, which predominantly leads to activation of target gene expression (Cadigan & Waterman 2012). Secondly, repression of *Tcf7l1* activity by disassociation from DNA (Shy et al. 2013), which promotes self-renewal. Our results therefore suggest that *Tcf7l1* acts downstream and independently of LIF-specific effects, which is unusual in the context of our knockout screen and merits further investigation. Moreover, the effect of *Tcf7l1* knockout on pluripotency is largely redundant with addition CHIR99021 (Fig 69), suggesting that *Tcf7l1* repression may be the primary mechanism through which it affects pluripotency.

4.4.5 Promoting collaboration and transparency through an interactive web interface

We have developed mESCexplorer, a companion app that will give access to all data, analyses, and interactive visualizations related to this project (see appendix for more details). While web interfaces for standardized bioinformatics applications have been common for some time—NCBI has offered the ability to run BLAST on its website for more than 20 years (Altschul et al. 1997)—it has only recently become feasible and accepted to build specialized applications for individual research projects. Despite significant benefits, foremost among them transparency and ease of access (Perkel 2018), adoption of interactive tools in regular research publications has been slow. The introduction of living figures by F1000 research was one of the first steps in that direction (Singh Chawla 2015). Recently, F1000 has begun supporting plotly, a more standardized and actively developed commercial platform for the implementation and hosting of interactive charts (Ingraham 2017)(https://plot.ly). Beyond simple visualizations, full or partial analytic workflows can be transformed into the backend of web applications by frameworks such as Dash (https://plot.ly/products/dash/) or the Shiny web framework for R (Chang et al. 2015). Implementing mESCexplorer in R shiny serves a dual purpose: collaborators benefit from a central resource to access data and analyses during the development of the project in a way that is less limited to the specific viewpoint of the bioinformatician, hopefully helping to reduce the analytical bottleneck (Dasgupta et al. 2015). After publication, the purpose of the tool will be to increase transparency of the analysis and increase confidence in the results, as well as making them easier to re-use in future projects.

Despite these benefits, some caveats and potential improvements must be kept in mind. Firstly, maintenance and availability of the web app is associated with running costs. Due to the limited lifespan and high turnover of research projects, resources developed for these projects may fall into disrepair. A partial solution is the bundling of the Shiny app into a standalone package that is offered as a simple download, which is cheaper to maintain. This package may then be executed on any machine that runs R and has the requisite dependencies installed. However,

these dependencies represent additional possible points of failure; more thorough solutions, such as Docker containers that package the complete software environment needed (Boettiger 2014), should be considered in future projects, but may require a higher amount of effort for initial setup. Secondly, mESCexplorer is very specialized and desired capabilities may not be added after the initial development, nor potential issues resolved. It is therefore important to provide data and results in easy-to-process data formats outside of the app.

4.5 Methods

4.5.1 Identifying naive pluripotency network entry points using stability selection

We extracted each gene's expression in the N2B27 condition and calculated the correlations of marker gene expression (*Nanog, Tfcp2l1, Esrrb, Tbx3, Klf4, Prdm14, Zfp42*) to that of the remaining genes. To obtain a stable prediction of the most influential naive marker for each gene, we used stability selection (Meinshausen & Bühlmann 2010; Shah & Samworth 2013). On each iteration, *knockout:RC9* log-fold-changes were randomly assigned to two equal-sized partitions. We then calculated the R^2 value by correlating each gene with every marker, separately on both partitions of log-fold-changes. For each gene and partition, the marker ($q = 1$) with the highest R^2 value was selected (out of $p = 7$ possible markers). For each marker, we determine how often it is selected as best-correlated to the given gene in $B = 500$ iterations (1000 sub-samples). The significance of the selection frequency was calculated from all parameters using the function stabsel_parameters() in the stabs R package, under assumption of r-concavity (Shah & Samworth 2013). A selection frequency of 100% was achieved in only 3 cases, indicating that the chosen sampling resolution was sufficient.

5 Appendix

5.1 R tools

5.1.1 SETHRO: simple enrichment testing for highly redundant ontologies

Gene Ontology (GO) terms can be highly redundant. This is a consequence of the hierarchical structure of GO, which is a directed acyclic graph (DAG)(Ashburner et al. 2000; Gene Ontology Consortium 2001). Thus, specific terms are always subsets of more general terms. Additional redundancies can appear in parallel branches of the GO graph, especially between the three primary domains / root nodes (Molecular Function, Biological Process, and Cellular Component). There is a strong incentive to try to reduce these redundancies: firstly, it alleviates the multiple testing problem, as fewer tests need to be performed. Secondly, it facilitates the interpretation of results by making lists of terms enriched in a geneset of interest more concise.

One possibility to combat redundancy in GO is to use GO slims, which are curated subsets of the GO graph for specific purposes or organisms (Anon n.d.); this approach however sacrifices a lot of information. The elim algorithm in the R package topGO penalizes redundancy by using the DAG structure directly: leaf terms are tested for enrichment first. If the result is significant, genes of interest are then penalized when calculating the significance of more general ancestor terms (Alexa & Rahnenführer 2009). We used topGO in the analysis presented in the first chapter of this thesis, but found that lists of enriched terms were still frequently large and contained many redundant terms. Also, elim results in complicated dependencies that are not transparent to the user. A third approach is the clustering of annotations, as employed by DAVID (Huang et al. 2007). However, DAVID is server-based and does not integrate well into the analysis of large-scale results in the R environment.

We therefore developed SETHRO for internal use, which is shorthand for the combination of two conceptually simple methods:

1. Classical GO enrichment testing using Fisher tests. This is a standard method (Rivals et al. 2007) performed using the fisher.test function implemented in R, and is at the core of most software packages that perform functional enrichment analysis.
2. Clustering of GO terms based on their similarity in *genes of interest*.

To cluster GO terms, SETHRO first determines which genes of interest (e.g. genes significantly changed between conditions) are included in each term. Each term's genes of interest are encoded in a binary membership vector *m* whose length equals the number of all annotated genes. An element m_i equals 1 if the gene of interest *i* is annotated for that term, and 0 otherwise. SETHRO selects all terms with a minimum number of genes of interest (the default

value is 5). It then calculates the pairwise manhattan distances (l_0 norms) of all member vectors. For any pair of terms, this distance equals the number of genes of interest that are different between the two terms (missing or additional). SETHRO then clusters terms, using complete linkage hierarchical clustering – i.e. the linking function evaluates the distance between two clusters as the maximum distance between any of their elements. The thus obtained cluster hierarchy is then cut at a specified height *t* (the default value is 5) to obtain discrete clusters of GO terms. To summarize, a GO term cluster is a set of functional terms that contain at least *n* genes of interest and differ by at most *t* genes from each other.

After calculating term clusters, a representative term is chosen for each cluster. This is either the most specific term (lowest number of total member genes), or, in the case of ties, the most specific term with the shortest name. The output table of SETHRO provides information about all member terms of the cluster, its representative term, as well as all genes that are included in the representative term (or, separately, one of its co-terms). P-values are corrected for multiplicity using the Benjamini Hochberg method, considering representative terms only. This is justified under the assumption that the term clustering is a filter independent of the Fisher test statistic: while the clustering depends on the genes of interest, it uses no information about their distributions in the background gene set.

SETHRO is freely available as an R package on github:
https://github.com/robertsehlke/SETHRO

An implementation as Shiny web app with an example dataset is also available:
https://github.com/robertsehlke/shinySETGO (source code)
http://shiny.robertdoes.science/shinySETGO (running example)

5.1.2 CellPlot, an R package for the integrated visualization of functional enrichment and expression data

Functional analysis of gene sets of interest is commonly the first step in the characterization of results from gene perturbation experiments. Several ontologies exist for that purpose, the most prominent ones being Gene Ontology (Ashburner et al. 2000; Gene Ontology Consortium 2001) and KEGG pathways (Kanehisa et al. 2012). The output of a functional enrichment analysis is an enrichment score with associated p-value, indicating whether functional categories were enriched in the gene set of interest (compared to the background distribution among all genes). It is common practice to visualise the enrichment or negative log-transformed p-values in simple bar charts, discarding information about differential expression of individual genes. More advanced visualization tools like REVIGO (Supek et al. 2011) also do not integrate molecular data. The recently released GOplot package attempts to address this issue (Walter et al. 2015),

but accommodates only a limited number of terms to be displayed concurrently. We have developed CellPlot, an intuitive R package that integrates gene log-fold changes from perturbation experiments and the enrichment of their functional categories (e.g. Gene Ontology terms or KEGG pathways). CellPlot is a conceptually simple visualization package designed to fit seamlessly into most contexts where GO enrichments are presented.

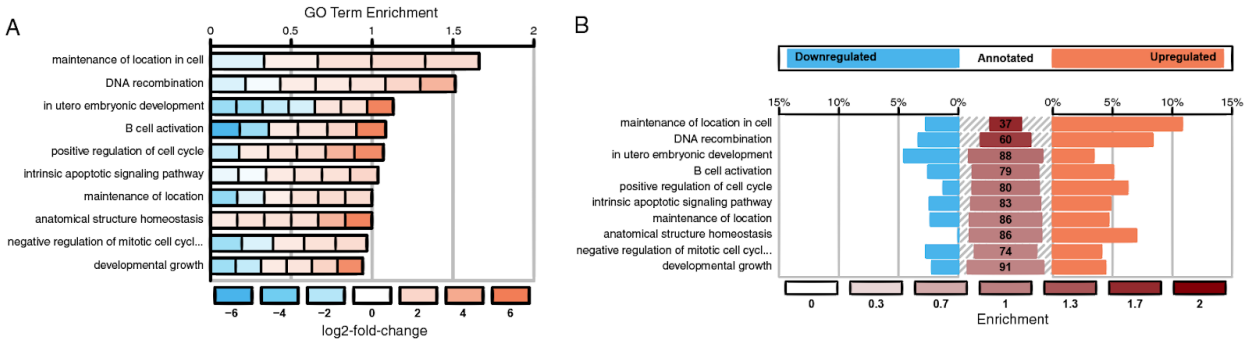


Fig 70**: Example output from the cell.plot() and sym.plot() functions.**
(A) *Example CellPlot based on differential expression results from* Golub et al. (Golub et al. 1999)*. The length of bars describes the log2-transformed GO term enrichment. Individual cells correspond to significant member genes and are coloured according to their observed log-fold-change between conditions.* ***(B)*** *Example SymPlot, based on the same data and analysis as in (A). The central column shows the number of all annotated genes in the respective term, with individual bars coloured according to the term's log2(enrichment). Bars to the left indicate the percentage of member genes that were in the gene set of interest and down-regulated, while bars to the right indicate percentage of genes of interest that were up-regulated.*

The function `cell.plot` combines the traditional bar charts with a one-dimensional heat map (Fig 70 A). Bars are segmented into a number of cells equal to the members of interest in the respective functional term. For each cell, the colour maps to some continuous property, e.g. gene expression fold changes of all genes in a GO term. Bar width traditionally represents enrichment or negative log-transformed p-value of the enrichment, but any vector of weights can be specified. The function `sym.plot` focuses on visualizing proportions within functional terms (Fig 70B). The middle column shows the overall size of each term along with the observed enrichment in the total gene set. Bars to the left and right of the middle column show the percentage of genes that are significantly down- or up-regulated. This makes it easy to spot homogeneous categories, e.g. GO terms where most member genes are up- or down-regulated in concert, and systemic biases. Additional functionality of the package includes the simultaneous plotting of multiple enrichment analysis for direct comparison using the `go.histogram` function (Fig 5), as well as direct integration with the topGO package via adapter functions.

The plotting functions offer a wide range of visual parameters to adjust their appearance, as well as basic options for filtering and sorting based on the number of members in each

functional term. We use a generic input scheme easily adaptable to different data formats. Required parameters are:

- Bar weights: a vector of positive values representing functional enrichment or significance.
- Cell weights: a list of vectors providing values for each cell of a term (e.g. gene fold changes). In the case of `sym.plot`, only the numbers of up- or down-regulated genes are required.
- Term size: a vector providing total size of functional categories (`sym.plot` only).

The CellPlot package is freely available on github (https://github.com/dieterich-lab/CellPlot). We include an example workflow using microarray data from Golub et al. (Golub et al. 1999), taken from the `multtest` R package (Pollard et al. 2005).

