KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Human Genetics

KU LEUVEN

DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

Regulatory mechanisms controlling insulin-like peptide production in the *Drosophila* insulin-producing cells

Mattias Winant

Promotor:	Prof. Dr. Patrick Callaerts
Co-promotor:	Prof. Dr. Korneel Hens
Jury members:	Prof. Dr. Julien Colombani
	Prof. Dr. John Creemers
	Prof. Dr. Georg Halder
	Prof. Dr. Christian Wegener

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Biomedical Sciences

13/11/2023

Acknowledgments

Deze thesis was niet mogelijk geweest zonder een aantal belangrijke personen en dus dient dit stukje om iedereen uitvoerig te bedanken.

Prof. Dr**. Patrick** Callaerts**,** de eerste keer les dat ik bij jou les volgde en je vol overtuiging en passie de fruitvlieg als modelorganisme uit de doeken deed, wist ik dat *Drosophila* een blijvend deel van mijn academische carrière zou uitmaken. Door de te grote vraag voor jouw labrotaties kon ik jammer genoeg geen rotatie doen in jouw lab. Toch mocht ik, zonder dat je goed en wel wist wat voor vlees je in de kuip had, beginnen aan een masterproef bij jou en toen werd het helemaal duidelijk, I'm a fly guy. Bedankt om in mijn academische kwaliteiten te geloven. Veel dank voor je eindeloos optimisme, na elke vergadering kwam ik beter gezind en positiever naar buiten dan toen ik binnen ging.

Prof. Dr. **Korneel** Hens, bedankt voor de vele data die je vrij aan ons beschikbaar hebt gesteld. Zonder jouw input was dit doctoraat niet hetzelfde geweest.

Many thanks to my internal jury members, Prof. Dr. **Creemers** and Prof. Dr. **Halder** for providing guidance and advice in the various stages of my PhD. Prof. Dr. **Colombani** and Prof. Dr. **Wegener**, I really appreciate the time and effort that you have put into travelling to attend my defence and for the critical assessment of the manuscript.

Dr. **Kurt** Buhler**,** I cannot emphasize enough how important you were in making me, for better or for worse, into the scientist that I am now (mostly for better!). Your tenacity in overcoming setbacks and accomplishing your goals truly inspired me. Your ability to craft a story (scientific or otherwise) and to clearly visualize it had a big impact on me. The data as visualized in this thesis is enormously influenced by you. Also, many thanks for being such a great cat sitter!

Daarnaast wil ik ook mijn collega's van het LBDG bedanken. Dr. **Veerle** Vulsteke, de lab mama, zonder jou had ik ongetwijfeld de hele tijd zonder reagentia gezeten en had ik nooit geweten hoe de pH meter werkt. Veel dank voor alles wat je doet om het labo vlot te laten draaien. Dr. **Sofie** De Groef, je hebt je rol als cynische postdoc voorbeeldig volbracht. Nee, jouw advies, wetenschappelijk én culinair¸waren enorm belangrijk voor dit doctoraat. Bij jou kon ik altijd terecht om hersenspinsels en te ver gedreven hypotheses af te toetsen. Jouw kennis over de beta cel kwam steeds goed van pas. **Cissy**, ik zal nooit vergeten dat ik je de eerste keer in het Engels aansprak en dat je zei "In het Nederlands mag ook hoor" (Oeps, er werd mij verteld dat er een nieuwe Italiaanse studente ging komen). Veel succes met het schrijven van jouw thesis, ook al gaat dit over een duidelijk minder cool organisme dan *Drosophila*! **Lenz**, toen ik je voor het eerst ontmoette, dacht ik niet dat we goede vrienden

zouden worden, maar ik ben blij dat dit niet het geval bleek. Toen je vertrok was er toch beduidend minder animo op het lab en moest ik me meer laten gelden als man (om het genderevenwicht in balans houden etc.). Bedankt voor alle gesprekken over de twee belangrijkste W-woorden: Whisky en Wetenschap. **Tom**, bedankt voor je hulp met RNA extracties en qPCR, jouw handen zijn goud waard. Dr **Jason** Clements, even though we were not at the lab at the same time, many thanks for clearing the way for me to do research on the IPCs in the lab. **Emily**, veel succes met je doctoraat, maar ik ben er zeker van dat je dit enorm goed gaat doen. Tijdens je labrotatie had ik al door dat er een PhD in jou zit. Ook bedankt aan alle andere studenten, laborotaties en masterproeven, die ik in meer of mindere mate begeleid heb. Jullie moeilijke vragen hebben mij gedwongen om kritischer te zijn voor mezelf en mijn uitleg zo duidelijk en precies mogelijk te doen.

Tijdens mijn opleiding biomedische wetenschappen en de doctoraatsjaren heb ik ook veel gehad aan **Karen** en **Yoeri**, alhoewel ik nu Dr. Yu en Dr. Sleyp moet zeggen. Lunchpauzes met jullie waren altijd heel leuk, minder leuk was dat jullie voor mij hebben afgelegd (strebers!). Veel succes in jullie verdere loopbaan!

Ik bedank ook graag **David**, niet enkel een neef, maar vooral een hele goede vriend. Bedankt voor al jouw zelfgebrouwen biertjes. Ik wens je heel veel succes met jouw doctoraat. **Midas,** consultant par excellence, bedankt om een goede vriend te zijn. Op whisky-avonden, kwam je steevast op de proppen met boeiende onderwerpen en “hot takes”. Dr. **Sebastian** Kurten, ik heb grote bewondering voor hoe positief jij in het leven staat en hoe snel je jezelf de Nederlandse taal meester hebt gemaakt. Vrienden van de zaalvoetbal (**Timmie**, **Thomas**, **Ocula, Fré, Jerre,** …) bedankt om mijn wel zeer bedenkelijke voetbalkwaliteiten te tolereren.

Dit doctoraat was ook nooit tot stand gekomen zonder de steun van mijn familie die na al die jaren nog steeds niet echt weten wat ik doe. **Mama** en **Papa** bedankt om voor een omgeving te zorgen waar ik zonder problemen kon gaan studeren. Ook veel dank aan mijn broer en zus, **Kristof** en **Elien**, en hun partners, **Karen** en **Jochem** voor de gezellige zondagavonden in Oud-Heverlee. Alsook bedankt voor de financiële kater door mij niet één maar twee keer peter te maken van twee prachtige petekindjes: **Kaat** en **Gloria**. **Marleen** en **Patrick,** bedankt voor de leuke etentjes en gesprekken in Lier.

Saving the best for last. **Inge**, liefje, de steun die jij mij hebt gegeven doorheen het doctoraat is niet in woorden te beschrijven. Ook wanneer je het zelf moeilijk had met je eigen doctoraat stond je altijd paraat om te luisteren naar mijn gezaag. Bedankt om mijn psycholoog, chauffeur, agenda, maar vooral mijn vriendin te zijn.

Table of contents

Acknowledgments i

List of abbreviations v

List of gene symbols vii

Table of contents iii

Summary v

1 General introduction 1

Insulin signaling in humans 1

The beta cell and its transcriptional identity 2

Insulin processing and secretion 4

Insulin-IGF signaling pathway 7

IIS in disease 8

Insulin signaling in *Drosophila melanogaster* 10

The Insulin-Producing Cells 11

Regulation of the IPCs 16

dilp expression in other tissues 18

2 Objectives 20

3 Genome-wide analysis identifies Homothorax and Extradenticle as regulators of insulin in *Drosophila* Insulin-Producing cells. 21

Abstract 21

Introduction 22

Results 23

Discussion 35

Methods 38

Supplementary information 45

4 The *TCF7L2* homolog, Pangolin, is essential for *Drosophila* Insulin-producing cell development. 57

Abstract 57
Introduction 58
Results 59
Discussion 66
Methods 68
Supplementary information 71
5 Deciphering the transcription factor coding regulating expression of the *Drosophila dilp2*, *-3* and *-5* genes. 79
Abstract 79
Introduction 80
Results 81
Discussion 89
Supplemental information 95
6 Discussion 101
References 107
Scientific acknowledgments and personal contribution 123
Conflict of interest statement 123
Publications 124

List of abbreviations

AA	Amino acid
ADP	Adenosine diphosphate
ANOVA	Analysis of Variance
AO	Aorta
ATP	Adenosine triphosphate
cAMP	Cyclic adenosine monophosphate
CC	Corpora cardiaca
CRM	Cis-regulatory Module
CRN	Combined recurrent nerve
D	Dorsal
D. melanogaster	*Drosophila melanogaster*
DGRP	*Drosophila* Genome Reference Panel
DNA	Deoxyribonucleic acid
dpc	Days post conception
ER	Endoplasmatic reticulum
ES	esopghagus
ETC	Electron transport chain
FADH	Flavin adenine dinucleotide
GDP	Guanosine diphosphate
GMC	Ganglion mother cells
GPCR	G-protein coupled receptor
GTP	Guanosine triphosphate
GWAS	Genome-wide association study
HG	Hypocerebral ganglion
IHC	immunohistochemistry
IIS	Insulin and Insulin-like Growth factor Signalling
IPC	Insulin-producing cell
kDA	kilodalton
L1-3	Larval instar 1 -3
LDCV	Large Dense Core Vesicles
MPC	Multipotent pancreatic progenitors
mRNA	Messenger ribonucleic acid
NADH	Nicotinamide adenine dinucleotide
NSC	neurosecretory cells
P	Posterior
PI	Pars intercerebralis
PIP3	phosphatidylinositol (4,5)-trisphosphate
PV	Proventriculus
qPCR	Quantitative polymerase chain reaction
RDGN	Retinal determination gene network
RG	Ring gland
RNAi	RNA interference
SEZ	Suboesophageal zone
SP	Signal peptide
T1D	Type 1 diabetes
T2D	Type 2 diabetes

TAG	thoracicoabdominal ganglion
TCA	Tricarboxylic acid
TFBS	Transcription factor binding site
TSS	Transcription start site
VGCC	Voltage gated calcium channels
Y1H	Yeast-1-Hybrid