5.1.3 mESCexplorer, a web interface for interactive data analysis and exploration

To facilitate the interpretation of our analysis results within our group of collaborators, as well as to improve transparency of results for publication, we have developed mESCexplorer (Fig 71), using the shiny web framework for R (Chang et al. 2015).

5.1.3.1 General app layout

mESCexplorer consists of multiple elements that allow the user to access summary statistics and visualizations of several parts of our analysis outlined in previous sections of this chapter. It consists of two main parts:

Data tabs (on the left), displaying statistics and auxiliary information on the level of individual genes, gene sets, and pre-computed gene clusters (Interaction_{N2} and constitutive gene sets, as described earlier in this chapter).

Auxiliary tabs (on the right), showing visualization and additional information for items selected in data tabs (genes, gene sets, or clusters). Central to this section of the app is the knockout sample t-SNE map, which shows an interactive t-SNE projection of all knockout samples in relation to the wild type. Currently, the projection can be based on naive marker expression, N2B27 expression of *RC9*-differentiation-associated genes ("N2 (diff.)"), expression of all genes in 2i, or expression of all genes in both N2B27 and 2i. The most important features of the knockout map are **(a)** the selection of knockout samples or groups of samples to calculate aggregate statistics in the data tabs, and **(b)** the visualization of several knockout-wise statistics by colour, such as individual genes log-fold-changes, gene set combined p-values, or averaged z-values of clusters (as explained in the following subsections).

5.1.3.2 Gene level information

Differential expression analysis results of individual genes are accessible in the “Genes” tab. Two types of information are available in the main table:

The first type depends on the active selection of knockout samples. If more than one knockout sample is selected, values averaged over the selection will be shown. These conditional statistics include p-values of *knockout:RC9* interaction, KO_{N2} *vs.* $RC9_{N2}$, and KO_{2i} *vs.* $RC9_{2i}$ coefficients, as well as the coefficients (log-fold-changes) themselves. Additionally, z-values over the distribution of all knockouts are shown for KO_{N2} *vs.* $RC9_{N2}$ and KO_{2i} *vs.* $RC9_{2i}$ log-fold-changes.

The second type are global statistics that are independent of the knockout sample selection. These currently include only the naive marker R^2 value, i.e. the shared variance of the gene expression vector with the expression patterns of naive pluripotency markers.

Additional information is shown below the main table upon selection of a table row (gene). This includes a line chart of the gene’s expression level in 2i and N2B27, for the *RC9* wild type as well as any selected knockout samples. Also, the user may investigate the neighborhood of the selected gene in either its 2i or N2B27 co-expression network (see section 5.1.3.3). Further tabs to the right of the line chart display all GO annotations and reactome pathways the gene is annotated with.

5.1.3.3 Geneset level information

Aggregated differential expression results on the basis of gene sets are available in the “Genesets” tab. Generally, gene sets are genes grouped by a common annotation. The user begins by selecting the “Gene set source”: N2 network modules, 2i network modules, GO terms, Reactome pathways, or others. Summary p-values of gene sets are calculated from one of three gene-level comparisons, to be selected in the “Comparison” drop-down menu: “Interaction” (*knockout:RC9*), “N2, KO vs. WT” (KO_{N2} *vs.* $RC9_{N2}$), and “2i, KO vs. WT” (KO_{2i} *vs.* $RC9_{2i}$). We made of the Fisher method (Fridley et al. 2010) to combine the p-values of individual genes into a gene set p-value. If more than one knockout sample is selected, the combined p-value in all selected samples is averaged.

Network modules are non-overlapping sets based on a co-expression network learned on N2B27 or 2i expression data using the GENIE3 package (Huynh-Thu et al. 2010). GENIE3 is a random forest based algorithm suited to detect linear as well as non-linear associations (edges) between gene expression patterns. We thresholded the edge weights calculated by GENIE3 at 10% of the weight of the strongest edge to obtain a discrete directed network. We then partitioned this network into modules using the infomap algorithm as implemented in igraph (Rosvall & Bergstrom 2008). Enriched biological functions in each network module and their BH-corrected significances, i.e. GO terms and reactome pathways, were calculated using

SETHRO. Upon selecting a network module in the main table, the full list of annotations (with ≥ 5 genes) are displayed on the right-hand side of the app under “Network module annotations”.

The remaining choices of gene sets, GO terms, Reactome pathways, and Others, can be mutually overlapping. Others contains curated sets of genes, such as naive pluripotency markers, or all genes contained in the “PluriNet” (Müller et al. 2008).

5.1.3.4 Pre-computed clusters

Results of the cluster analysis presented in this thesis are accessible in the “Pre-computed clusters” tab. The clustering, encompassing the constitutive knockout response clusters and $Interaction_{N2}$ clusters, can be selected by drop down menu (“Clustering”). This will initialize a heatmap showing the mean of KO_{N2} *vs.* $RC9_{N2}$ log-fold-changes for each cluster and knockout, along with the corresponding naive marker log-fold-changes (similar to Fig 62). A cluster is then selected for further inspection. Cluster averaged log-fold-changes, or their corresponding z-values from row-normalization, are visualized on the knockout t-sne map. The results of a GO enrichment analysis of the cluster genes (using SETHRO) are displayed in the main table. Via the “go to cluster genes” button it is furthermore possible to switch to the “Genes” tab and inspect, sort, and filter all cluster genes.

5.1.3.5 Development and outlook

Current development of mESCexplorer focuses on two areas:
Firstly, providing comprehensive documentation, as well as downloads of plain data files underlying the analysis, directly from within the app. This step is crucial to ensure the data and analysis are easily accessible as a resource for other researchers in the field.

Secondly, the streamlining of analysis tools. Currently, multiple redundant ways of analysis are integrated in the app, reflecting its use as a collaborative tool that is frequently being adapted to new developments in the underlying analysis. The final objective is focus mESCexplorer on reproducing the workflow presented in the eventual publication of this project.

In summary, mESCexplorer is a web app designed to provide transparent access to data analysed in this project. It will be released as a freely available resource upon publication of the project.

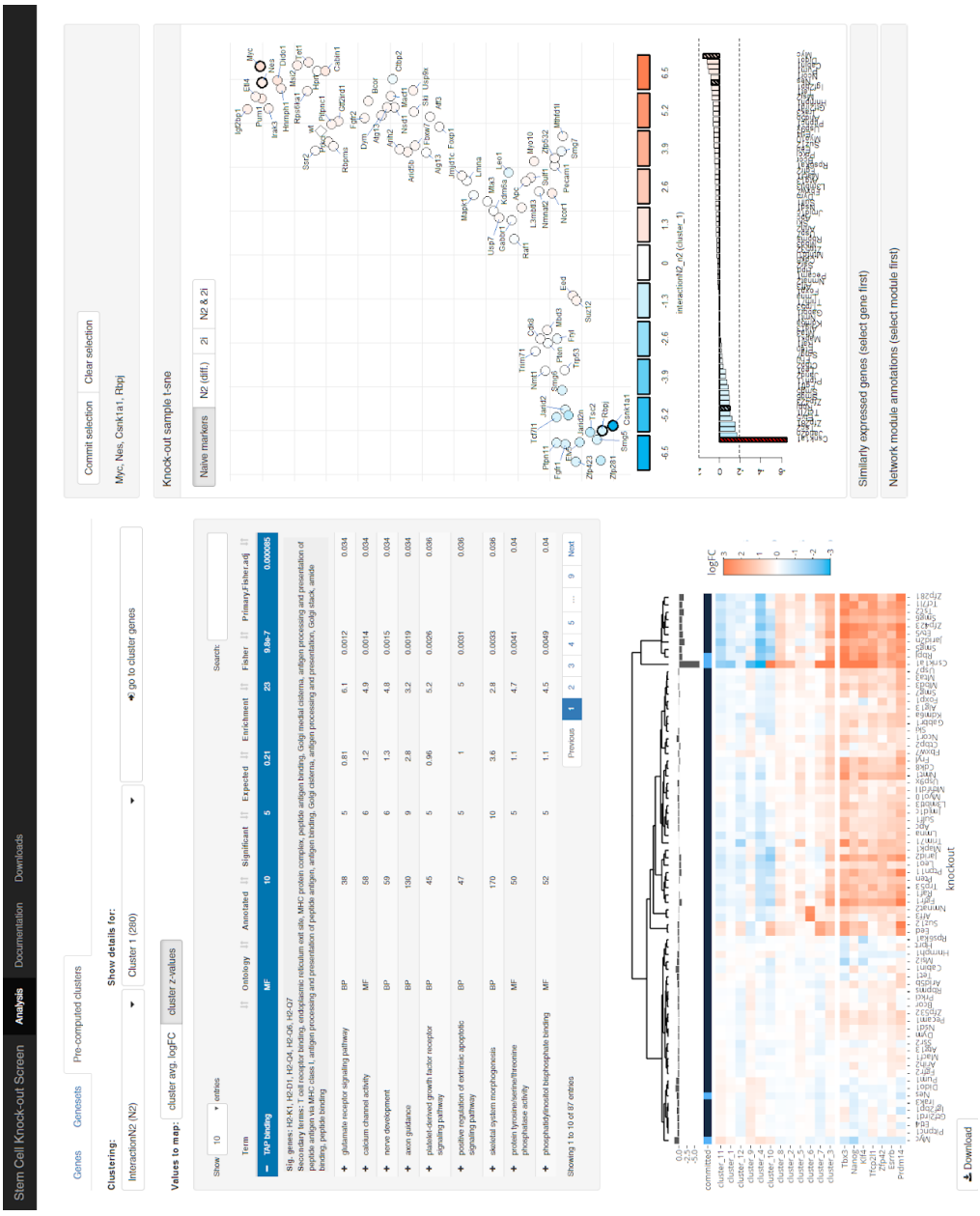


Fig 71***: the mESCexplorer web tool.***

Screenshot from the R Shiny web tool developed for this project. Selected is cluster one of the InteractionN2 subset clustering.

5.2 Supplementary files

Included with the digital copy of this thesis is an archive of supplementary files:

Supplementary_Files.zip

It can be verified that the archive is unchanged from publication using this MD5 checksum:

28A0C62C94D465CB3DFA2475CA4E7214

5.2.1 Files: Dynamics of fly insulin signalling

These files relate to chapter 3: Dynamics of fly insulin signalling across omics and age. Note that additional material, especially relating to chapter 3.3.2, has been published alongside (Tain et al. 2017).

Subfolder: **/Drosophila IIS**

#	File	Description
#1	Dilp post-transcriptional regulation.xlsx	● Table 1: differential expression, $dilp^{2\text{-}3,5}$ vs. w^{Dah} of proteins and matching transcripts ● Table 2: differential expression, $dilp^{2\text{-}3,5}$ vs. w^{Dah} , all transcripts ● Table 3: GO enrichment analysis, $dilp^{2\text{-}3,5}$ vs. w^{Dah}, significant protein/transcript pairs by quadrant
#2	Dilp and Ablation, proteomics integration.xlsx	● Table 1: differential expression in $dilp^{2\text{-}3,5}$ vs. w^{Dah} and matching w^{Dah}/Abl. vs. w^{Dah} ● Table 2: insulin model integration, correlation analysis
#3	Dilp age-dependent proteomics.xlsx	● Table 1: differential expression, TMT $dilp^{2\text{-}3,5}$ vs. w^{Dah}, d10->d30->d50 (age trends and interaction) ● Table 2: differential expression, TMT $dilp^{2\text{-}3,5}$ vs. w^{Dah}, individual insulin response at each time point
#4	integration_dilp_ablation_Fig16_top20_terms.pdf	Top 20 GO terms of each quadrant as shown in Fig. 16
#5	Dilp miRNA analysis.xlsx	● Table 1: all miRNA target predictions ● Table 2: differential expression of miRNA and targets, $dilp^{2\text{-}3,5}$ vs. w^{Dah}

5.2.2 Files: Regulatory networks of mouse stem cell pluripotency

These files relate to chapter 4: Regulatory networks of mouse stem cell pluripotency.