List of gene symbols

4EBP1	eukaryotic translation initiation factor 4E-binding protein 1
GAD65	65 kDa glutamic acid decarboxylase
achi	achintya
ALS	Acid labile solute
adipokinetic hormone	AKH
AS160	Akt substrate of 160 kDa
ara	araucan
Arm	Armadillo
arr	arrow
bap	bagpipe
Baldspot	Baldspot
bbc	bb in a boxcar
bsh	brain-specific homeobox
bru1/3	bruno 1/3
bun	bunched
cnc	cap-n-collar
CPE	carboxypeptidase E
CHES-1-like	Checkpoint suppressor 1-like
CCK	cholecystokinin
CkIIα-i1	CKII-alpha subunit interactor-1
croc	crocodile
CtBP	C-terminal Binding Protein
Dac	Dachshund
dimm	dimmed
dan	distal antenna
DH44	Diuretic Hormone 44
dInR	*Drosophila* Insulin receptor
dilp1-8	*Drosophila* insulin-like peptide
Dsk	Drosulfakinin
drm	drumstick
Mek	Dual specificity mitogen-activated protein kinase kinase 1
E2f2	E2F transcription factor 2
E23	Early gene at 23
Eip74EF	Ecdysone-induced protein 74EF
ena	enabled
en	engrailed
esn	espinas
exd	extradenticle
ey	eyeless
eya	eyes absent
Fit	Female-independent of transformer
Fit1	Fermitin 1
FGF	Fibroblast Growth Factor
fkh	fork head
foxo	forkhead box
FoxP	Forkhead box P
GLP-1	Glucagon-like peptide 1
GCK	Glucokinase
GLUT1/3	Glucose transporter 1/3
Glut4EF	Glucose transporter 4 enhancer factor
GluRIB	Glutamate receptor IB
GluClalpha	Glutamate-gated chloride channel subunit alpha
GSK-3	Glycogen synthase kinase 3

GramD1B	GRAM domain containing 1B
GFP	green fluorescent protein
Gbp1/2	Growth blocking protein 1/2
grb2	Growth Factor Receptor Bound Protein 2
GH	Growth Hormone
Gyc88E	Guanylyl cyclase at 88E
HP1b	Heterochromatin Protein 1b
hth	homothorax
Hr38	Hormone receptor-like in 38
Hr78	Hormone-receptor-like in 78
INS	Insulin
IR	insulin receptor
INSL	Insulin-like
IGF	Insulin-like growth factor
IGFBP1-7	Insulin-like growth factor binding protein1-7
IGFR	Insulin-like growth factor receptor
Irs	Insulin-Receptor Substrate
ISL1	ISL LIM Homeobox 1
kek5	kekkon 5
kirre	kin of irre
lmd	lame duck
LRP1	LDL receptor protein 1
Lgr3/4	Leucine-rich repeat-containing G-protein coupled receptor 3/4
Lar	Leukocyte-antigen-related-like
Lst	Limostatin
Ldsdh1	Lipid droplet subset dehydrogenase 1
MAFA	MAF BZIP Transcription Factor A
mTOR	mammalian target of rapamycin
Max	Max
mgl	Megalin
MEIS1/2	Meis homeobox 1/2
Mio	Missing oocyte
Marf	Mitochondrial assembly regulatory factor
MAPK	Mitogen-activated protein kinase
ERK	Mitogen-activated protein kinase 1
Mp	Multiplexin
Myo31DF	Myosin 31DF
CG43340	Myosin-7a binding protein
nerfin-1	nervous fingers 1
NGN3	neurogenin3
NEUROD1	Neuronal Differentiation 1
Nplp2	Neuropeptide-like precursor 2
NKX2.2	NK2 Homeobox 2
NKX6.1	NK6 Homeobox 1
nub	nubbin
Nf-YB	Nuclear factor Y-box B
opa	odd paired
optix	Optix
CG33970	oskyddad
p70S6K	p70 ribosomal protein S6 kinase
PAX4	Paired box transcription factor 4
PAX6	Paired box transcription factor 6
PDX	pancreatic and duodenal homeobox 1
pan	pangolin
ptc	patched

PBX1/2	PBX Homeobox 1/2
Prep1/2	PBX Regulated protein 1/2
PI3K	phosphatidylinositol (3,4,5)-trisphosphate Kinase
PKA	protein kinase 1
PC1/3	proprotein convertase 1
AKT	protein kinase B
ptth	Prothoracicotropic hormone
pum	pumilio
PDK	Pyruvate dehydrogenase kinase
Raf	RAF proto-oncogene serine/threonine-protein kinase
Ras	Rat sarcoma virus protein
REST	repressor element 1 silencing transcription factor
rp49	ribosomal protein 49
rps13	ribosomal protein S13
sd	scalloped
Sdr	Secreted decoy of InR
SPoCk	Secretory Pathway Calcium atpase
SHC	SHC-transforming protein 1
sNPF	Short Neuropeptide F
Shrm	Shroom
sick	sickie
Side	similar to Deadpan
sna	snail
SS	Somatostatin
SOS	Son of sevenless homolog 1
shh	Sonic Hedgehog
sun	stunted
Sug	Sugarbabe
sfl	sulfateless
Sdc	Syndecan
tkv	thickveins
tio	tiptop
TfAP-2	Transcription factor AP-2
TNF-α	Tumor necrosis factor alpha
toy	twin of eyeless
toe	twin of eyg
vis	vismay
Vsx1	Visual system homeobox 1
wb	wing blister
wor	worniu

Summary

Insulin is an important hormone and is released from the beta cells in the pancreas when there is a surplus of nutritional energy available in the blood. This activates peripheral tissues like adipose tissue and muscles to store this energy for later use. Dysregulation of insulin signaling results in the development of diabetes and contributes to the development of cancer.

In the commonly used model organism the fruit fly, *Drosophila melanogaster*, a similar mechanism is found. *Drosophila* insulin-like peptides (Dilps) which are structurally and functionally related to human insulin are produced, among other tissues, by the Insulin-Producing Cells (IPCs). In contrast to the human beta cell, IPCs produce up to four different dilps that are independently regulated. The IPCs, located in the brain, play a major role in the regulation of growth, metabolism, fecundity, stress response and fecundity. Interestingly, the IPCs show remarkable similarity to the beta cells, ranging from the way in which dilps/insulin are stored and secreted, to the expression and role of common transcription factors such as Eyeless (Ey)/PAX6.

The research presented in this thesis aims to deepen our understanding of the genetics of IPC development and of the transcriptional control of *dilp2*, *-3* and *-5.* In Chapter 3, we utilized publicly available transcriptome data to identify genetic variants associated with *dilp2*, *-3*, *and -5* transcript levels in a defined genetic population representative of the natural variation. Genes affected by these variants were further analyzed for their role in controlling *dilp2*, *-3*, *and -5* expression in the IPCs using an RNAi-based approach. The transcription factor Homothorax (Hth) emerged as a significant regulator of *dilp2* and *-3* expression. Its co-factor Extradenticle (Exd) was shown to have an equally important role in adult IPCs. Previous work had already established a role for Ey and Dachshund (Dac) in the IPCs. Collectively, these four transcription factors are best known for their role in the developing eye in the retinal determination gene network (RDGN). We then characterized the role of other members of the RDGN and found that many are functional in the IPCs to promote *dilp2* and *-5* expression. This regulatory network is thus re-used in the IPCs.

In Chapter 4, we defined a role for the *Drosophila TCF7L2* homolog *pangolin* (*pan)* in the differentiation and function of *Drosophila* IPCs. Variants in *TCF7L2* have been robustly and independently found to increase type 2 diabetes risk and several studies have indicated a role for *TCF7L2* in beta cell development. Here, we find that expression of a dominant negative form of Pan had massive deleterious effects on IPC development. Several key transcription factors such as Dimmed, Ey and (larval) Hth are severely reduced. This also resulted in the near absence of *dilp2*, *-3* and *-5* expression in the IPCs and severely affected their

morphological differentiation. This shows that TCF7L2/Pan has a conserved role in the differentiation of the IPCs.

Lastly in Chapter 5, we performed a Yeast-1-hybrid study of the proximal promoter of *dilp2*, *-3* and *-5* to identify the transcription factors directly involved in their regulation. This revealed that there is almost no overlap between the transcription factors regulating these three dilps showing that they are independently regulated. We also provide evidence for compensation and interdependence of *dilp* expression. Interestingly, we defined a role for a second *PAX6* homolog, *twin of eyeless* (*toy*), in the regulation of *dilp2* and IPC morphology. This further highlights the conserved association between PAX6 and insulin.

In summary, these results identified several transcription factors, including Hth, Pan and Toy with a role in *dilp* transcriptional control and/or IPC development. Their homologs have similar roles in the beta cell, providing additional evidence for an evolutionary conserved regulatory network controlling insulin expression.

1 General introduction

Hormones are chemical messengers that regulate various bodily functions such as growth and development, metabolism, and reproduction. Hormones play a crucial role in maintaining organismal homeostasis[1]. Upon release in the bloodstream, they will bind to specific receptors to trigger a response in target cells and tissues. Insulin is one such key hormone. In mammals, it is produced by the beta cells of the pancreas and is released to signal a state of nutrient excess to other cell types that can then respond accordingly by promoting the storage of surplus metabolic fuel. Excess or deficient insulin secretion can result in imbalances that result or can contribute to disease states such as cancer and diabetes[2].

Insulin and the related insulin-like growth factors (IGFs) are part of a family of hormones that are conserved across metazoans. A large body of work has shown that its members have conserved roles in the regulation of metabolism, growth, longevity and reproduction from nematodes to vertebrates[3–5].

Insulin signaling in humans

The pancreas, located behind the stomach in the abdomen, contains the beta cells that produce insulin. The pancreas is a gland with both an exocrine and an endocrine function. The exocrine pancreas constitutes most of the pancreatic tissue and produces digestive enzymes that are released into the duodenum to facilitate the breakdown of consumed fats, proteins and carbohydrates. The endocrine pancreas, on the other hand, is located in clusters of cells (Islets of Langerhans) throughout the exocrine tissue[6]. The beta cells are located within these islets where they are typically found in the core of the islets and account for ~60% of the cells thus forming the largest population of the islet endocrine cell types[7]. The alpha cells constitute the second largest population of endocrine cells (~35%) and are found in the periphery of the islets. They produce and secrete glucagon, a hormone that increases blood sugar levels and acts to counterbalance insulin. Furthermore, delta cells produce and secrete somatostatin (SS), a hormone that inhibits the release of both insulin and glucagon. In addition to these three cell types, epsilon cells produce ghrelin which stimulates appetite and regulates energy balance, and Pancreatic Polypeptide (PP) cells that produce PP, which helps to regulate digestion and nutrient absorption[6,8].

The beta cell and its transcriptional identity

All cells within the pancreatic islets or all other cells within the human body for that matter share the same genome that is roughly 3 gigabases long and contains ~20 000 protein-coding genes[9,10]. These genes can be transcribed into mRNA and translated into the proteins which perform most of the functions within the cells. This flow of information from DNA to protein is a process termed the centra dogma of molecular biology[11]. Controlling this flow of information is essential as different cell types need to express different genes. The beta cells, for example, need to express insulin (*INS*) and the appropriate machinery for insulin release into the bloodstream while muscle cells will not express these genes. The spatial and temporal aspects of gene expression are regulated by transcription factors that can interact with their DNA binding domains to transcription factor binding sites (TFBS) that are located in Cis-Regulatory Modules (CRMs)[12]. Examples of CRMs include the proximal promoter and the more distally located enhancers[13]. They are relatively short DNA sequences (100s base pairs) that typically contain multiple TFBS for different transcription factors[14]. The proximal promoter is located within 1 kilobase upstream from the transcription start site (TSS) while the enhancers can be located more than 1 megabase from the gene it controls but can interact with the promoter to steer gene expression via a process called DNA looping which is mediated by co-factors such as the mediator complex and cohesin[15,16]. In the proximal promoter of human *INS***,** the majority of the TFBS are located within 300 base pairs from the TSS[17–19] (Figure 1). Transcription factors can act cooperatively to increase *INS* transcription as is the case for NEUROD1 and PDX1 via their interaction with the p300 coactivator[20]. Total *INS* expression is thus the result of the integration of the interactions of the expressed transcription factors with the proximal promoter and enhancers, and by interactions with other co-factors, other transcription factors and intracellular pathways that modify transcription factor affinities and activity[21].

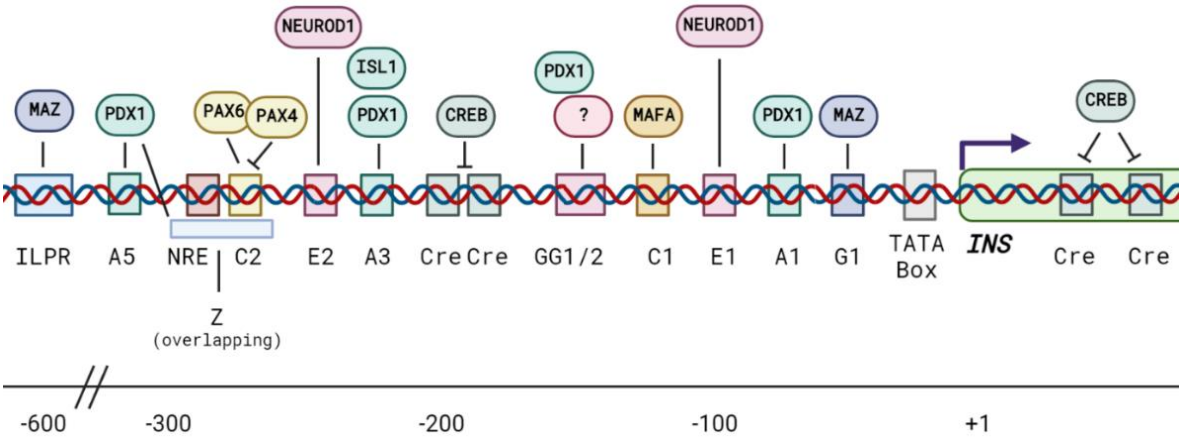


Figure 1. Transcriptional control of *INS*. The ~300bp region upstream of the TSS is the proximal promoter region and contains several TFBS that each can be bound by single or multiple transcription factors and controls the spatiotemporal expression of *INS*. Figure is based on Hay and Docherty[22]. *Made with BioRender*

Development of the beta cell

The developmental origin of the beta cell can be traced to 30 days post conception (30 dpc) with the appearance of a ventral and a dorsal pancreatic bud from the primitive digestive tube. The location of the dorsal bud is the result of exclusion of Sonic hedgehog (Shh) from the prepancreatic endoderm caused by its contact with the adjacent notochord[23,24]. The ventral bud is specified by a loss of FGF activity[25]. Both the ventral and dorsal bud are marked by the expression of the pancreatic and duodenal homeobox 1 (PDX1) transcription factor. PDX1 is essential for normal pancreas development as patients with PDX1 loss of function mutations present with pancreas agenesis[26]. The ventral bud will migrate posteriorly to join and fuse with the dorsal bud upon gut rotation at ~55 dpc[27]. Both the ventral and dorsal bud contain multipotent pancreatic progenitors (MPCs) that will split into exocrine progenitors and endocrine/ductal bi-progenitors based on the lower GATA4 expression by the latter at ~45 dpc[28]. This is followed by a single wave of transient Neurogenin3 (NGN3) expression which marks the specification to an endocrine cell fate[29]. These endocrine progenitors go on to generate FEV Transcription Factor^{+} cells that still lack any hormone expression but are an intermediate step towards more differentiated endocrine cells[30]. This is followed by further specification by the expression of transcription factors that program these cells to a specific endocrine cell fate. High levels of PAX4 and low ARX levels instruct the cell to take on a beta cell identity[31]. Other transcription factors that are essential for beta cell identity include PAX6, NKX2.2, NKX6.1, ISL1 and NEUROD1 (Figure 2)[27,32]. Several of these transcription factors are, besides their role in beta cell differentiation, additionally required in the differentiated beta cell where they are involved in *INS* transcriptional control (Figure 1).

The beta cells present at birth are terminally differentiated and are characterized by the production of a single hormone, insulin, and the machinery required for hormone processing and secretion, but remain functionally immature. Beta cell maturation occurs postpartum as the result of changes in nutrition and includes the acquisition of appropriate glucose stimulated insulin secretion[33]. The transcription factor MAFA plays a key role in beta cell maturation and controls transcription of genes associated with granule formation and release[34,35].

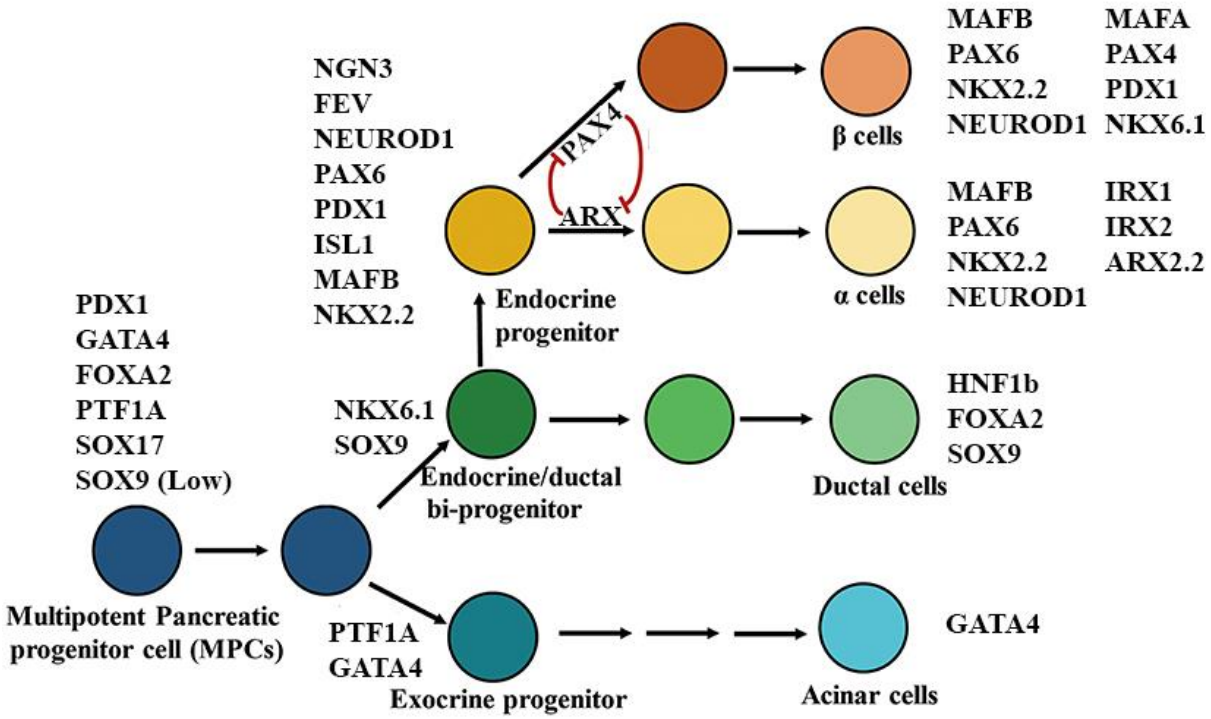


Figure 2. **Beta cell development from pancreatic progenitors.** A myriad of transcription factors orchestrate the differentiation of the different cell types that constitute the pancreas. An initial split between exocrine and endocrine/ductal progenitors is followed by a transient wave of NGN3 expression which gives rise to the endocrine progenitors which are already marked by several transcription factors (NEUROD1, PDX1, PAX6) that are found in mature beta cells. High PAX4 is essential for beta cell identity while ARX is necessary for alpha cell fate. Modified from Zhong and Yiang under the terms of the creative commons license CC BY[36]. Note that for the sake of simplicity delta, epsilon and PP cells which also derive from the endocrine progenitors are omitted from the figure.

Insulin processing and secretion

After *INS* is transcribed in the beta cells, the mRNA encoding preproinsulin is translated. Preproinsulin contains an N-terminal signal peptide (SP) which enables its translocation to the lumen of the endoplasmic reticulum (ER). The SP is subsequently removed to generate proinsulin which contains the remaining A-C-B sequence. This is followed by the generation of two disulfide bonds between the A and B chain and one disulfide bond within the A chain itself[37]. Proinsulin moves through the Golgi complex into the trans Golgi network and is sorted into secretory granules containing proprotein convertase 1 (PC1/3) and carboxypeptidase E (CPE) that act in concert to remove the C-peptide to generate mature insulin[38]. Upon accumulation in these granules, insulin forms dimers and in the presence of Zn^{2+} ions will start to self-assemble into stable hexamers. Mature granules have a dense core that contains the insulin hexamers and a clear periphery that contains the soluble C-peptide (Figure 3)[39].