Subfolder: **/mESC**

#	File	Description
#6	ConstitutiveResponse_clusters_GO_analysis.csv	GO analysis, clusters of the constitutive response
#7	ConstitutiveResponse_notChangedInDifferentiation_clusters_GO_analysis.csv	GO analysis, clusters of the constitutive response unchanged during normal diff.
#8	InhibitorScreen_fraction_undifferentiated.csv	Inhibitor screen, fraction of undifferentiated cells
#9	InhibitorScreen_interaction_coefficients.csv	Inhibitor screen, coefficients of the interaction term from beta regression
#10	InhibitorScreen_interaction_pvalues.csv	Inhibitor screen, p-values of the interaction term from beta regression
#11	Interaction_adjPvalues.csv	Interaction (KO/RC9:N2/2i), BH-adjusted p-values
#12	Interaction_logFoldChanges.csv	Interaction (KO/RC9:N2/2i), log2-fold-changes
#13	InteractionN2_clusters_GO_analysis.csv	GO analysis, clusters of the InteractionN2 subset
#14	KO2i_vs_RC9_2i_adjPvalues.csv	Knockout response in 2i, BH-adjusted p-values
#15	KO2i_vs_RC9_2i_logFoldChanges.csv	Knockout response in 2i, log2-fold-changes
#16	KON2_vs_RC9N2_adjPvalues.csv	Knockout response in N2B27, BH-adjusted p-values
#17	KON2_vs_RC9N2_logFoldChanges.csv	Knockout response in N2B27, log2-fold-changes
#18	N2_LIF_specific_clusters_GO_analysis.csv	GO analysis, clusters of the LIF-specific response in N2B27
#19	RC9_2iLIF_vs_2i_LIF_specific.csv	LIF-specific response (RC9/2iLIF vs. RC9/2i), log2-fold-changes and (adj.) p-values
#20	RC9_2iLIF_vs_2i_LIF_specific_GO_analysis.csv	GO analysis, LIF-specific response (RC9/2iLIF vs. RC9/2i)
#21	RC9_N2_vs_2i_bottom10pc_markerRegression_GO.csv	GO analysis, Normal differentiation response (RC9/N2 vs. RC9/2i), bottom 10% by marker R^2
#22	RC9_N2_vs_2i_top10pc_markerRegression_GO.csv	GO analysis, Normal differentiation response (RC9/N2 vs. RC9/2i), top 10% by marker R^2
#23	RC9_N2_vs_2i_normal_differentiation.csv	Normal differentiation response (RC9/N2 vs. RC9/2i), log2-fold-changes and (adj.) p-values

#24	RC9_N2_vs_2i_top10pc_significant_GO.csv	GO analysis, Normal differentiation response (RC9/N2 vs. RC9/2i), top 10% by significance
#25	MarkerRegression_and_stability Selection.xlsx	Naive marker regression and stability selection results (correlations, frequencies, p-values)

6 References

Aboobaker, A.A. et al., 2005. Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development. *Proceedings of the National Academy of Sciences of the United States of America*, 102(50), pp.18017–18022.

Adler, A.S. et al., 2012. CDK8 maintains tumor dedifferentiation and embryonic stem cell pluripotency. *Cancer research*, 72(8), pp.2129–2139.

Afschar, S. et al., 2016. Nuclear hormone receptor DHR96 mediates the resistance to xenobiotics but not the increased lifespan of insulin-mutant Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 113(5), pp.1321–1326.

Aitken, C.E. et al., 2016. Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. *eLife*, 5. Available at: http://dx.doi.org/10.7554/eLife.20934.

Albert, F.W. et al., 2014. Genetic influences on translation in yeast. *PLoS genetics*, 10(10), p.e1004692.

Alexa, A. & Rahnenführer, J., 2009. Gene set enrichment analysis with topGO. Available at: https://www.bioconductor.org/packages/3.7/bioc/vignettes/topGO/inst/doc/topGO.pdf.

Alexa, A., Rahnenführer, J. & Lengauer, T., 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* , 22(13), pp.1600–1607.

Alic, N. et al., 2014. Cell-nonautonomous effects of dFOXO/DAF-16 in aging. *Cell reports*, 6(4), pp.608–616.

Alic, N. et al., 2011. Genome-wide dFOXO targets and topology of the transcriptomic response to stress and insulin signalling. *Molecular systems biology*, 7(1), p.502.

Allen, L.F., Sebolt-Leopold, J. & Meyer, M.B., 2003. CI-1040 (PD184352), a targeted signal transduction inhibitor of MEK (MAPKK). *Seminars in oncology*, 30(5 Suppl 16), pp.105–116.

Altschul, S.F. et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), pp.3389–3402.

Amador-Noguez, D. et al., 2007. Alterations in xenobiotic metabolism in the long-lived Little mice. *Aging cell*, 6(4), pp.453–470.

Ameres, S.L. et al., 2010. Target RNA–Directed Trimming and Tailing of Small Silencing RNAs. *Science*, 328(5985), pp.1534–1539.

Anders, S., Pyl, P.T. & Huber, W., 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* , 31(2), pp.166–169.

Anon, GO Slim and Subset Guide | Gene Ontology Consortium. Available at: http://www.geneontology.org/page/go-slim-and-subset-guide [Accessed March 12, 2018].

Apostolou, E. et al., 2013. Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell stem cell*, 12(6), pp.699–712.

Aradska, J. et al., 2015. Gel-free mass spectrometry analysis of Drosophila melanogaster heads. *Proteomics*, 15(19), pp.3356–3360.

Arava, Y. et al., 2003. Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), pp.3889–3894.

Aronson, N.N., Jr & Kuranda, M.J., 1989. Lysosomal degradation of Asn-linked glycoproteins. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 3(14), pp.2615–2622.

Artieri, C.G. & Fraser, H.B., 2014. Evolution at two levels of gene expression in yeast. *Genome research*, 24(3), pp.411–421.

Ashburner, M. et al., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, 25(1), pp.25–29.

Athanasiadis, A., Rich, A. & Maas, S., 2004. Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS biology*, 2(12), p.e391.

Azimifar, S.B. et al., 2014. Cell-type-resolved quantitative proteomics of murine liver. *Cell metabolism*, 20(6), pp.1076–1087.

Bagga, S. et al., 2005. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell*, 122(4), pp.553–563.

Bai, H. et al., 2013. Activin signaling targeted by insulin/dFOXO regulates aging and muscle proteostasis in Drosophila. *PLoS genetics*, 9(11), p.e1003941.

Bai, H., Kang, P. & Tatar, M., 2012. Drosophila insulin-like peptide-6 (dilp6) expression from fat body extends lifespan and represses secretion of Drosophila insulin-like peptide-2 from the brain. *Aging cell*, 11(6), pp.978–985.

Bailey, T.L. et al., 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic acids research*, 37(Web Server issue), pp.W202–8.

Bailis, J.M. & Forsburg, S.L., 2004. MCM proteins: DNA damage, mutagenesis and repair. *Current opinion in genetics & development*, 14(1), pp.17–21.

Baker, D.J. et al., 2006. No decline in skeletal muscle oxidative capacity with aging in long-term calorically restricted rats: effects are independent of mitochondrial DNA integrity. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 61(7), pp.675–684.

Barrett, L.W., Fletcher, S. & Wilton, S.D., 2012. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and molecular life sciences: CMLS*, 69(21), pp.3613–3634.

Bartke, A., 2011. Pleiotropic effects of growth hormone signaling in aging. *Trends in endocrinology and*

metabolism: TEM, 22(11), pp.437–442.

Barzilai, N. et al., 2016. Metformin as a Tool to Target Aging. *Cell metabolism*, 23(6), pp.1060–1065.

Batista, P.J. & Chang, H.Y., 2013. Long noncoding RNAs: cellular address codes in development and disease. *Cell*, 152(6), pp.1298–1307.

Batlle-Morera, L., Smith, A. & Nichols, J., 2008. Parameters influencing derivation of embryonic stem cells from murine embryos. *Genesis* , 46(12), pp.758–767.

Bazzocchi, C. et al., 2007. Wolbachia surface protein (WSP) inhibits apoptosis in human neutrophils. *Parasite immunology*, 29(2), pp.73–79.

Beddington, R.S. & Robertson, E.J., 1999. Axis development and early asymmetry in mammals. *Cell*, 96(2), pp.195–209.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 57(1), pp.289–300.

Benjamini, Y. & Yekutieli, D., 2001. The Control of the False Discovery Rate in Multiple Testing under Dependency. *Annals of statistics*, 29(4), pp.1165–1188.

Beretta, L. et al., 1996. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. *The EMBO journal*, 15(3), pp.658–664.

Blüher, M., Kahn, B.B. & Kahn, C.R., 2003. Extended longevity in mice lacking the insulin receptor in adipose tissue. *Science*, 299(5606), pp.572–574.

Boettiger, C., 2014. An introduction to Docker for reproducible research, with examples from the R environment. *arXiv [cs.SE]*. Available at: http://arxiv.org/abs/1410.0846.

Bonawitz, N.D. et al., 2007. Reduced TOR signaling extends chronological life span via increased respiration and upregulation of mitochondrial gene expression. *Cell metabolism*, 5(4), pp.265–277.

Bonnay, F. et al., 2013. big bang gene modulates gut immune tolerance in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 110(8), pp.2957–2962.

Boyce, M. et al., 2005. A Selective Inhibitor of eIF2α Dephosphorylation Protects Cells from ER Stress. *Science*, 307(5711), pp.935–939.

Bradley, A. et al., 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature*, 309(5965), pp.255–256.

Brandt, T. et al., 2017. Changes of mitochondrial ultrastructure and function during ageing in mice and Drosophila. *eLife*, 6. Available at: http://dx.doi.org/10.7554/eLife.24662.

Bratic, A. & Larsson, N.-G., 2013. The role of mitochondria in aging. *The Journal of clinical investigation*, 123(3), pp.951–957.

Brockmann, R. et al., 2007. Posttranscriptional expression regulation: what determines translation rates?

PLoS computational biology, 3(3), p.e57.

Brogiolo, W. et al., 2001. An evolutionarily conserved function of the Drosophila insulin receptor and insulin-like peptides in growth control. *Current biology: CB*, 11(4), pp.213–221.

Broughton, S. et al., 2008. Reduction of DILP2 in Drosophila triages a metabolic phenotype from lifespan revealing redundancy and compensation among DILPs. *PloS one*, 3(11), p.e3721.

Broughton, S.J. et al., 2010. DILP-producing median neurosecretory cells in the Drosophila brain mediate the response of lifespan to nutrition. *Aging cell*, 9(3), pp.336–346.

Broughton, S.J. et al., 2005. Longer lifespan, altered metabolism, and stress resistance in Drosophila from ablation of cells making insulin-like ligands. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8), pp.3105–3110.

Brown, C.J. et al., 2018. Proteome changes in the aging Drosophila melanogaster head. *International journal of mass spectrometry*, 425, pp.36–46.

Brunner, E. et al., 2007. A high-quality catalog of the Drosophila melanogaster proteome. *Nature biotechnology*, 25(5), pp.576–583.

Buch, S. et al., 2008. Opposing effects of dietary protein and sugar regulate a transcriptional target of Drosophila insulin-like peptide signaling. *Cell metabolism*, 7(4), pp.321–332.

Burdon, T. et al., 1999. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Developmental biology*, 210(1), pp.30–43.

Cadigan, K.M. & Waterman, M.L., 2012. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harbor perspectives in biology*, 4(11). Available at: http://dx.doi.org/10.1101/cshperspect.a007906.

Cadiñanos, J. & Bradley, A., 2007. Generation of an inducible and optimized piggyBac transposon system. *Nucleic acids research*, 35(12), p.e87.

Cardenas, M.E. et al., 1999. The TOR signaling cascade regulates gene expression in response to nutrients. *Genes & development*, 13(24), pp.3271–3279.

Carlevaro-Fita, J. et al., 2016. Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA* , 22(6), pp.867–882.

Carlson, K.A., Zhang, C. & Harshman, L.G., 2016. A dataset for assessing temporal changes in gene expression during the aging process of adult Drosophila melanogaster. *Data in brief*, 7, pp.1652–1657.

Caro, L.G. & Palade, G.E., 1964. PROTEIN SYNTHESIS, STORAGE, AND DISCHARGE IN THE PANCREATIC EXOCRINE CELL. AN AUTORADIOGRAPHIC STUDY. *The Journal of cell biology*, 20, pp.473–495.

Carrieri, C. et al., 2012. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature*, 491(7424), pp.454–457.

Cattie, D.J. et al., 2016. Mutations in Nonessential eIF3k and eIF3l Genes Confer Lifespan Extension and

Enhanced Resistance to ER Stress in Caenorhabditis elegans. *PLoS genetics*, 12(9), p.e1006326.

Cech, T.R. & Steitz, J.A., 2014a. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. *Cell*, 157(1), pp.77–94.

Cech, T.R. & Steitz, J.A., 2014b. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell*, 157(1), pp.77–94.

Celeste, A. et al., 2002. Genomic instability in mice lacking histone H2AX. *Science*, 296(5569), pp.922–927.

Cenik, C. et al., 2015. Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome research*, 25(11), pp.1610–1621.

Cesana, M. et al., 2011. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*, 147(2), pp.358–369.

Chambers, I. et al., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, 113(5), pp.643–655.

Chang, W. et al., 2015. Shiny: web application framework for R. *R package version 0. 11*, 1(4), p.106.

Chase, B.A. & Kankel, D.R., 1987. A genetic analysis of glutamatergic function in Drosophila. *Journal of neurobiology*, 18(1), pp.15–41.

Chen, B. et al., 2016. Genome-wide identification and developmental expression profiling of long noncoding RNAs during Drosophila metamorphosis. *Scientific reports*, 6, p.23330.

Cheng, C. et al., 2015. An approach for determining and measuring network hierarchy applied to comparing the phosphorylome and the regulome. *Genome biology*, 16, p.63.

Chen, H. et al., 2016. Age-associated de-repression of retrotransposons in the Drosophila fat body, its potential cause and consequence. *Aging cell*, 15(3), pp.542–552.

Chen, J.-F. et al., 2006. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature genetics*, 38(2), pp.228–233.

Chen, Q. et al., 2006. Ump1 extends yeast lifespan and enhances viability during oxidative stress: central role for the proteasome? *Free radical biology & medicine*, 40(1), pp.120–126.

Cho, K.-O. et al., 2011. Wolbachia Bacteria Reside in Host Golgi-Related Vesicles Whose Position Is Regulated by Polarity Proteins O. Neyrolles, ed. *PloS one*, 6(7), p.e22703.

Chondrogianni, N. et al., 2000. Fibroblast cultures from healthy centenarians have an active proteasome. *Experimental gerontology*, 35(6-7), pp.721–728.

Çiçek, I.Ö. et al., 2016. Hedgehog Signaling Strength Is Orchestrated by the mir-310 Cluster of MicroRNAs in Response to Diet. *Genetics*, 202(3), pp.1167–1183.

Cohen, E. & Dillin, A., 2008. The insulin paradox: aging, proteotoxicity and neurodegeneration. *Nature reviews. Neuroscience*, 9(10), pp.759–767.

Copeland, J.M. et al., 2009. Extension of Drosophila life span by RNAi of the mitochondrial respiratory chain. *Current biology: CB*, 19(19), pp.1591–1598.

Cribari-Neto, F. & Zeileis, A., 2010. Beta Regression in R. *Journal of Statistical Software, Articles*, 34(2), pp.1–24.

Dasgupta, A. et al., 2015. Reducing the Analytical Bottleneck for Domain Scientists: Lessons from a Climate Data Visualization Case Study. *Computing in science & engineering*, 18(1), pp.92–100.

Daum, B. et al., 2013. Age-dependent dissociation of ATP synthase dimers and loss of inner-membrane cristae in mitochondria. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), pp.15301–15306.

Dayon, L. et al., 2008. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Analytical chemistry*, 80(8), pp.2921–2931.

De Cecco, M. et al., 2013. Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging cell*, 12(2), pp.247–256.

Deeb, S.J. et al., 2012. Super-SILAC allows classification of diffuse large B-cell lymphoma subtypes by their protein expression profiles. *Molecular & cellular proteomics: MCP*, 11(5), pp.77–89.

Deming, W.E., 1943. Statistical adjustment of data. Available at: http://psycnet.apa.org/record/1944-00642-000 [Accessed March 13, 2018].

Demontis, F. & Perrimon, N., 2010. FOXO/4E-BP signaling in Drosophila muscles regulates organism-wide proteostasis during aging. *Cell*, 143(5), pp.813–825.

Dennis, S. et al., 2012. C. elegans germ cells show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related retrotransposon. *PLoS pathogens*, 8(3), p.e1002591.

Deshmukh, A.S. et al., 2015. Deep proteomics of mouse skeletal muscle enables quantitation of protein isoforms, metabolic pathways, and transcription factors. *Molecular & cellular proteomics: MCP*, 14(4), pp.841–853.

Dillin, A. et al., 2002. Rates of behavior and aging specified by mitochondrial function during development. *Science*, 298(5602), pp.2398–2401.

Diop, S.B. et al., 2015. PGC-1/Spargel Counteracts High-Fat-Diet-Induced Obesity and Cardiac Lipotoxicity Downstream of TOR and Brummer ATGL Lipase. *Cell reports*. Available at: http://dx.doi.org/10.1016/j.celrep.2015.02.022.

Di, Y., Schafer, D.W. & Cumbie, J.S., 2011. The NBP negative binomial model for assessing differential gene expression from RNA-Seq. *Statistical applications in genetics and molecular biology*. Available at: https://www.degruyter.com/view/j/sagmb.2011.10.issue-1/sagmb.2011.10.1.1637/sagmb.2011.10.1.1637.xml.

Dobin, A. et al., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* , 29(1), pp.15–21.

Doble, B.W. & Woodgett, J.R., 2003. GSK-3: tricks of the trade for a multi-tasking kinase. *Journal of cell*

science, 116(Pt 7), pp.1175–1186.

Dutta, D. et al., 2011. Self-renewal versus lineage commitment of embryonic stem cells: protein kinase C signaling shifts the balance. *Stem cells*, 29(4), pp.618–628.

Eiges, R. et al., 2001. Establishment of human embryonic stem cell-transfected clones carrying a marker for undifferentiated cells. *Current biology: CB*, 11(7), pp.514–518.

Elkon, R., Ugalde, A.P. & Agami, R., 2013. Alternative cleavage and polyadenylation: extent, regulation and function. *Nature reviews. Genetics*, 14(7), pp.496–506.

Elling, U. et al., 2011. Forward and reverse genetics through derivation of haploid mouse embryonic stem cells. *Cell stem cell*, 9(6), pp.563–574.

Ellis, B.C., Graham, L.D. & Molloy, P.L., 2014. CRNDE, a long non-coding RNA responsive to insulin/IGF signaling, regulates genes involved in central metabolism. *Biochimica et biophysica acta*, 1843(2), pp.372–386.

El Ouaamari, A. et al., 2008. miR-375 Targets 3'-Phosphoinositide–Dependent Protein Kinase-1 and Regulates Glucose-Induced Biological Responses in Pancreatic β-Cells. *Diabetes*, 57(10), pp.2708–2717.

Engin, F. et al., 2013. Restoration of the unfolded protein response in pancreatic β cells protects mice against type 1 diabetes. *Science translational medicine*, 5(211), p.211ra156.

Enright, A.J. et al., 2003. MicroRNA targets in Drosophila. *Genome biology*, 5(1), p.R1.

Ernst, P. et al., 2004. An Mll-dependent Hox program drives hematopoietic progenitor expansion. *Current biology: CB*, 14(22), pp.2063–2069.

Esko, J.D. & Selleck, S.B., 2002. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annual review of biochemistry*, 71, pp.435–471.

Essers, P. et al., 2016. Reduced insulin/insulin-like growth factor signaling decreases translation in Drosophila and mice. *Scientific reports*, 6, p.30290.

Evans, M.J. & Scarpulla, R.C., 1990. NRF-1: a trans-activator of nuclear-encoded respiratory genes in animal cells. *Genes & development*, 4(6), pp.1023–1034.

Ewald, C.Y. et al., 2015. Dauer-independent insulin/IGF-1-signalling implicates collagen remodelling in longevity. *Nature*, 519(7541), pp.97–101.

Facchinetti, V. et al., 2008. The mammalian target of rapamycin complex 2 controls folding and stability of Akt and protein kinase C. *The EMBO journal*, 27(14), pp.1932–1943.

Faherty, C.S. & Maurelli, A.T., 2008. Staying alive: bacterial inhibition of apoptosis during infection. *Trends in microbiology*, 16(4), pp.173–180.

Farquhar, M.G. & Palade, G.E., 1963. Junctional complexes in various epithelia. *The Journal of cell biology*, 17, pp.375–412.

Ferguson, M. et al., 2005. Age-associated decline in mitochondrial respiration and electron transport in Drosophila melanogaster. *Biochemical Journal*, 390(Pt 2), pp.501–511.

Fernyhough, P. et al., 1989. Stabilization of tubulin mRNAs by insulin and insulin-like growth factor I during neurite formation. *Brain research. Molecular brain research*, 6(2-3), pp.109–120.

Fidalgo, M. et al., 2016. Zfp281 Coordinates Opposing Functions of Tet1 and Tet2 in Pluripotent States. *Cell stem cell*, 19(3), pp.355–369.

Firestein, R. et al., 2008. CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. *Nature*, 455(7212), pp.547–551.

Fischer, K. et al., 2014. High Pressure Freezing/Freeze Substitution Fixation Improves the Ultrastructural Assessment of Wolbachia Endosymbiont – Filarial Nematode Host Interaction. *PloS one*, 9(1), p.e86383.

Fisher, R.A., 1915. Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population. *Biometrika*, 10(4), pp.507–521.

Fletcher, L.M. et al., 2000. Role for the microtubule cytoskeleton in GLUT4 vesicle trafficking and in the regulation of insulin-stimulated glucose uptake. *Biochemical Journal*, 352 Pt 2, pp.267–276.

Fontana, L., 2007. Aging, Adiposity, and Calorie Restriction. *JAMA: the journal of the American Medical Association*, 297(9), p.986.

Fontana, L., Partridge, L. & Longo, V.D., 2010. Extending healthy life span--from yeast to humans. *Science*, 328(5976), pp.321–326.

Fridley, B.L., Jenkins, G.D. & Biernacka, J.M., 2010. Self-contained gene-set analysis of expression data: an evaluation of existing and novel methods. *PloS one*, 5(9). Available at: http://dx.doi.org/10.1371/journal.pone.0012693.

Gan, Z. et al., 2006. RNA Editing by ADAR2 Is Metabolically Regulated in Pancreatic Islets and β-Cells. *The Journal of biological chemistry*, 281(44), pp.33386–33394.

Garmhausen, M., 2016. *A Systems Perspective on the Exit from Pluripotency*. text.thesis.doctoral. Universität zu Köln. Available at: http://kups.ub.uni-koeln.de/7636/ [Accessed January 14, 2018].

Géminard, C., Rulifson, E.J. & Léopold, P., 2009. Remote control of insulin secretion by fat cells in Drosophila. *Cell metabolism*, 10(3), pp.199–207.

Gene Ontology Consortium, 2001. Creating the gene ontology resource: design and implementation. *Genome research*, 11(8), pp.1425–1433.

Gershman, B. et al., 2007. High-resolution dynamics of the transcriptional response to nutrition in Drosophila: a key role for dFOXO. *Physiological genomics*, 29(1), pp.24–34.

G Hendrickson, D. et al., 2016. Widespread RNA binding by chromatin-associated proteins. *Genome biology*, 17, p.28.

Giannakou, M.E. et al., 2004. Long-lived Drosophila with overexpressed dFOXO in adult fat body. *Science*,

305(5682), p.361.

Gitter, A. et al., 2009. Backup in gene regulatory networks explains differences between binding and knockout results. *Molecular systems biology*, 5, p.276.

Glozak, M.A. et al., 2005. Acetylation and deacetylation of non-histone proteins. *Gene*, 363, pp.15–23.

Goalstone, M.L. & Draznin, B., 1999. Effect of insulin on farnesyltransferase gene transcription and mRNA stability. *Biochemical and biophysical research communications*, 254(1), pp.243–247.

Golub, T.R. et al., 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, 286(5439), pp.531–537.

Gordiyenko, Y. et al., 2014. eIF2B is a decameric guanine nucleotide exchange factor with a γ2ε2 tetrameric core. *Nature communications*, 5, p.3902.

Gramates, L.S. et al., 2017. FlyBase at 25: looking to the future. *Nucleic acids research*, 45(D1), pp.D663–D671.

Grand, E.K. et al., 2004. Targeting FGFR3 in multiple myeloma: inhibition of t(4;14)-positive cells by SU5402 and PD173074. *Leukemia*, 18(5), pp.962–966.

Graveley, B.R. et al., 2011. The developmental transcriptome of Drosophila melanogaster. *Nature*, 471(7339), pp.473–479.

Grönke, S. et al., 2005. Brummer lipase is an evolutionary conserved fat storage regulator in Drosophila. *Cell metabolism*, 1(5), pp.323–330.