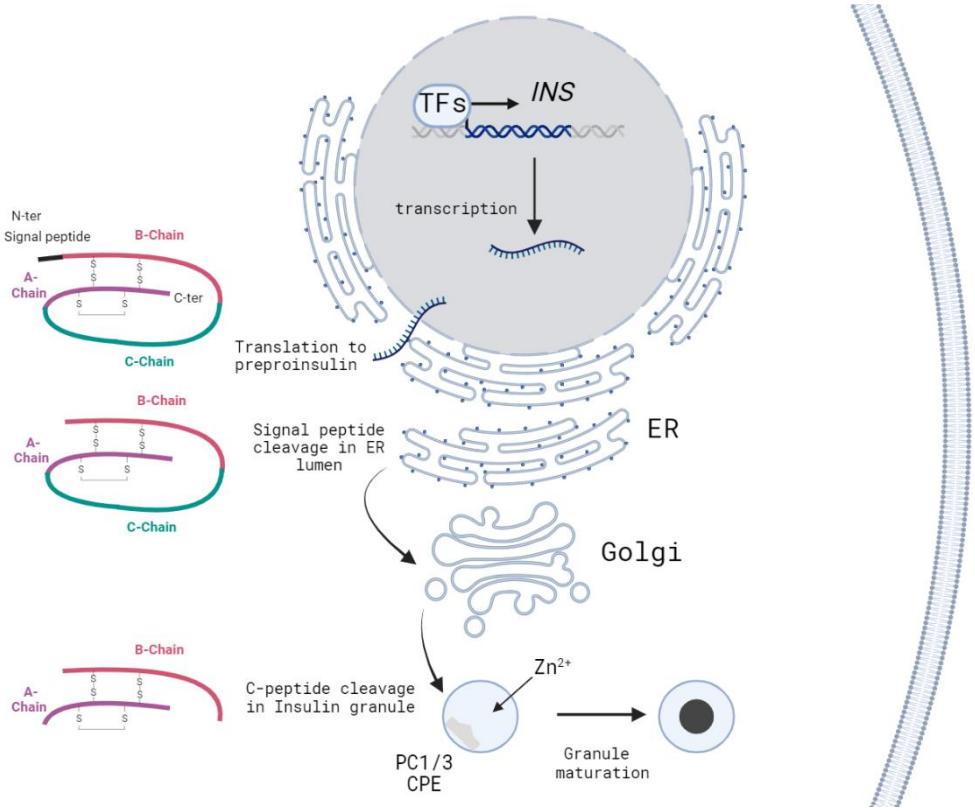


Figure 3. Production of mature insulin in the beta cell. The *INS* mRNA is translated and preproinsulin is transferred to the ER lumen, the SP is cleaved, and proinsulin moves via the Golgi into secretory granules where the C-peptide is removed. Zn+ influx results in the generation of insulin hexamers in mature granules. *Made with BioRender.*

The release of these granules is tightly regulated in function of the physiological status of the animal. Glucose is the major regulator of insulin secretion. The sensing of glucose occurs directly by the beta cells through the uptake of glucose from the blood vessels surrounding the pancreatic islets via the GLUT1 and GLUT3 transporters, in a process termed glucose-stimulated insulin secretion[40]. Following entry into the cell glucose is phosphorylated by Hexokinase (GCK) to trap it in the beta cell. Subsequent metabolization results in the generation of mitochondrial NADH/FADH which fuels the electron transport chain (ETC) to increased ATP/ADP ratios which causes the ATP sensitive K^+-channels to close. This closure causes the cell to depolarize and the activation of voltage-gated calcium channels (VGCC). The subsequent Ca^{2+} influx stimulates the fusion of insulin granules with the outer membrane of the beta cells resulting in the release of the readily releasable insulin granules into circulation[18,41]. This initial release is called the triggering phase and lasts around 5-10 minutes. The second phase of insulin release is more sustained with lower-level secretion and is facilitated by the amplifying pathway. This pathway increases the number of insulin granules that are readily releasable. It is stimulated by various nutritional signals like lipids and by hormonal signals such as Glucagon-like peptide 1 (GLP-1) that stimulate intracellular cAMP

production with subsequent activation of Protein Kinase A (PKA). This results in PKA recruitment to the insulin granules where it stimulates secretion[41–43].

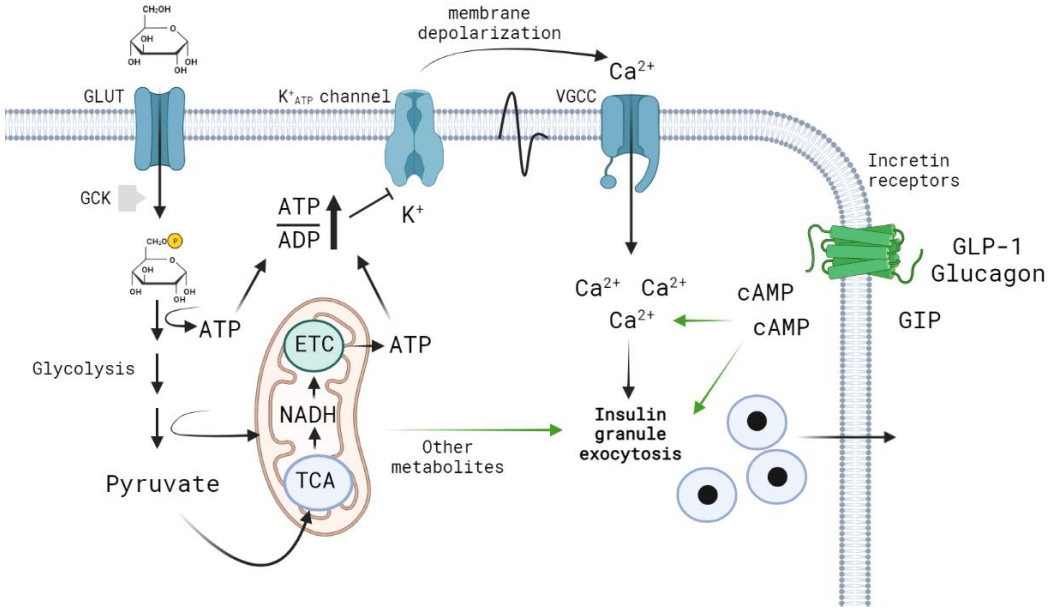


Figure 4. Glucose stimulated insulin secretion. Glucose uptake by the beta cells results in an increased ATP/ADP ratio that closes ATP sensitive K^+-channels. Subsequent opening of VGCCs results in Ca^{2+} influx which stimulates insulin granule fusion with the outer membrane. Incretin receptors such as the GLP-1 receptor modulate and enhance sustained insulin secretion. *Made with BioRender.*

Insulin is the founding member of the insulin superfamily of genes that encode for proteins with shared characteristic structural domains. The A and B chains are relatively invariant across the family and contain the conserved spaced cysteine residues that are considered the "insulin signature". The motif CC-3X-C-8X-C in the A chain in particular is characteristic of the insulin superfamily[44]. Of note are the insulin-like growth factors 1 and -2 (IGF-1 and IGF-2) that together with insulin are involved in what is called Insulin and Insulin-like growth factor Signaling (IIS). Structurally, IGF-1 and -2 contain a shorter C-chain that is not cleaved out in mature IGF, a small additional D-chain also present in mature IGF and a large E-peptide that is processed and plays an important role in the regulation of IGF bioavailability[45].

IGF-1 is expressed by a variety of tissues but the highest levels are found in the liver where its biosynthesis is under the control of Growth Hormone (GH), insulin and nutrients[46]. IGF-1 is primarily an adult growth signal while IGF-2 is important during gestation but remains abundant in adults[47]. In circulation, the vast majority of IGF is bound by IGF binding proteins 1-7 (IGFBP1-7) that have a higher affinity for IGF than the IGF-1 receptor (IGF1R). The majority of IGF-1 and -2 in circulation is bound by IGFBP3 in a complex with IGF acid labile solute (IGFALS)[48]. The classical view of IGFBPs posits that they limit IGF bioavailability but

also significantly extend IGF half-life allowing IGFs to spread in circulation. This passive role of IGFBPs was later challenged by the finding that they have IGF-independent functions and that they are produced by a variety of tissues suggesting that they could play a role in the local regulation of IGF action[49]. The remaining members of the insulin superfamily consist of Relaxin-like 1-3 and Insulin-like 3-6 that signal via G-protein coupled receptors (GPCRs) to regulate processes related to reproduction like INSL3, for example, which stimulates germ cell survival[50].

Insulin-IGF signaling pathway

IIS is mediated by three receptors in vertebrates: the insulin receptor (IR), IGF-1 and IGF-2 receptors (IGF1R and IGF2R). IR and IGF1R are evolutionarily related and share significant sequence similarity. They both activate the same IIS pathway but have biases in their downstream activation. IGF2R does not have any signaling potential due to the lack of a tyrosine kinase domain and is evolutionarily unrelated to IGF1R and IR. It acts mainly to regulate circulating IGF-2 levels[51]. Alternative splicing of *IR* mRNA results in the generation of two possible isoforms termed IR-A and IR-B[52]. IR-B differs from IR-A due to the inclusion of the 36bp exon 11. These isoforms differ only slightly in their affinities for insulin, but IR-A has significantly higher affinity for both IGF-1 and -2. IR-A thus acts to convey a mitogenic, cell-proliferation program while IR-B is more associated with a metabolic program. The relative abundance of both isoforms is dependent on the cellular and regulatory context[53]. Both IR isoforms are proteolytically cleaved by different members of the proprotein convertase family depending on the specific tissue and result in α- and β-chains that are disulfide-linked[54]. They then form homodimers (e.g. IR-A/IR-A) or heterodimers (e.g. IR-A/IR-B). IGF1R is similarly processed and can additionally form hybrid heterodimers with both IR-A and IR-B[55]. The importance of these hybrid receptors is relatively poorly understood, but they are functional and can activate the shared downstream pathway[56]. As mentioned before, both IR and IGF1R largely share the two main canonical downstream pathways, RAS/MAPK and PI3K/AKT, but IGF1R preferentially activates the RAS/MAPK while IR has a bias towards PI3K/AKT[57]. This bias is consistent with the growth stimulating effects of the IGFs and a more metabolic role for insulin (Figure 5).

The conformational change as a result of ligand binding to the receptor dimers allows for the autophosphorylation of three tyrosine residues of its activation loop which itself results in phosphorylation of tyrosine residues on Insulin-Receptor Substrate proteins (IRS) and SHC. The signal then splits into two main signaling pathways. The first emanates from the binding of phosphorylated IRS to the membrane bound phosphatidylinositol (3,4,5)-trisphosphate Kinase (PI3K) which converts phosphatidylinositol (4,5)-bisphosphate (PIP2) to

phosphatidylinositol (3,4,5)-trisphosphate (PIP3). PIP3 subsequently activates PDK which in turn activates Protein Kinase B (AKT). AKT is upstream of several important insulin effectors: (1) Glycogen synthase kinase 3 (GSK-3) which is involved in the regulation of glycogen synthesis; (2) Mammalian target of Rapamycin (mTOR) a major regulator of protein synthesis, which itself acts as a nutrient sensor and stimulates the eukaryotic translation initiation factor 4E-binding protein 1 (4EBP1) and p70 ribosomal protein S6 kinase (p70S6K) via their phosphorylation; (3) FOXO transcription factors that regulate the expression of genes involved in adipogenesis, mitogenesis, gluconeogenesis and cell survival; and (4) GLUT4 translocation to the cell membrane via Akt substrate of 160 kDa (AS160)[58–60].