Grönke, S. et al., 2003. Control of fat storage by a Drosophila PAT domain protein. *Current biology: CB*, 13(7), pp.603–606.

Grönke, S. et al., 2007. Dual lipolytic control of body fat storage and mobilization in Drosophila. *PLoS biology*, 5(6), p.e137.

Grönke, S. et al., 2010. Molecular Evolution and Functional Characterization of Drosophila Insulin-Like Peptides E. Rulifson, ed. *PLoS genetics*, 6(2), p.e1000857.

Guhaniyogi, J. & Brewer, G., 2001. Regulation of mRNA stability in mammalian cells. *Gene*, 265(1-2), pp.11–23.

Gundelfinger, E.D. et al., 1984. Structure and evolution of the 7SL RNA component of the signal recognition particle. *The EMBO journal*, 3(10), pp.2325–2332.

Guo, S., 2014. Insulin signaling, resistance, and the metabolic syndrome: insights from mouse models into disease mechanisms. *The Journal of endocrinology*, 220(2), pp.T1–T23.

Hagen, T.M. et al., 1997. Mitochondrial decay in hepatocytes from old rats: membrane potential declines, heterogeneity and oxidants increase. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7), pp.3064–3069.

Halaschek-Wiener, J. et al., 2005. Analysis of long-lived C. elegans daf-2 mutants using serial analysis of

gene expression. *Genome research*, 15(5), pp.603–615.

Ha, M. & Kim, V.N., 2014. Regulation of microRNA biogenesis. *Nature reviews. Molecular cell biology*, 15(8), pp.509–524.

Hansen, M. et al., 2007. Lifespan extension by conditions that inhibit translation in Caenorhabditis elegans. *Aging cell*, 6(1), pp.95–110.

Hara, K. et al., 1998. Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism. *The Journal of biological chemistry*, 273(23), pp.14484–14494.

Harding, H.P. et al., 2000. Perk is essential for translational regulation and cell survival during the unfolded protein response. *Molecular cell*, 5(5), pp.897–904.

Harman, D., 1956. Aging: a theory based on free radical and radiation chemistry. *Journal of gerontology*, 11(3), pp.298–300.

Harper, M.E. et al., 1998. Age-related increase in mitochondrial proton leak and decrease in ATP turnover reactions in mouse hepatocytes. *The American journal of physiology*, 275(2 Pt 1), pp.E197–206.

Haruta, T. et al., 2000. A rapamycin-sensitive pathway down-regulates insulin signaling via phosphorylation and proteasomal degradation of insulin receptor substrate-1. *Molecular endocrinology* , 14(6), pp.783–794.

Hastie, T. et al., 2001. impute: Imputation for microarray data. *Bioinformatics* , 17(6), pp.520–525.

Hebert, S.L. et al., 2015. Mitochondrial Aging and Physical Decline: Insights From Three Generations of Women. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 70(11), pp.1409–1417.

van Heesch, S. et al., 2014. Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. *Genome biology*, 15(1), p.R6.

Heintz, C. et al., 2016. Splicing factor 1 modulates dietary restriction and TORC1 pathway longevity in C. elegans. *Nature*, 541(7635), pp.102–106.

Hietakangas, V. & Cohen, S.M., 2007. Re-evaluating AKT regulation: role of TOR complex 2 in tissue growth. *Genes & development*, 21(6), pp.632–637.

Hochstrasser, M. & Varshavsky, A., 1990. In vivo degradation of a transcriptional regulator: the yeast alpha 2 repressor. *Cell*, 61(4), pp.697–708.

Hoffmann, J.A., 2003. The immune response of Drosophila. *Nature*, 426(6962), pp.33–38.

Hong, Y.H. et al., 2003. AMP-activated Protein Kinase Regulates HNF4α Transcriptional Activity by Inhibiting Dimer Formation and Decreasing Protein Stability. *The Journal of biological chemistry*, 278(30), pp.27495–27501.

Hosler, B.A. et al., 1989. Expression of REX-1, a gene containing zinc finger motifs, is rapidly reduced by retinoic acid in F9 teratocarcinoma cells. *Molecular and cellular biology*, 9(12), pp.5623–5629.

Hsu, P.P. et al., 2011. The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science*, 332(6035), pp.1317–1322.

Huang, D.W. et al., 2007. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome biology*, 8(9), p.R183.

Hughes, M.E. et al., 2012. Deep sequencing the circadian and diurnal transcriptome of Drosophila brain. *Genome research*, 22(7), pp.1266–1281.

Huynh-Thu, V.A. et al., 2010. Inferring regulatory networks from expression data using tree-based methods. *PloS one*, 5(9). Available at: http://dx.doi.org/10.1371/journal.pone.0012776.

Ikeya, T. et al., 2009. The endosymbiont Wolbachia increases insulin/IGF-like signalling in Drosophila. *Proceedings. Biological sciences / The Royal Society*, 276(1674), pp.3799–3807.

Ingraham, T., 2017. So long static – we now support interactive Plotly figures in our articles! - F1000 Blogs. *F1000 Blogs*. Available at: https://blog.f1000.com/2017/07/19/so-long-static-we-now-support-interactive-ploty-figures-in-our-articles/ [Accessed March 10, 2018].

Isken, O. & Maquat, L.E., 2008. The multiple lives of NMD factors: balancing roles in gene and genome regulation. *Nature reviews. Genetics*, 9(9), pp.699–712.

Jeyaprakash, A. & Hoy, M.A., 2000. Long PCR improves Wolbachia DNA amplification: wsp sequences found in 76% of sixty-three arthropod species. *Insect molecular biology*, 9(4), pp.393–405.

Jiang, N. et al., 2013. Dietary and genetic effects on age-related loss of gene silencing reveal epigenetic plasticity of chromatin repression during aging. *Aging*, 5(11), pp.813–824.

Johnson, A.E. & van Waes, M.A., 1999. The Translocon: A Dynamic Gateway at the ER Membrane. *Annual review of cell and developmental biology*, 15(1), pp.799–842.

Jones, B.C. et al., 2016. A somatic piRNA pathway in the Drosophila fat body ensures metabolic homeostasis and normal lifespan. *Nature communications*, 7, p.13856.

Jordan, S.D. et al., 2011. Obesity-induced overexpression of miRNA-143 inhibits insulin-stimulated AKT activation and impairs glucose metabolism. *Nature cell biology*, 13(4), pp.434–446.

Jünger, M.A. et al., 2003. The Drosophila forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling. *Journal of biology*, 2(3), p.20.

Jüschke, C. et al., 2013. Transcriptome and proteome quantification of a tumor model provides novel insights into post-transcriptional gene regulation. *Genome biology*, 14(11), p.r133.

Kaeberlein, M. et al., 2005. Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. *Science*, 310(5751), pp.1193–1196.

Kaletsky, R. et al., 2016. The C. elegans adult neuronal IIS/FOXO transcriptome reveals adult phenotype regulators. *Nature*, 529(7584), pp.92–96.

Kalkan, T. & Smith, A., 2014. Mapping the route from naive pluripotency to lineage specification.

Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369(1657), p.20130540.

Kanehisa, M. et al., 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic acids research*, 40(Database issue), pp.D109–14.

Kang, H. et al., 2010. Relative actin nucleation promotion efficiency by WASP and WAVE proteins in endothelial cells. *Biochemical and biophysical research communications*, 400(4), pp.661–666.

Kang, S. et al., 2003. Insulin can block apoptosis by decreasing oxidative stress via phosphatidylinositol 3-kinase- and extracellular signal-regulated protein kinase-dependent signaling pathways in HepG2 cells. *European journal of endocrinology / European Federation of Endocrine Societies*, 148(1), pp.147–155.

Karolina, D.S. et al., 2011. MicroRNA 144 impairs insulin signaling by inhibiting the expression of insulin receptor substrate 1 in type 2 diabetes mellitus. *PloS one*, 6(8), p.e22839.

Karouzou, M.V. et al., 2007. Drosophila cuticular proteins with the R&R Consensus: annotation and classification with a new tool for discriminating RR-1 and RR-2 sequences. *Insect biochemistry and molecular biology*, 37(8), pp.754–760.

Karpievitch, Y.V., Dabney, A.R. & Smith, R.D., 2012. Normalization and missing value imputation for label-free LC-MS analysis. *BMC bioinformatics*, 13 Suppl 16, p.S5.

Katti, P. et al., 2017. Overexpression of miRNA-9 Generates Muscle Hypercontraction Through Translational Repression of Troponin-T in Drosophila melanogaster Indirect Flight Muscles. *G3* , 7(10), pp.3521–3531.

Keenan, R.J. et al., 2003. The Signal Recognition Particle. Available at: http://dx.doi.org/10.1146/annurev.biochem.70.1.755 [Accessed November 17, 2017].

Kenyon, C. et al., 1993. A C. elegans mutant that lives twice as long as wild type. *Nature*, 366(6454), pp.461–464.

Kenyon, C., 2011. The first long-lived mutants: discovery of the insulin/IGF-1 pathway for ageing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 366(1561), pp.9–16.

Kenyon, C.J., 2010. The genetics of ageing. *Nature*, 464(7288), pp.504–512.

Kikis, E.A., Gidalevitz, T. & Morimoto, R.I., 2010. Protein homeostasis in models of aging and age-related conformational disease. *Advances in experimental medicine and biology*, 694, pp.138–159.

Kim, D.-K. et al., 2014. Lefty1 and lefty2 control the balance between self-renewal and pluripotent differentiation of mouse embryonic stem cells. *Stem cells and development*, 23(5), pp.457–466.

Kim, M.-S. et al., 2014. A draft map of the human proteome. *Nature*, 509(7502), pp.575–581.

Kino, T. et al., 2010. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Science signaling*, 3(107), p.ra8.

Kirstein-Miles, J. et al., 2013. The nascent polypeptide-associated complex is a key regulator of proteostasis. *The EMBO journal*, 32(10), pp.1451–1468.

Kiss, T., 2002. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, 109(2), pp.145–148.

Komar, A.A. & Hatzoglou, M., 2011. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. *Cell cycle* , 10(2), pp.229–240.

Kooijman, R., 2006. Regulation of apoptosis by insulin-like growth factor (IGF)-I. *Cytokine & growth factor reviews*, 17(4), pp.305–323.

Kruegel, U. et al., 2011. Elevated proteasome capacity extends replicative lifespan in Saccharomyces cerevisiae. *PLoS genetics*, 7(9), p.e1002253.

Kubota, H., Hynes, G. & Willison, K., 1995. The Chaperonin Containing t-complex polypeptide 1 (TCP-1). *European journal of biochemistry / FEBS*, 230(1), pp.3–16.

Kunath, T. et al., 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* , 134(16), pp.2895–2902.

Kuraishi, T. et al., 2011. Genetic evidence for a protective role of the peritrophic matrix against intestinal bacterial infection in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 108(38), pp.15966–15971.

Lackner, D.H. et al., 2012. Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. *Genome biology*, 13(4), p.R25.

Lakkaraju, A.K.K. et al., 2008. SRP keeps polypeptides translocation-competent by slowing translation to match limiting ER-targeting sites. *Cell*, 133(3), pp.440–451.

Laprise, P. et al., 2006. The FERM protein Yurt is a negative regulatory component of the Crumbs complex that controls epithelial polarity and apical membrane size. *Developmental cell*, 11(3), pp.363–374.

Law, C.W. et al., 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome biology*, 15(2), p.R29.

Lecker, S.H., Goldberg, A.L. & Mitch, W.E., 2006. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *Journal of the American Society of Nephrology: JASN*, 17(7), pp.1807–1819.

Leeb, M. et al., 2014. Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. *Cell stem cell*, 14(3), pp.385–393.

Leeb, M. & Wutz, A., 2011. Derivation of haploid embryonic stem cells from mouse embryos. *Nature*, 479(7371), pp.131–134.

Lee, C. & Longo, V., 2016. Dietary restriction with and without caloric restriction for healthy aging. *F1000Research*, 5. Available at: http://dx.doi.org/10.12688/f1000research.7136.1.

Leek, J.T. et al., 2010. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature reviews. Genetics*, 11(10), pp.733–739.

Lee, R.C., Feinbaum, R.L. & Ambros, V., 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75(5), pp.843–854.

Lee, S.H., 2015. Intestinal permeability regulation by tight junction: implication on inflammatory bowel diseases. *Intestinal research*, 13(1), pp.11–18.