The RAS-MAPK pathway is the second major pathway downstream of IIS. It starts with the recruitment of SOS/GRB2 to IRS and SHC proteins which converts membrane bound RAS-GDP to RAS-GTP and results via RAF/MEK in the phosphorylation of ERK1-2. This phosphorylation allows its translocation to the nucleus and ERK1-2 phosphorylates an array of transcription factors to regulate transcription into a pro cell proliferation and differentiation program[58,61].

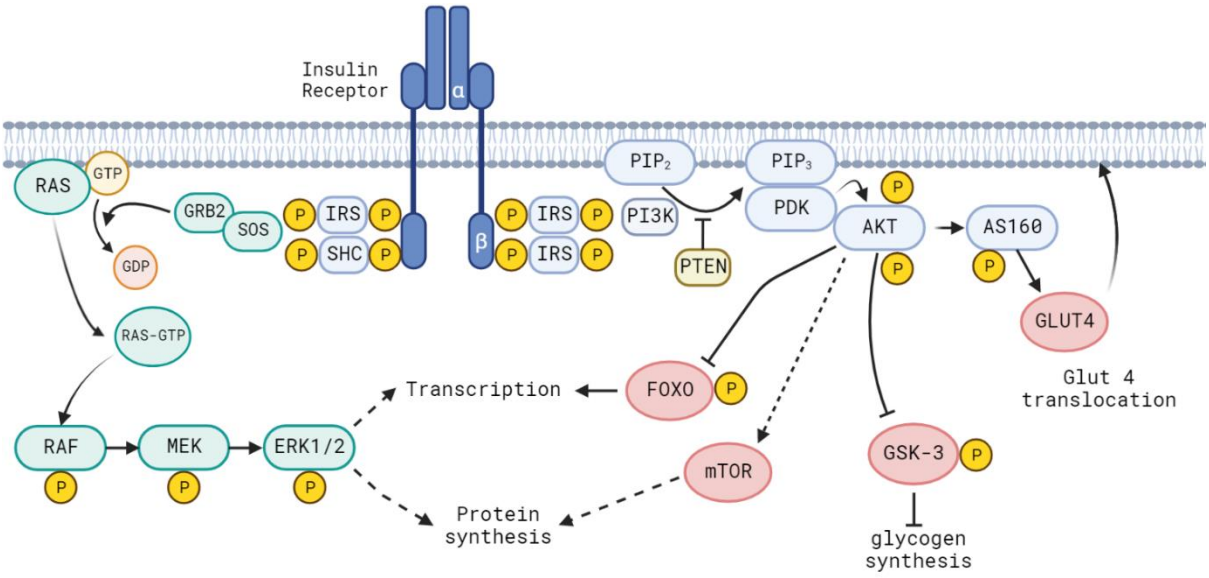


Figure 5. **The Insulin signaling pathway**. Ligand binding results in the activation of two major canonical pathways. One culminating in ERK1/2 (Left) and the other in AKT (Right). Dashed lines indicate indirect interactions. Note that not all interactions are shown for the sake of clarity. Based on Boucher *et al.*[58] *Made with BioRender.*

IIS in disease

The association between IIS and disease is best known in the form of diabetes, a disease that has three main types. Type 1 diabetes (T1D) is the result of an autoimmune-mediated destruction of the beta cells[62]. This results in an absolute insulin deficiency and its

symptomatic onset usually occurs in childhood. Autoantibodies against insulin and/or 65 kDa glutamic acid decarboxylase (GAD65) can usually be detected several months before the patient begins to present with symptoms. Genome-wide association studies (GWAS) revealed that most variants associated with T1D are involved in the immune system[63].

Type 2 diabetes (T2D), on the other hand, is characterized by the development of insulin resistance in the peripheral tissues (muscles, liver, and adipose tissue). This insulin resistance results in a compensatory increase in insulin release from the beta cells. T2D then develops when beta cells can no longer adequately compensate for the increased insulin resistance, a process called beta cell failure which, like in T1D, leads to a loss of glucose homeostasis[64]. The majority of T2D patients are overweight or obese and the development of insulin resistance is unequivocally linked to the increased fat accumulation as a result of a disequilibrium between caloric intake and expenditure[65]. Accumulation of visceral fat, in particular, seems to be key in the development of T2D. Secretion of pro-inflammatory adipokines from visceral fat reduces peripheral insulin sensitivity and can inhibit insulin secretion from the beta cells[66]. T2D has a large genetic component with an estimated heritability between 30% and 70%. GWAS of T2D patients have identified variants in more than 700 genes that are associated with T2D[67]. Identified variants have relatively low effect sizes and are often implicated in beta cell functions such as insulin secretion[68]

A third form of diabetes, called gestational diabetes, occurs during pregnancy when there is increased peripheral insulin resistance due to hormonal changes and increased inflammation with the goal of keeping circulating glucose sufficiently high for the developing fetus. Like in T2D, glucose homeostasis is lost if the beta cells are unable to compensate for insulin resistance[69]. Importantly, this defective beta cell response is already present before pregnancy but will manifest clinically by the increased insulin resistance caused by pregnancy[70]. The similarity with T2D also extends to the genetic level where the same variants detected in T2D GWA studies have similar effect sizes in gestational diabetes[71].

Due to the canonical role of IIS in growth regulation and cellular survival, it is no coincidence that IIS contributes to tumor development and progression[72]. The compensatory increase of insulin due to insulin resistance and the associated increased IGF-1 production promotes tumoral growth. The hyperglycemia following beta cell failure additionally promotes cancer cell proliferation as glucose is their preferred fuel source[73]. High glucose levels also induce the generation of reactive oxygen species (ROS) that stimulate tumor progression and high glucose protects cancer cells from apoptosis[74]. Tumor cells tend to overexpress IR-A and IGF1R biasing IIS towards the mitogenic pro-survival side of the pathway[75]. Furthermore,

cancer cells often gain the ability to produce and secrete IGF-2 which allows the tumor to provide its own proliferation signals[76].

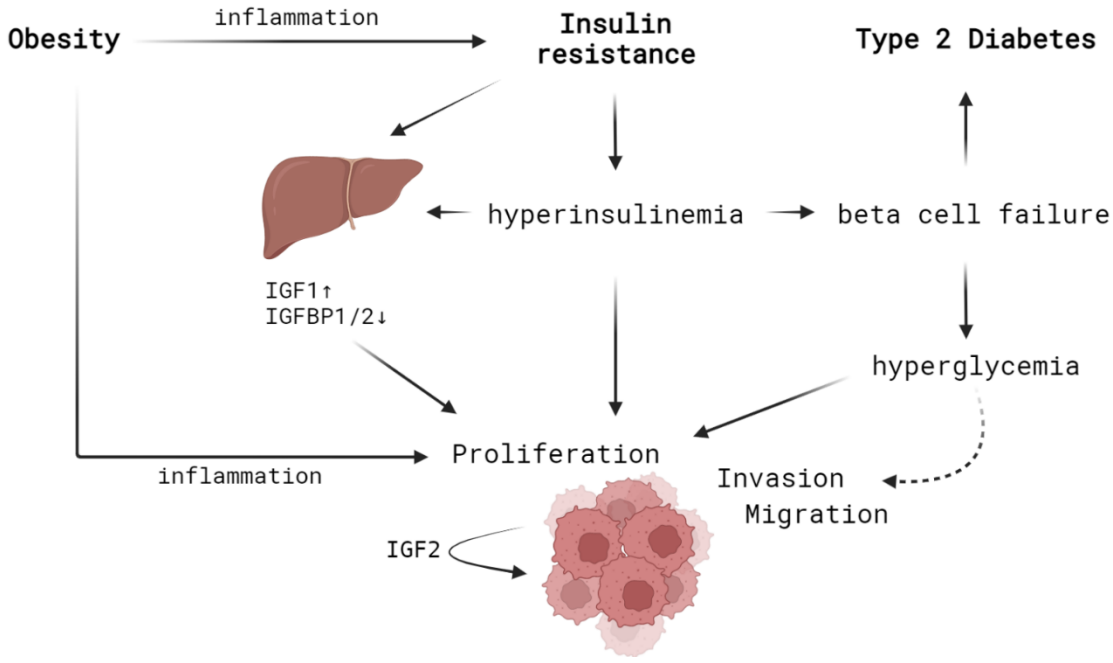


Figure 6. IIS in cancer and diabetes. The interactions between obesity, T2D and cancer are partially a result of dysregulation of IIS as result of insulin resistance. *Made with BioRender.*

Insulin signaling in *Drosophila melanogaster*

The insulin signaling pathway is remarkably conserved across metazoans and is well-understood in the commonly used model organism *Drosophila melanogaster*[77]*.* Its genome encodes for a single Insulin Receptor (dInR) that is similar to the mammalian IR with regard to its organization and processing. *dInR* encodes for two distinct isoforms, one of which has a larger β-subunit containing a region with high IRS homology[78]. A separate gene encodes for a single IRS protein ortholog called *chico* and activation further downstream such as ERK activation, dFOXO translocation and activation of TOR are also conserved[77,79,80]*.* In the mammalian insulin signaling pathway there are often multiple paralogs (e.g. AKT1-3) that only have a single ortholog in the fly (Akt). On the ligand side, however, eight *Drosophila* insulin-like Peptides or Dilps are encoded in the *D. melanogaster* genome. *dilp1-4* are clustered together on the third chromosome and *dilp5* is located 20kb further along the chromosome while *dilp6*, -7 and -*8* are dispersed in the genome[3]. The spaced conserved cysteine residues in the A chain, typical of insulin-related genes is found in all Dilps (Figure 7 A)[3,81]. Dilp1-6 are all thought to signal via dInR and its downstream pathway while Dilp7 and -8 activate the GPCR receptors Lgr4 and Lgr3, respectively[82–84]. The protein structure of Dilp5 has been

solved and was shown to be highly similar to human insulin despite relatively low shared AA identity: 27.8% (Figure 7 B). Dilp5 was also shown to bind the human IR and was able to lower blood glucose levels in rats[85]. Dilp and mammalian insulin members are thought to derive from a single ancestral insulin gene[86]. The accurate determination of their phylogenetic relationships, however, is difficult due to their divergent AA sequences and short protein lengths (Figure 7 C)[87]. The spatiotemporal expression profiles of *dilp1-8* are complex and dynamic which shows that they are differentially regulated[3].