Lehane, M.J., 1997. Peritrophic matrix structure and function. *Annual review of entomology*, 42, pp.525–550.

Leiser, S.F. et al., 2015. Cell nonautonomous activation of flavin-containing monooxygenase promotes longevity and health span. *Science*, 350(6266), pp.1375–1378.

Levy, J.R. & Hug, V., 1992. Regulation of insulin receptor gene expression. Cell cycle-mediated effects on insulin receptor mRNA stability. *The Journal of biological chemistry*, 267(35), pp.25289–25295.

Lewis, S.E. et al., 1985. The effects of aging and chronic dietary restriction on whole body growth and protein turnover in the rat. *Experimental gerontology*, 20(5), pp.253–263.

Li, B. & Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*, 12, p.323.

de Lichtenberg, U. et al., 2005. Dynamic complex formation during the yeast cell cycle. *Science*, 307(5710), pp.724–727.

Li, G.-W. et al., 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell*, 157(3), pp.624–635.

Li, J.J., Bickel, P.J. & Biggin, M.D., 2014. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ*, 2, p.e270.

Limmer, S. et al., 2014. The Drosophila blood-brain barrier: development and function of a glial endothelium. *Frontiers in neuroscience*, 8, p.365.

Lin, X., 2004. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development* , 131(24), pp.6009–6021.

Liu, K.-C. et al., 2013. Ovarian cancer stem-like cells show induced translineage-differentiation capacity and are suppressed by alkaline phosphatase inhibitor. *Oncotarget*, 4(12), pp.2366–2382.

Liu, Y. & Aebersold, R., 2016. The interdependence of transcript and protein abundance: new data--new complexities. *Molecular systems biology*, 12(1), p.856.

Liu, Y., Beyer, A. & Aebersold, R., 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*, 165(3), pp.535–550.

Liu, Y., Zhou, J. & White, K.P., 2014. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics* , 30(3), pp.301–304.

Li, V.S.W. et al., 2012. Wnt signaling through inhibition of β-catenin degradation in an intact Axin1 complex. *Cell*, 149(6), pp.1245–1256.

Li, X. et al., 2007. Akt/PKB regulates hepatic metabolism by directly inhibiting PGC-1alpha transcription coactivator. *Nature*, 447(7147), pp.1012–1016.

Llimargas, M. et al., 2004. Lachesin is a component of a septate junction-based mechanism that controls tube size and epithelial integrity in the Drosophila tracheal system. *Development* , 131(1), pp.181–190.

Long, Y. et al., 2017. How do lncRNAs regulate transcription? *Science advances*, 3(9), p.eaao2110.

López-Otín, C. et al., 2013. The Hallmarks of Aging. *Cell*, 153(6), pp.1194–1217.

Maaten, L. van der & Hinton, G., 2008. Visualizing Data using t-SNE. *Journal of machine learning research: JMLR*, 9(Nov), pp.2579–2605.

MacInnes, A.W., 2016. The role of the ribosome in the regulation of longevity and lifespan extension. *Wiley interdisciplinary reviews. RNA*, 7(2), pp.198–212.

Macneil, L.T. & Walhout, A.J.M., 2011. Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. *Genome research*, 21(5), pp.645–657.

Madigan, J.P., Chotkowski, H.L. & Glaser, R.L., 2002. DNA double-strand break-induced phosphorylation of Drosophila histone variant H2Av helps prevent radiation-induced apoptosis. *Nucleic acids research*, 30(17), pp.3698–3705.

Mah, I.K. et al., 2015. Atypical PKC-iota Controls Stem Cell Expansion via Regulation of the Notch Pathway. *Stem cell reports*, 5(5), pp.866–880.

Manning, B.D. et al., 2002. Identification of the tuberous sclerosis complex-2 tumor suppressor gene product tuberin as a target of the phosphoinositide 3-kinase/akt pathway. *Molecular cell*, 10(1), pp.151–162.

Mann, M. et al., 2013. The coming age of complete, accurate, and ubiquitous proteomes. *Molecular cell*, 49(4), pp.583–590.

Mansfield, E.R. & Helms, B.P., 1982. Detecting Multicollinearity. *The American statistician*, 36(3a), pp.158–160.

Marhfour, I. et al., 2012. Expression of endoplasmic reticulum stress markers in the islets of patients with type 1 diabetes. *Diabetologia*, 55(9), pp.2417–2420.

Marques, J.T. et al., 2010. Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila. *Nature structural & molecular biology*, 17(1), pp.24–30.

Marr, M.T., 2nd et al., 2007. IRES-mediated functional coupling of transcription and translation amplifies insulin receptor feedback. *Genes & development*, 21(2), pp.175–183.

Martello, G. et al., 2012. Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell

self-renewal. *Cell stem cell*, 11(4), pp.491–504.

Masui, S. et al., 2008. Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. *BMC developmental biology*, 8, p.45.

Matsuda, T. et al., 1999. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *The EMBO journal*, 18(15), pp.4261–4269.

Maxwell, P.H., Burhans, W.C. & Curcio, M.J., 2011. Retrotransposition is associated with genome instability during chronological aging. *Proceedings of the National Academy of Sciences of the United States of America*, 108(51), pp.20376–20381.

Mazin, P. et al., 2013. Widespread splicing changes in human brain development and aging. *Molecular systems biology*, 9, p.633.

McCarthy, D.J. & Smyth, G.K., 2009. Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* , 25(6), pp.765–771.

McCormick, P.J. et al., 2003. Cotranslational Protein Integration into the ER Membrane Is Mediated by the Binding of Nascent Chains to Translocon Proteins. *Molecular cell*, 12(2), pp.329–341.

McCracken, K.W. et al., 2011. Generating human intestinal tissue from pluripotent stem cells in vitro. *Nature protocols*, 6(12), pp.1920–1928.

McElwee, J.J. et al., 2007. Evolutionary conservation of regulated longevity assurance mechanisms. *Genome biology*, 8(7), p.R132.

McManus, C.J. et al., 2014. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. *Genome research*, 24(3), pp.422–430.

Meinshausen, N. & Bühlmann, P., 2010. Stability selection. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 72(4), pp.417–473.

Metaxakis, A. et al., 2014. Lowered insulin signalling ameliorates age-related sleep fragmentation in Drosophila. *PLoS biology*, 12(4), p.e1001824.

Min, K.T. & Benzer, S., 1997. Wolbachia, normally a symbiont of Drosophila, can be virulent, causing degeneration and early death. *Proceedings of the National Academy of Sciences of the United States of America*, 94(20), pp.10792–10796.

Miron, M., Lasko, P. & Sonenberg, N., 2003. Signaling from Akt to FRAP/TOR targets both 4E-BP and S6K in Drosophila melanogaster. *Molecular and cellular biology*, 23(24), pp.9117–9126.

Miura, S. et al., 2002. Functional conservation for lipid storage droplet association among Perilipin, ADRP, and TIP47 (PAT)-related proteins in mammals, Drosophila, and Dictyostelium. *The Journal of biological chemistry*, 277(35), pp.32253–32257.

Mizuno, H., Tobita, M. & Uysal, A.C., 2012. Concise review: adipose-derived stem cells as a novel tool for future regenerative medicine. *Stem cells*, 30(5), pp.804–810.

Mohr, C. & Hartmann, B., 2014. Alternative splicing in *Drosophila* neuronal development. *Journal of*

neurogenetics, 28(3-4), pp.199–215.

Moll, P. et al., 2014. QuantSeq 3′ mRNA sequencing for RNA quantification. *Nature methods*, 11. Available at: http://dx.doi.org/10.1038/nmeth.f.376.

Mootha, V.K. et al., 2003. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics*, 34(3), pp.267–273.

Moreau, J.-F. et al., 1988. Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells. *Nature*, 336, p.690.

Mortazavi, A. et al., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7), pp.621–628.

Mukherjee, C. et al., 2012. Identification of cytoplasmic capping targets reveals a role for cap homeostasis in translation and mRNA stability. *Cell reports*, 2(3), pp.674–684.

Mukherjee, S. & Duttaroy, A., 2013. Spargel/dPGC-1 is a new downstream effector in the insulin-TOR signaling pathway in Drosophila. *Genetics*, 195(2), pp.433–441.

Mulgrew, N.M. et al., 2014. c-Met inhibition in a HOXA9/Meis1 model of CN-AML. *Developmental dynamics: an official publication of the American Association of Anatomists*, 243(1), pp.172–181.

Müller, F.-J. et al., 2008. Regulatory networks define phenotypic classes of human stem cell lines. *Nature*, 455(7211), pp.401–405.

Murali, T. et al., 2011. DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. *Nucleic acids research*, 39(Database issue), pp.D736–43.

Murata, S., Yashiroda, H. & Tanaka, K., 2009. Molecular mechanisms of proteasome assembly. *Nature reviews. Molecular cell biology*, 10(2), pp.104–115.

Murphy, C.T. et al., 2003. Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature*, 424(6946), pp.277–283.

Muzzey, D., Sherlock, G. & Weissman, J.S., 2014. Extensive and coordinated control of allele-specific expression by both transcription and translation in Candida albicans. *Genome research*, 24(6), pp.963–973.

Nagaraj, N. et al., 2012. System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Molecular & cellular proteomics: MCP*, 11(3), p.M111.013722.

Nandi, A. et al., 2004. Mouse models of insulin resistance. *Physiological reviews*, 84(2), pp.623–647.

Narayan, V. et al., 2016. Deep Proteome Analysis Identifies Age-Related Processes in C. elegans. *Cell Systems*, 3(2), pp.144–159.

Newton, I.L.G., Savytskyy, O. & Sheehan, K.B., 2015. Wolbachia utilize host actin for efficient maternal transmission in Drosophila melanogaster. *PLoS pathogens*, 11(4), p.e1004798.

Nichols, J. & Smith, A., 2009. Naive and primed pluripotent states. *Cell stem cell*, 4(6), pp.487–492.

Nichols, J. & Smith, A., 2012. Pluripotency in the embryo and in culture. *Cold Spring Harbor perspectives in biology*, 4(8), p.a008128.

Niwa, H. et al., 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes & development*, 12(13), pp.2048–2060.

Nolte-'t Hoen, E.N.M. et al., 2015. The role of microRNA in nutritional control. *Journal of internal medicine*, 278(2), pp.99–109.

O'Brien, R.M. et al., 1990. Identification of a sequence in the PEPCK gene that mediates a negative effect of insulin on transcription. *Science*, 249(4968), pp.533–537.

Ogawa, K. et al., 2006. Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. *Biochemical and biophysical research communications*, 343(1), pp.159–166.

Oh, S.W. et al., 2006. Identification of direct DAF-16 targets controlling longevity, metabolism and diapause by chromatin immunoprecipitation. *Nature genetics*, 38(2), pp.251–257.

Ohtsuka, S., Nakai-Futatsugi, Y. & Niwa, H., 2015. LIF signal in mouse embryonic stem cells. *JAK-STAT*, 4(2), p.e1086520.

Ori, A. et al., 2015. Integrated Transcriptome and Proteome Analyses Reveal Organ-Specific Proteome Deterioration in Old Rats. *Cell Systems*, 1(3), pp.224–237.

Pan, K.Z. et al., 2007. Inhibition of mRNA translation extends lifespan in Caenorhabditis elegans. *Aging cell*, 6(1), pp.111–119.

Park, S. et al., 2011. Structural defects in the regulatory particle-core particle interface of the proteasome induce a novel proteasome stress response. *The Journal of biological chemistry*, 286(42), pp.36652–36666.

Parodi, A.J., 2000. Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *Biochemical Journal*, 348 Pt 1, pp.1–13.

Partridge, L. et al., 2011. Ageing in Drosophila: The role of the insulin/Igf and TOR signalling network. *Experimental gerontology*, 46(5), pp.376–381.

Pathare, G.R. et al., 2012. The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together. *Proceedings of the National Academy of Sciences of the United States of America*, 109(1), pp.149–154.

Pérez, V.I. et al., 2009. Protein stability and resistance to oxidative stress are determinants of longevity in the longest-living rodent, the naked mole-rat. *Proceedings of the National Academy of Sciences of the United States of America*, 106(9), pp.3059–3064.

Perkel, J.M., 2018. Data visualization tools drive interactivity and reproducibility in online publishing. *Nature*, 554(7690), pp.133–134.

Pietri, J.E., DeBruhl, H. & Sullivan, W., 2016. The rich somatic life of Wolbachia. *MicrobiologyOpen*, 5(6),

pp.923–936.

Pletcher, S.D. et al., 2002. Genome-wide transcript profiles in aging and calorically restricted Drosophila melanogaster. *Current biology: CB*, 12(9), pp.712–723.

Pollard, K.S., Dudoit, S. & van der Laan, M.J., 2005. Multiple Testing Procedures: the multtest Package and Applications to Genomics. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Statistics for Biology and Health. Springer, New York, NY, pp. 249–271.

Porras, A. et al., 2003. Long-term treatment with insulin induces apoptosis in brown adipocytes: role of oxidative stress. *Endocrinology*, 144(12), pp.5390–5401.

Poy, M.N. et al., 2004. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature*, 432(7014), pp.226–230.

Proud, C.G., 2002. Regulation of mammalian translation factors by nutrients. *European journal of biochemistry / FEBS*, 269(22), pp.5338–5349.

Proudfoot, N.J., 2011. Ending the message: poly(A) signals then and now. *Genes & development*, 25(17), pp.1770–1782.

Puigserver, P. & Spiegelman, B.M., 2003. Peroxisome Proliferator-Activated Receptor-γ Coactivator 1α (PGC-1α): Transcriptional Coactivator and Metabolic Regulator. *Endocrine reviews*, 24(1), pp.78–90.

Rapoport, T.A., 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature*, 450(7170), pp.663–669.

Rask-Madsen, C. & Kahn, C.R., 2012. Tissue-specific insulin signaling, metabolic syndrome, and cardiovascular disease. *Arteriosclerosis, thrombosis, and vascular biology*, 32(9), pp.2052–2059.

Rea, S.L., Ventura, N. & Johnson, T.E., 2007. Relationship between mitochondrial electron transport chain dysfunction, development, and life extension in Caenorhabditis elegans. *PLoS biology*, 5(10), p.e259.

Reczko, M. et al., 2012. Functional microRNA targets in protein coding sequences. *Bioinformatics* , 28(6), pp.771–776.

Reid, D.W. & Nicchitta, C.V., 2015. Diversity and selectivity in mRNA translation on the endoplasmic reticulum. *Nature reviews. Molecular cell biology*, 16(4), pp.221–231.

Reid, D.W. & Nicchitta, C.V., 2012. Primary role for endoplasmic reticulum-bound ribosomes in cellular translation identified by ribosome profiling. *The Journal of biological chemistry*, 287(8), pp.5518–5527.

Rera, M. et al., 2011. Modulation of longevity and tissue homeostasis by the Drosophila PGC-1 homolog. *Cell metabolism*, 14(5), pp.623–634.

Rera, M., Clark, R.I. & Walker, D.W., 2012. Intestinal barrier dysfunction links metabolic and inflammatory markers of aging to death in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 109(52), pp.21528–21533.

Rhodenizer, D. et al., 2008. Genetic and environmental factors impact age-related impairment of

negative geotaxis in Drosophila by altering age-dependent climbing speed. *Experimental gerontology*, 43(8), pp.739–748.

Rice, K.M., Turnbow, M.A. & Garner, C.W., 1993. Insulin stimulates the degradation of IRS-1 in 3T3-L1 adipocytes. *Biochemical and biophysical research communications*, 190(3), pp.961–967.

Rimkus, T.K. et al., 2016. Targeting the Sonic Hedgehog Signaling Pathway: Review of Smoothened and GLI Inhibitors. *Cancers*, 8(2). Available at: http://dx.doi.org/10.3390/cancers8020022.

Ristow, M. & Schmeisser, S., 2011. Extending life span by increasing oxidative stress. *Free radical biology & medicine*, 51(2), pp.327–336.

Rivals, I. et al., 2007. Enrichment or depletion of a GO category within a class of genes: which test? *Bioinformatics* , 23(4), pp.401–407.

Robinson, M.D. & Oshlack, A., 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology*, 11(3), p.R25.

Rogakou, E.P. et al., 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *The Journal of biological chemistry*, 273(10), pp.5858–5868.

Rogers, M.B., Hosler, B.A. & Gudas, L.J., 1991. Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development* , 113(3), pp.815–824.

van Rooij, E. et al., 2009. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Developmental cell*, 17(5), pp.662–673.

Ross, R.J., Weiner, M.M. & Lin, H., 2014. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature*, 505(7483), pp.353–359.

Rosvall, M. & Bergstrom, C.T., 2008. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(4), pp.1118–1123.

Ruby, J.G. et al., 2007. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. *Genome research*, 17(12), pp.1850–1864.

Rulifson, E.J., Kim, S.K. & Nusse, R., 2002. Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. *Science*, 296(5570), pp.1118–1120.

Sadoul, K. et al., 2008. Regulation of protein turnover by acetyltransferases and deacetylases. *Biochimie*, 90(2), pp.306–312.

Saltiel, A.R. & Ronald Kahn, C., 2001. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature*, 414(6865), pp.799–806.

Santos, L.E. & Ferreira, S.T., 2017. Crosstalk between endoplasmic reticulum stress and brain inflammation in Alzheimer's disease. *Neuropharmacology*. Available at: http://dx.doi.org/10.1016/j.neuropharm.2017.11.016.

Saraogi, I. & Shan, S.-O., 2011. Molecular Mechanism of Co-translational Protein Targeting by the Signal Recognition Particle. *Traffic* , 12(5), pp.535–542.

Sarot, E. et al., 2004. Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the Drosophila melanogaster flamenco gene. *Genetics*, 166(3), pp.1313–1321.

Saxena, A. & Carninci, P., 2011. Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 33(11), pp.830–839.

Scarpulla, R.C., 2008. Transcriptional paradigms in mammalian mitochondrial biogenesis and function. *Physiological reviews*, 88(2), pp.611–638.

Scheuner, D. et al., 2005. Control of mRNA translation preserves endoplasmic reticulum function in beta cells and maintains glucose homeostasis. *Nature medicine*, 11(7), pp.757–764.

Schmidt, A. et al., 1998. The TOR nutrient signalling pathway phosphorylates NPR1 and inhibits turnover of the tryptophan permease. *The EMBO journal*, 17(23), pp.6924–6931.

Schröder, M. & Kaufman, R.J., 2005. ER stress and the unfolded protein response. *Mutation research*, 569(1-2), pp.29–63.

Schwanhäusser, B. et al., 2011. Global quantification of mammalian gene expression control. *Nature*, 473(7347), pp.337–342.

Seifert, A. et al., 2015. Role of Hox genes in stem cell differentiation. *World journal of stem cells*, 7(3), pp.583–595.

Selman, C. et al., 2009. Ribosomal protein S6 kinase 1 signaling regulates mammalian life span. *Science*, 326(5949), pp.140–144.

Serbus, L.R. et al., 2015. The impact of host diet on Wolbachia titer in Drosophila. *PLoS pathogens*, 11(3), p.e1004777.

Shahbazian, D. et al., 2006. The mTOR/PI3K and MAPK pathways converge on eIF4B to control its phosphorylation and activity. *The EMBO journal*, 25(12), pp.2781–2791.

Shah, M. et al., 2016. A Transcript-Specific eIF3 Complex Mediates Global Translational Control of Energy Metabolism. *Cell reports*, 16(7), pp.1891–1902.

Shah, R.D. & Samworth, R.J., 2013. Variable selection with error control: another look at stability selection. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 75(1), pp.55–80.

Sheehan, K.B. et al., 2016. Identification and Characterization of a Candidate Wolbachia pipientis Type IV Effector That Interacts with the Actin Cytoskeleton. *mBio*, 7(4). Available at: http://dx.doi.org/10.1128/mBio.00622-16.

Shigenaga, M.K., Hagen, T.M. & Ames, B.N., 1994. Oxidative damage and mitochondrial decay in aging. *Proceedings of the National Academy of Sciences of the United States of America*, 91(23), pp.10771–10778.

Shi, Z. et al., 2017. Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. *Molecular cell*, 67(1), pp.71–83.e7.

Shy, B.R. et al., 2013. Regulation of Tcf7l1 DNA binding and protein stability as principal mechanisms of Wnt/β-catenin signaling. *Cell reports*, 4(1), pp.1–9.

Sicard, M. et al., A host as an ecosystem: Wolbachia coping with environmental constraints. Available at: http://dx.doi.org/10.1111/1462-2920.12573.

Singh Chawla, D., 2015. "Living figures" make their debut. *Nature*, 521(7550), p.112.

Slack, C. et al., 2011. dFOXO-independent effects of reduced insulin-like signaling in Drosophila. *Aging cell*, 10(5), pp.735–748.

Slack, C. et al., 2015. The Ras-Erk-ETS-Signaling Pathway Is a Drug Target for Longevity. *Cell*, 162(1), pp.72–83.

Smyth, G.K., 2005. Limma: linear models for microarray data. In Springer, pp. 397–420.

Smyth, G.K., 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology*, 3, p.Article3.

Sokol, N.S. & Ambros, V., 2005. Mesodermally expressed Drosophila microRNA-1 is regulated by Twist and is required in muscles during larval growth. *Genes & development*, 19(19), pp.2343–2354.

Sokolova, V. et al., 2015. Proteasome Activation is Mediated via a Functional Switch of the Rpt6 C-terminal Tail Following Chaperone-dependent Assembly. *Scientific reports*, 5, p.14909.

Sonenberg, N. & Hinnebusch, A.G., 2009. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*, 136(4), pp.731–745.

Sood, R. et al., 2000. Pancreatic eukaryotic initiation factor-2alpha kinase (PEK) homologues in humans, Drosophila melanogaster and Caenorhabditis elegans that mediate translational control in response to endoplasmic reticulum stress. *Biochemical Journal*, 346 Pt 2, pp.281–293.

Stadtman, E.R., 2001. Protein oxidation in aging and age-related diseases. *Annals of the New York Academy of Sciences*, 928, pp.22–38.

Starr, R. et al., 1997. A family of cytokine-inducible inhibitors of signalling. *Nature*, 387(6636), pp.917–921.

Sterner, D.E. & Berger, S.L., 2000. Acetylation of histones and transcription-related factors. *Microbiology and molecular biology reviews: MMBR*, 64(2), pp.435–459.

Stern, J.R., Coon, M.J. & del Campillo, A., 1956. Enzymes of fatty acid metabolism. III. Breakdown and synthesis of beta-keto fatty acids. *The Journal of biological chemistry*, 221(1), pp.1–14.

Stouffer, S.A. et al., 1949. The American soldier: Adjustment during army life. (Studies in social psychology in World War II). Available at: http://psycnet.apa.org/psycinfo/1950-00790-000 [Accessed November 27, 2017].

Stout, G.J. et al., 2013. Insulin/IGF-1-mediated longevity is marked by reduced protein metabolism. *Molecular systems biology*, 9, p.679.

Sugimura, R. et al., 2017. Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature*, 545(7655), pp.432–438.

Supek, F. et al., 2011. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PloS one*, 6(7), p.e21800.

Tabar, V. & Studer, L., 2014. Pluripotent stem cells in regenerative medicine: challenges and recent progress. *Nature reviews. Genetics*, 15(2), pp.82–92.

Tain, L.S. et al., 2017. A proteomic atlas of insulin signalling reveals tissue-specific mechanisms of longevity assurance. *Molecular systems biology*, 13(9), p.939.

Takahashi, K. & Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4), pp.663–676.

Tang, T. et al., 2010. Adverse interactions between micro-RNAs and target genes from different species. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29), pp.12935–12940.

Tavernarakis, N. & Driscoll, M., 2002. Caloric restriction and lifespan: a role for protein turnover? *Mechanisms of ageing and development*, 123(2-3), pp.215–229.

Tawo, R. et al., 2017. The Ubiquitin Ligase CHIP Integrates Proteostasis and Aging by Regulation of Insulin Receptor Turnover. *Cell*, 169(3), pp.470–482.e13.

Teixeira, L., Ferreira, A. & Ashburner, M., 2008. The bacterial symbiont Wolbachia induces resistance to RNA viral infections in Drosophila melanogaster. *PLoS biology*, 6(12), p.e2.

Teleman, A.A., 2009. Molecular mechanisms of metabolic regulation by insulin in Drosophila. *Biochemical Journal*, 425(1), pp.13–26.

Teleman, A.A. et al., 2008. Nutritional control of protein biosynthetic capacity by insulin via Myc in Drosophila. *Cell metabolism*, 7(1), pp.21–32.

The UniProt Consortium, 2017. UniProt: the universal protein knowledgebase. *Nucleic acids research*, 45(D1), pp.D158–D169.