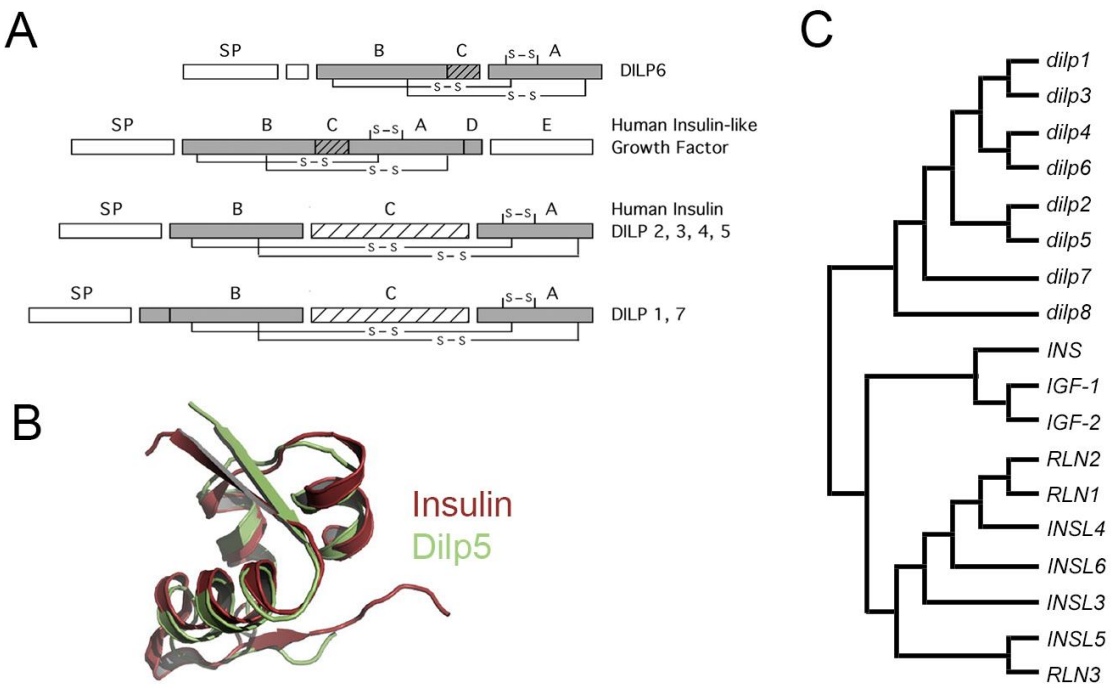


Figure 7. Insulin and Dilp comparison. (A) Comparison of the structural organization of Dilp1-7 and human insulin and IGF. Reproduced from Brogiolo *et al* (License number: 5523070473772)[3]. (B) Crystal structure of Dilp5 superimposed on human insulin. From Sajid *et al* under the terms of the creative commons license CC BY[85]. SP: Signal peptide. (C) Hypothetical evolutionary relationships between members of human and *Drosophila* insulin family, Based on Grönke *et al* and Chan *et al*[88,89].

The Insulin-Producing Cells

The Insulin-Producing Cells (IPCs) are a major source of Dilp production and express *dilp1*, *-2*, *-3* and *-5*. Initial studies revealed that genetic ablation of these cells results in a major developmental delay and growth retardation while also increasing circulating sugar levels, reminiscent of diabetes[90]. Subsequent publications revealed that IPCs are responsible for or involved in various aspects of physiology. Metabolically, reducing IPC activity was found to be sufficient to increase stored glycogen and triglyceride levels and of considerable interest were the findings that attenuating IIS activity increases lifespan[89,91,92]. The IPCs are additionally directly involved in the maintenance of germ stem cell division and reduced IPC activity

decreases lifetime fecundity[93]. Moreover, the IPCs have been shown to play a role in stress resistance (oxidative damage, starvation), sleep and feeding behavior[92,94,95]. In summary, IPCs are major players in the regulation of physiology.

Developmentally, the IPCs derive from a pair of neuroblasts (NB), one per hemisphere, out of a group of Castor$^+$, Vsx1$^+$, Optix$^+$ cells in the anterior neuroectoderm whose earliest progenitors arise from the area corresponding to the anterior stripe of *giant* expression[96,97]. The first neuroblast to delaminate from this group will start to express Dachshund (Dac) and will then undergo 5 asymmetric cell divisions to generate 5 ganglion mother cells (GMCs). The IPCs arise from the division of these first 4 GMCs which are marked by the expression of Dac, Eyeless (Ey) and Tiptop (Tio). The embryonic stage 17, just before larval hatching, is the onset of *dilp2* expression in the IPCs[90,97]. Expression of *dilp5* then follows in late L1 larval stage followed by *dilp3* in mid-late L3[98,99]. Lastly, expression of *dilp1* starts early in pupal development and is maintained for 2-3 days post adult eclosion[100]. A distinct temporal aspect of *dilp* expression in the IPCs can thus be observed where different *dilp* genes are turned on throughout development in a sequential manner.

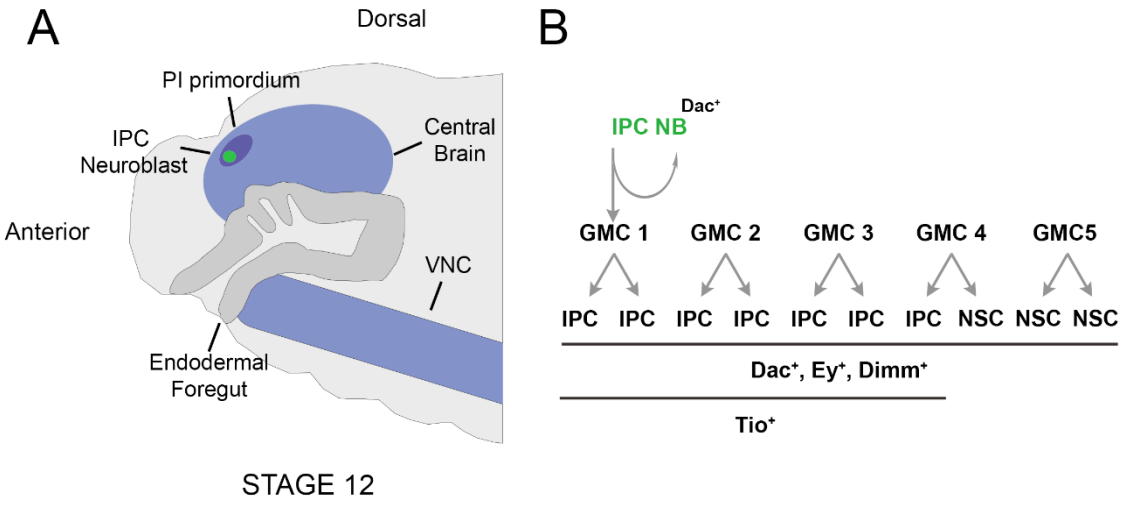


Figure 8. IPC embryonic development. (A) Anterior part of a stage 12 embryo showing one lobe of the prospective central brain (blue) with the Pars intercrebralis Primordium (dark blue) containing a single Dac$^+$ NB (green). (B) Generation of 5 GMCs from this NB by asymmetric cell division. Subsequent symmetric cell divisions generate 10 - 12 neurosecretory cells (NSC) of which 7 are IPCs and express Dac, Ey, Dimm and Tio. Figure is based on Hartenstein[101] and Hwang and Rulifson[102].

The cell bodies of the IPCs are located in the pars intercerebralis (PI). Their neurites in the larval stage extend laterally and posteriorly of which some remain ipsilateral to terminate in the suboesophageal ganglion while others project contralaterally and terminate outside the brain on the aorta, corpora cardiaca (CC) and esophagus (Figure 9 A)[90,103]. These projections

outside the brain are the probable release site for Dilp secretion. The CC produces Adipokinetic hormone (AKH) a hormone analogous to glucagon. It acts to counterbalance Dilp function and thus acts to avoid hypoglycaemia[104,105]. In the larval stage, the CC consists of 14 cells and forms the ring gland together with the ecdysone-producing prothoracic gland and the Juvenile Hormone-producing corpora allata[106]. This connection between the alpha cell-like CC and the beta cell-like IPCs has been termed the fly pancreas[107].

During metamorphosis the posterior IPC neurites from both hemispheres will from a single bundle that has a fan shape within the suboesophageal ganglion while others will terminate on the aorta, CC, crop and proventriculus via the esophageal foramen (Figure 9 B). During this period the IPC cell bodies from both hemispheres will merge and form a single pyramid-shaped cluster[103,108].

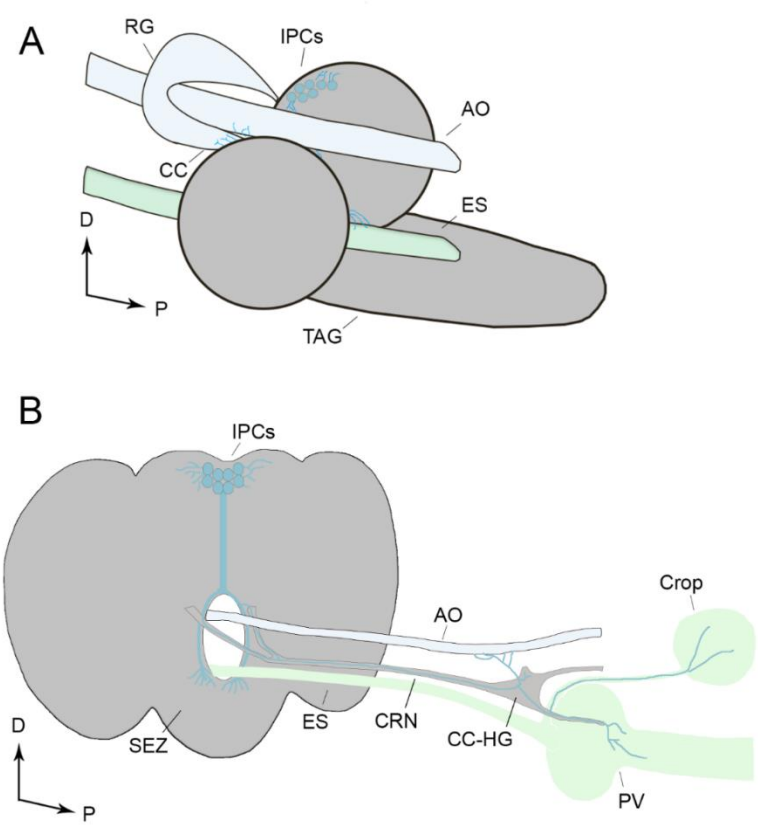


Figure 9. Anatomy of the IPCs within the larval and adult brain. (A) The larval IPCs are located in the pars intercerebralis of both hemispheres and extend their neurites to the AO, CC and ES while some stay inside the brain. Note that to allow for a full visual appreciation of the innervation pattern only the IPCs and neurites of a single hemisphere are shown. (B) Anatomy of the IPCs in the adult stage. IPCs neurites terminate via the CRN on the aorta, proventriculus/anterior midgut and crop. The TAG is not depicted in the adult stage. Based on Nässel and Zandwala[109]; AO, Aorta; CC: Corpora cardiaca, CC-HG Corpora cardiaca- hypocerebral ganglion, CRN: Combined recurrent nerve, ES: esophagus, PV: proventriculus, SEZ: Suboesophageal zone, RG: Ring gland; TAG: thoracicoabdominal ganglion. D, Dorsal and P, Posterior.