Thompson, A. et al., 2003. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Analytical chemistry*, 75(8), pp.1895–1904.

Tian, G. et al., 2011. An asymmetric interface between the regulatory and core particles of the proteasome. *Nature structural & molecular biology*, 18(11), pp.1259–1267.

Tiefenböck, S.K. et al., 2010. The Drosophila PGC-1 homologue Spargel coordinates mitochondrial activity to insulin signalling. *The EMBO journal*, 29(1), pp.171–183.

Tiku, V. et al., 2016. Small nucleoli are a cellular hallmark of longevity. *Nature communications*, 8,

p.16083.

Tillmar, L. & Welsh, N., 2004. Glucose-induced binding of the polypyrimidine tract-binding protein (PTB) to the 3'-untranslated region of the insulin mRNA (ins-PRS) is inhibited by rapamycin. *Molecular and cellular biochemistry*, 260(1-2), pp.85–90.

Tillmar, L. & Welsh, N., 2002. Hypoxia may increase rat insulin mRNA levels by promoting binding of the polypyrimidine tract-binding protein (PTB) to the pyrimidine-rich insulin mRNA 3'-untranslated region. *Molecular medicine* , 8(5), pp.263–272.

Ting, L. et al., 2011. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nature methods*, 8(11), pp.937–940.

Tipney, H. & Hunter, L., 2010. An introduction to effective use of enrichment analysis software. *Human genomics*, 4(3), pp.202–206.

Toivonen, J.M. et al., 2007. No Influence of Indy on Lifespan in Drosophila after Correction for Genetic and Cytoplasmic Background Effects. *PLoS genetics*, 3(6), p.e95.

Tomko, R.J., Jr & Hochstrasser, M., 2013. Molecular architecture and assembly of the eukaryotic proteasome. *Annual review of biochemistry*, 82, pp.415–445.

Tonoki, A. et al., 2009. Genetic evidence linking age-dependent attenuation of the 26S proteasome with the aging process. *Molecular and cellular biology*, 29(4), pp.1095–1106.

Trajkovski, M. et al., 2011. MicroRNAs 103 and 107 regulate insulin sensitivity. *Nature*, 474(7353), pp.649–653.

Trombetta, E.S. & Parodi, A.J., 2003. Quality control and protein folding in the secretory pathway. *Annual review of cell and developmental biology*, 19, pp.649–676.

Trott, J. & Martinez Arias, A., 2013. Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency. *Biology open*, 2(10), pp.1049–1056.

Tsakiri, E.N. et al., 2013. Differential regulation of proteasome functionality in reproductive vs. somatic tissues of Drosophila during aging or oxidative stress. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 27(6), pp.2407–2420.

Tsurudome, K. et al., 2010. The Drosophila miR-310 cluster negatively regulates synaptic strength at the neuromuscular junction. *Neuron*, 68(5), pp.879–893.

Ulloa, L. & Tabibzadeh, S., 2001. Lefty Inhibits Receptor-regulated Smad Phosphorylation Induced by the Activated Transforming Growth Factor-β Receptor. *The Journal of biological chemistry*, 276(24), pp.21397–21404.

Ungvari, Z. et al., 2013. Testing predictions of the oxidative stress hypothesis of aging using a novel invertebrate model of longevity: the giant clam (Tridacna derasa). *The journals of gerontology. Series A, Biological sciences and medical sciences*, 68(4), pp.359–367.

Van Der Kelen, K. et al., 2009. Translational control of eukaryotic gene expression. *Critical reviews in*

biochemistry and molecular biology, 44(4), pp.143–168.

Van Meter, M. et al., 2014. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nature communications*, 5, p.5011.

Vanunu, O. et al., 2010. Associating genes and protein complexes with disease via network propagation. *PLoS computational biology*, 6(1), p.e1000641.

Veraksa, A., Bauer, A. & Artavanis-Tsakonas, S., 2005. Analyzing protein complexes in Drosophila with tandem affinity purification-mass spectrometry. *Developmental dynamics: an official publication of the American Association of Anatomists*, 232(3), pp.827–834.

Vilchez, D. et al., 2012. RPN-6 determines C. elegans longevity under proteotoxic stress conditions. *Nature*, 489(7415), pp.263–268.

Vilchez, D., Saez, I. & Dillin, A., 2014. The role of protein clearance mechanisms in organismal ageing and age-related diseases. *Nature communications*, 5, p.5659.

Virbasius, J.V., Virbasius, C.A. & Scarpulla, R.C., 1993. Identity of GABP with NRF-2, a multisubunit activator of cytochrome oxidase expression, reveals a cellular role for an ETS domain activator of viral promoters. *Genes & development*, 7(3), pp.380–392.

Voronin, D.A., Dudkina, N.V. & Kiseleva, E.V., 2004. A New Form of Symbiotic Bacteria Wolbachia Found in the Endoplasmic Reticulum of Early Embryos of Drosophila melanogaster. *Doklady biological sciences: proceedings of the Academy of Sciences of the USSR, Biological sciences sections / translated from Russian*, 396(1-6), pp.227–229.

Vosseller, K. et al., 2002. Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(8), pp.5313–5318.

Wagner, G.P., Kin, K. & Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in biosciences = Theorie in den Biowissenschaften*, 131(4), pp.281–285.

Walter, P. & Blobel, G., 1981. Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *The Journal of cell biology*, 91(2 Pt 1), pp.557–561.

Walter, P. & Johnson, A.E., 1994. Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annual review of cell biology*, 10, pp.87–119.

Walter, W., Sánchez-Cabo, F. & Ricote, M., 2015. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics* , 31(17), pp.2912–2914.

Wang, D. et al., 2014. Knockdown expression of eukaryotic initiation factor 5 C-terminal domain containing protein extends lifespan in Drosophila melanogaster. *Biochemical and biophysical research communications*, 446(2), pp.465–469.

Wang, Z. et al., 2012. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human

embryonic stem cells. *Cell stem cell*, 10(4), pp.440–454.

Wang, Z. et al., 2015. Evolution of gene regulation during transcription and translation. *Genome biology and evolution*, 7(4), pp.1155–1167.

Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), pp.236–244.

Welham, M.J. et al., 2007. Phosphoinositide 3-kinases and regulation of embryonic stem cell fate. *Biochemical Society transactions*, 35(Pt 2), pp.225–228.

Welte, M.A. et al., 2005. Regulation of lipid-droplet transport by the perilipin homolog LSD2. *Current biology: CB*, 15(14), pp.1266–1275.

Werner, J.A., 2014. *Identification and characterization of differentially expressed microRNAs in adult tissues of the long-lived Drosophila dilp2-3,5 mutant*. text.thesis.doctoral. Universität zu Köln. Available at: http://kups.ub.uni-koeln.de/5548/ [Accessed March 7, 2018].

Werren, J.H. & Windsor, D.M., 2000. Wolbachia infection frequencies in insects: evidence of a global equilibrium? *Proceedings. Biological sciences / The Royal Society*, 267(1450), pp.1277–1285.

Westlake, W.J., 1976. Symmetrical confidence intervals for bioequivalence trials. *Biometrics*, 32(4), pp.741–744.

Whitaker, R. et al., 2014. Dietary switch reveals fast coordinated gene expression changes in Drosophila melanogaster. *Aging*, 6(5), pp.355–368.

Whitlock, M.C., 2005. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *Journal of evolutionary biology*, 18(5), pp.1368–1373.

Whitney, E.M. et al., 2006. Transcriptional profiling of the cell cycle checkpoint gene krüppel-like factor 4 reveals a global inhibitory function in macromolecular biosynthesis. *Gene expression*, 13(2), pp.85–96.

Wilkinson, G.N. & Rogers, C.E., 1973. Symbolic Description of Factorial Models for Analysis of Variance. *Journal of the Royal Statistical Society. Series C, Applied statistics*, 22(3), pp.392–399.

Winchester, B., 2005. Lysosomal metabolism of glycoproteins. *Glycobiology*, 15(6), p.1R–15R.

Wiza, C., Nascimento, E.B.M. & Ouwens, D.M., 2012. Role of PRAS40 in Akt and mTOR signaling in health and disease. *American journal of physiology. Endocrinology and metabolism*, 302(12), pp.E1453–60.

Wolkow, C.A. et al., 2000. Regulation of C. elegans life-span by insulinlike signaling in the nervous system. *Science*, 290(5489), pp.147–150.

Won, K.A. et al., 1998. Maturation of human cyclin E requires the function of eukaryotic chaperonin CCT. *Molecular and cellular biology*, 18(12), pp.7584–7589.

Wood, J.G. et al., 2016. Chromatin-modifying genetic interventions suppress age-associated transposable element activation and extend life span in Drosophila. *Proceedings of the National Academy of*

Sciences of the United States of America, 113(40), pp.11277–11282.

Wood, J.G. et al., 2010. Chromatin remodeling in the aging genome of Drosophila. *Aging cell*, 9(6), pp.971–978.

Wu, Y. et al., 2013. CHIR99021 promotes self-renewal of mouse embryonic stem cells by modulation of protein-encoding gene and long intergenic non-coding RNA expression. *Experimental cell research*, 319(17), pp.2684–2699.

Yamamoto, R. & Tatar, M., 2011. Insulin receptor substrate chico acts with the transcription factor FOXO to extend Drosophila lifespan. *Aging cell*, 10(4), pp.729–732.

Yang, J. & Tower, J., 2009. Expression of hsp22 and hsp70 transgenes is partially predictive of drosophila survival under normal and stress conditions. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 64(8), pp.828–838.

Yang, L. et al., 2010. Deficiency in RNA editing enzyme ADAR2 impairs regulated exocytosis. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 24(10), pp.3720–3732.

Yang, Q. et al., 2006. TSC1/TSC2 and Rheb have different effects on TORC1 and TORC2 activity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(18), pp.6811–6816.

Yen, T.C. et al., 1989. Liver mitochondrial respiratory functions decline with age. *Biochemical and biophysical research communications*, 165(3), pp.944–1003.

Ye, P. et al., 2015. An mTORC1-Mdm2-Drosha axis for miRNA biogenesis in response to glucose- and amino acid-deprivation. *Molecular cell*, 57(4), pp.708–720.

Ying, Q.-L. et al., 2003. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nature biotechnology*, 21(2), pp.183–186.

Ying, Q.-L. et al., 2008. The ground state of embryonic stem cell self-renewal. *Nature*, 453(7194), pp.519–523.

Young, R.A., 2011. Control of the embryonic stem cell state. *Cell*, 144(6), pp.940–954.

Yu, T. et al., 2015. MicroRNA-193a-3p and -5p suppress the metastasis of human non-small-cell lung cancer by downregulating the ERBB4/PIK3R3/mTOR/S6K2 signaling pathway. *Oncogene*, 34(4), pp.413–423.

Zarse, K. et al., 2012. Impaired insulin/IGF1 signaling extends life span by promoting mitochondrial L-proline catabolism to induce a transient ROS signal. *Cell metabolism*, 15(4), pp.451–465.

Zhang, M. et al., 2013. Structure of succinyl-CoA:3-ketoacid CoA transferase from Drosophila melanogaster. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 69(Pt 10), pp.1089–1093.

Zhao, Y. et al., 2015. Ubl4A is required for insulin-induced Akt plasma membrane translocation through promotion of Arp2/3-dependent actin branching. *Proceedings of the National Academy of Sciences*

of the United States of America, 112(31), pp.9644–9649.

Zid, B.M. et al., 2009. 4E-BP extends lifespan upon dietary restriction by enhancing mitochondrial activity in Drosophila. *Cell*, 139(1), pp.149–160.

Zubarev, R.A., 2013. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics*, 13(5), pp.723–726.

Zug, R. & Hammerstein, P., 2015. Wolbachia and the insect immune system: what reactive oxygen species can tell us about the mechanisms of Wolbachia-host interactions. *Frontiers in microbiology*, 6, p.1201.

7 Erklärung zur Dissertation

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit—einschließlich Tabellen, Karten und Abbildungen—die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation nach keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie—abgesehen von unten angegebenen Teilpublikationen—noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Linda Partridge und Prof. Dr. Andreas Beyer betreut worden.

Teilpublikationen:

Tain, L.S. et al., 2017. A proteomic atlas of insulin signalling reveals tissue-specific mechanisms of longevity assurance. Molecular systems biology, 13(9), p.939.

Ort, Datum Unterschrift