The larval and adult IPCs are typical neurosecretory cells with large cell bodies to facilitate the significant peptide machinery required to produce sufficient amounts of Dilp protein. Like in mammals, there is most likely a maturation of Dilp protein by processing enzymes, as isolated Dilp2 and Dilp5 only contain the A- and B-chain[85,110]. The role of specific enzymes involved in this process are as of yet not fully identified, but the PC2 homolog *amontillado* is enriched in the IPCs and IPC-specific knockdown results in reduced adult size consistent with possible involvement in Dilp processing[103]. Mature Dilp protein is then stored in Large Dense Core Vesicles (LDCV) which are transported along IPC axons where they can be released into the hemolymph or within the brain[103,111,112]. The large cell bodies and LDCVs are the result of the transcription factor Dimmed (Dimm) which acts as a scaling factor to promote, but not specify a pro-secretory cellular function [111,113–115].

dilp expression in the IPCs

The function and role of Dilp1 is the least well understood of all IPC-derived Dilps. In normal conditions its expression is limited to the pupal stage and during a few days post eclosion. Experimental induction of reproductive diapause, when feeding is strongly reduced, can also induce *dilp1* expression which then persists for several weeks[100]. Dilp1 is thought to provide a pro-growth signal in a period where feeding is restricted (e.g. metamorphosis, reproductive diapause). Dilp6 plays a similar but larger role (see below)[116–118]. There is also an interesting epistatic relationship between Dilp1 and Dilp2 in the adult stage, where loss of *dilp2* expression strongly induces *dilp1* expression[119]. *dilp2* deletion mutants show increased lifespan, and it was shown that this phenotype entirely depends on the induction of *dilp1* expression. As increased lifespan is usually associated with reduced systemic IIS, Dilp1 would presumably need to act as a dInR antagonist in the adult stage, while still acting as an agonist in the pupal stage (e.g. stimulation of growth). The exact function and mechanisms that would facilitate this remain unresolved[118].

In contrast to Dilp1, Dilp2 is the most studied Dilp. There are likely several reasons for this, including the fact that it is most homologous to human insulin (35% amino acid similarity) and that it is also the first one to be detected in the IPCs[3,90]. *dilp2* expression is high during the L1 stage to moderately drop during larval development and pupation but is by far the most abundant Dilp in the adult IPCs[92,116]. Of all single IPC Dilp mutant flies, those with deleted *dilp2* show the strongest phenotypes. They take longer to develop and emerge as smaller flies[89]. Additionally, it is the most potent growth inducer of all Dilps. Weak overexpression of Dilp2 results in much larger flies (51%) as opposed to more moderate size phenotypes for other Dilps (25-35%)[98]. Dilp2 also seems to be uniquely necessary for the regulation of circulating trehalose levels[89,120]. Expression of *dilp2* is mostly independent of nutritional conditions during

the larval stage but Dilp2 secretion is highly dependent on the presence of AA[98,99,121]. Dilp2 is additionally detected in larval surface glia where it plays a role in neuroblast reactivation[122]. Further but more limited evidence exists for the expression of *dilp2* in imaginal discs, salivary glands and trachea-associated cells[3,123].

Expression of *dilp3* in the IPCs is significantly lower than *dilp2* and starts in the mid-late L3 stage and remains high during pupation. Its expression is strongly reduced upon starvation in larvae[3,98,116]. Dilp3 mutants do not show any major phenotypes besides reduced lifetime fecundity. Interestingly, the loss of *dilp3* results in an accompanying drop in *dilp2* and *dilp5* levels which shows that Dilp3 directly or indirectly acts to promote *dilp2* and *-5* expression. This stands in contrast with most other *dilp* mutants which show compensatory upregulation of other dilps and it is this compensatory upregulation that makes it hard to identify clear roles for specific Dilps[89]. Expression of *dilp3* was also detected in the visceral muscle of the larval midgut where it is locally activated to stimulate intestinal stem cell proliferation[124].

The expression of *dilp5* is highly dependent on nutrition. Like *dilp3* it is highly reduced upon starvation and Dilp5 protein will accumulate in starved larvae[99]. *dilp5* mutants do not show any phenotypes under normal food conditions but are developmentally delayed and smaller in restricted but not starved food conditions[99]. Expression of *dilp5* is additionally detected in the principal cells of the Malpighian tubules where it plays an autocrine role in the regulation of stress responses[125]. In initial in-situ hybridization studies *dilp5* was also detected in the ovarian follicle cells[98].

Dilp1, -2, -3 and -5 deriving from the IPCs have partially overlapping functions, as evidenced by the increased severity of growth and developmental timing phenotypes upon deletion of multiple *dilp* genes. Yet several lines of evidence indicate that they are not fully redundant, like the unique role of Dilp2 in trehalose regulation and the independent transcriptional regulation (see below)[89]. Of considerable interest is the question as to how the different Dilps deriving from the same cells and activating the same receptor could possibly elicit different responses. One possibility is the selective release of Dilps, which was shown for Dilp3 where it is selectively secreted in response to high trehalose levels in the larval stage[126]. Another possibility is that it could be due to differential activation of the same receptor (dInR) as it was shown that Dilp2 and Dilp5 elicit largely shared but also distinct global protein phosphorylation patterns[127]. These biases are reminiscent of those seen in the context of IR and IGF1R downstream activation in mammals.

In addition to the Dilps, other neuropeptides are expressed in the IPCs. In larvae and adult there is heterogeneous expression of a cholecystokinin (CCK)-like peptide termed Drosulfakinin (Dsk) in the IPCs. Its knockdown in the IPCs leads to increased intake of food

and an upregulation of *dilp2*, *-3* and *-5*[128]. It was also shown that two Diuretic hormone 44 (DH44) neurons located adjacent to the IPCs start to express *dilp2* but not *dilp3* or *-5* in the adult stage and it is these neurons that form the projections to the crop[129].

Regulation of the IPCs

The key role of the IPCs in regulating metabolism necessitates a meticulous regulation that can differ between the larval and adult stage. Autonomous sensing of glucose by the IPCs, for example, is restricted to the adult stage while larval IPCs depend on the CC cells to sense circulating glucose levels[126,130,131]. The larval IPCs can, however, autonomously sense AA and secrete Dilp2 and -5 in response to leucine but no such mechanism has been reported in adults[132,133]. This switch between the larval and adult IPCs seems to partially depend on the presence of the transcription factor HNF4 which is implicated in the regulation of genes associated with glucose-stimulated insulin secretion[134].

A large number of non-autonomous signals regulate the IPCs (Figure 10). One important group of non-autonomous signals are those that are associated with the nutritional status of the fly. The *Drosophila* fat body, an organ equivalent to the mammalian adipose tissue and liver, is the principal nutrient sensor and relays humoral signals to the IPCs to ensure appropriate IIS levels[135].The two peptides (SunA/B) encoded by the *stunted* gene, for example, are secreted by the larval fat body in response to high dietary AA levels which are detected via TOR signaling. SunA/B act directly on the IPCs via the Methuselah receptor to stimulate Dilp2 and -5 secretion[136]. Two additional peptides, Growth-blocking peptides 1 and 2 (GBP1/2), are similarly involved in the regulation of Dilp2 and -5 secretion in response to dietary AA but act via intermediate neurons that inhibit the IPCs[137,138]. In the adult stage, female specific secretion of Female-independent of transformer (Fit) induces Dilp2 secretion in response to high AA[139]. Besides AA, the IPCs are also regulated by dietary sugars and fats which stimulate Unpaired2 (Upd2) secretion from the fat body that in turn stimulates Dilp2 and -5 secretion from the IPCs via a reduction of synaptic contacts between inhibitory GABAergic neurons and the IPCs[140,141]. Strikingly, human Leptin can compensate for the phenotypes observed in *upd2* mutants[140]. Starvation induces the release of Eiger, a TNF-α ortholog, from the larval fat body to inhibit *dilp2* and *-5* expression[142]. Besides their role in larval glucose sensing, the CC cells additionally control the IPCs in the adult stage. Upon starvation, they secrete limostatin (Lst), a decretin, to inhibit Dilp2 secretion from the IPCs[143].

Neuronal control of the *dilps* has also been shown. Short Neuropeptide F (sNPF) stimulates *dilp2* and *-5* expression as well as Dilp2 secretion in response to dietary sugar[144,145]. Serotonin, on the other hand, inhibits *dilp2* and *-5* expression while octopamine inhibits *dilp3* expression[146,147]. Furthermore, recent work has shown that the IPCs are inhibited during

locomotion and that this inhibition occurs before the onset of locomotion, showing that it acts as an anticipatory signal to provide sufficient mobilized nutrients during flight or walking[148]

The transcriptional control of *dilp* expression is less well understood (Figure 10). The PAX6 ortholog, Ey, was found to directly regulate *dilp5* expression, a process dependent on physical interaction with Dac[149,150]. This interaction is subject to nutritional control by dFOXO as starvation results in the nuclear translocation of dFOXO. This results in Dac being outcompeted by dFOXO for Ey binding and this leads to the loss of *dilp5* expression[99]. Additional transcription factors that are known to affect *dilp* expression include *sugarbabe* (Sug) whose overexpression significantly reduces *dilp2*, *-3* and *-5* mRNA levels but no direct regulation was shown[151]. The transcription factor Missing oocyte (Mio) was shown to inhibit *dilp3* expression while *lame duck* (*lmd*), the *Drosophila* ortholog of *GLIS3*, promotes *dilp2* expression[130,152].

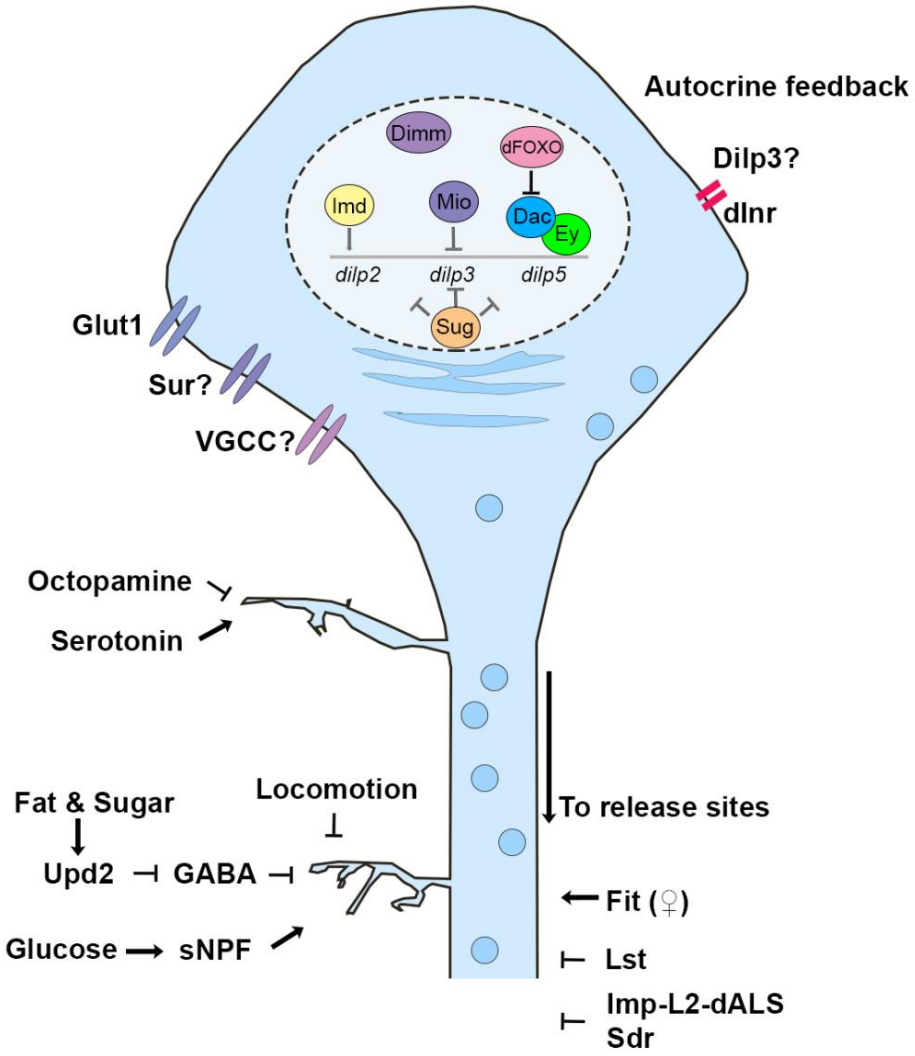


Figure 10. Cell autonomous and non-autonomous control of the IPCs. Dilp secretion and expression are regulated by transcription factors, humoral signals deriving from CC and Fat body, and from neurons within the brain. Only the non-autonomous regulators which have been identified in adults are shown. Figure is based on Alfa and Kim[153].

The activity of the IPC-derived Dilps is also regulated in circulation via their interactions with circulating proteins that modulate their activity. Dilp2 and Dilp5 can form ternary complexes with Imp-L2 and dALS, which are orthologs of IGFBP7 and IGFALS, respectively[154,155]. This association results in the sequestration of Dilp2 and -5, thereby decreasing their biological effects. Additionally, it was recently shown that Imp-L2 is necessary for the uptake and storage of Dilp2 in the CC to be locally used to stimulate growth in response to nutrient shortage. This shows that these circulating proteins can play a key role in the local regulation of Dilp functions[156]. Secreted decoy of InR (Sdr) is another Dilp interactor which seems to preferentially bind Dilp3 and antagonizes Dilp function[157].

Based on the total body of work on IPC regulation and function it appears that the IPCs share several features with the beta cells, some of which include the presence of shared transcription factors (Ey/PAX6), the packaging of insulin into LDCVs which are released by Ca^{2+} flux and granule secretion caused by autonomous glucose sensing[111,130,149]. One striking difference is their developmental origin as the IPCs are ectodermal while the beta cells derive from the endoderm[23,97]. Despite this, it seems that the beta cells are very neuron-like. For example, they do not express the repressor element 1 silencing transcription factor (REST) which is necessary to repress a neuronal phenotype[158]. Furthermore, several other species from Mollusca to Tunicata express insulin-like peptides in neurons, lending support for the hypothesis that the evolutionary origin of the beta cell is a neuronal insulin-producing cell[159,160].

dilp expression in other tissues

The other members of the *Drosophila* insulin family: *dilp4*, *-6*, *-7* and *-8* are expressed in other tissues and cell types and some aspects of their regulation and functions are summarized below.

Expression of *dilp4* is high during the embryonic stage and *dilp4* mutants exhibit aberrant wound healing upon damage[3,161]. Dilp4 acts in an autocrine manner via the IIS pathway to facilitate actin remodeling and cellular repair. In the adult stage, Dilp4 was recently shown to be expressed in adult cortex glia where it acts in an autocrine manner to increase glial glucose uptake to shuttle glucose towards neurons to meet their increased metabolic demands in the context of long term memory formation[162].

Dilp6 shares certain characteristics with vertebrate IGF-1 that distinguishes them both from insulin as both their peptide sequences do not contain obvious cleavage sites for an internal C-peptide[3,116]. It is produced in the fat body where it provides a growth promoting signal in non-feeding conditions such as the wandering larval stage and during pupal development, while also promoting growth in starved larvae[116,117]. Expression of *dilp6* in the fat body is

partially controlled by dFOXO signaling, which is itself controlled by the IIS pathway[163]. Dilp6 is also produced by glia in the larval stage where it stimulates neuroblast re-activation after periods of quiescence. Dilp6 is itself regulated by AA availability via autonomous TOR sensing and non-autonomously via signals from the fat body[122,164]. Glial Dilp6 is also involved in the regulation of *dilp5* expression in the IPCs. IPC-derived Dilp5 in turn stimulates glial Dilp6 release establishing a positive feedback loop[99]. There is thus a striking contrast where, depending on its tissue of origin, Dilp6 is induced by starvation in the fat body to act in an endocrine manner to buffer growth defects as well as in normal nutritional conditions in glial cells to act in a paracrine manner on the IPCs to stimulate systemic growth.

Dilp7 is produced in the abdominal neuromeres of the ventral nerve cord in larvae and adults[165,166]. The posterior Dilp7+ dMP2 neurons innervate the hindgut muscles. Silencing these neurons results in an increase in feeding behavior in response to low nutrition conditions[167]. A subset of these dMP2 neurons arise post-embryonically with a female bias in their number and function. These post-embryonic Dilp7+ neurons innervate seminal vesicles in males and the oviduct in females. Silencing these neurons results in female but not male infertility[168]. The exact function of Dilp7 in these neurons is not yet fully understood as RNAi-mediated KD of *dilp7* or Δ*dilp7* mutants show no decrease in fecundity[89,168]. Overexpression of Dilp7 using Dilp7-GAL4 alters the preference of females on which media they preferentially lay eggs[166]. Recent findings showed that Dilp7 acts as a co-transmitter on downstream abdominal leucokinin neurons to facilitate noxious light avoidance behaviour[84]. Dilp7 acts, at least in this context, via the relaxin-like Lgr4 receptor which has co-evolved with Dilp7 throughout the arthropod branch[84,169]. Previously, Dilp7 was presumed to signal via the InR as ubiquitous overexpression results in enhanced body size[98].

Dilp8 is a relaxin and was discovered as a non-autonomous signal secreted by damaged imaginal discs to delay pupariation, allowing the damaged disc to regenerate[81,170]. Dilp8 also acts in this context as a growth limiting factor on other organs to inhibit their growth and ensure that the animal attains symmetry in the adult stage[81]. This is supported by the observation that flies without functional Dilp8 exhibit increased left-right asymmetry[81]. Upon secretion by the imaginal discs Dilp8 acts on a pair of Lgr3+ neurons that are in contact with Ptth-neurons which in turn are connected to the ecdysone-producing prothoracic gland. Additionally, they are also in contact with the IPCs where they reduce *dilp3* and *dilp5* expression which in turn results in lower Juvenile hormone levels and attenuation of organismal growth[82,83]. Dilp8 can thus coordinate tissue growth with developmental timing, resulting in developmental stability and robustness. Dilp8 is also produced in the follicle cells of the ovaries where it regulates ovulation and fecundity, reminiscent of the function of mammalian relaxins[171].

2 Objectives

The IPCs play a key role in the regulation of many physiological processes and the production and secretion of *dilp2, -3* and *-5* depends on the integration of several identified non-autonomous signals. However, the regulatory networks active in the IPCs themselves which control IPC development, function and *dilp* expression are comparatively poorly understood. The overall aim of this PhD project was therefore to identify and characterize regulatory mechanisms active in the IPCs with a focus on transcriptional control and to compare these results with the situation in the mammalian beta cell.

To achieve this aim, we have specified three main objectives:

1. To identify cell-autonomous regulators of IPC development and function using a novel GWAS approach based on the natural variation in *dilp2*, *-3* and *-5* levels across the DGRP. This will also serve as a proof of concept for further research using a similar approach using variation of transcript levels as a method for candidate gene identification. From this screen, we will characterize the role of a candidate gene in depth.

2. To characterize the role of Pangolin in the development and function of the IPCs. Pangolin is the homolog of *TCF7L2,* a type 2 diabetes risk gene. This will expand the evidence related to the evolutionary relationship between IPCs and the beta cells and

3. To identify and validate the transcription factors that are involved in the direct regulation of *dilp2*, *-3* and *-5* in the IPCs using a Yeast-1-Hybrid approach.

3 Genome-wide analysis identifies Homothorax and Extradenticle as regulators of insulin in *Drosophila* insulin-producing cells.

Winant, M.[#1], Buhler, K.[#1], Clements, J.[#1], De Groef, S.[1], Hens, K.[2], Vulsteke, V.[1], Callaerts P.[1]

1 Laboratory of Behavioral and Developmental Genetics, Department of Human Genetics, KU Leuven—University of Leuven, Leuven, Belgium,
2 Centre for Functional Genomics, Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brooks University, Headington Campus, Oxford, United Kingdom

# These authors contributed equally

This chapter has been published in *PLOS Genetics* 18, e1010380 (2022).

Abstract

Drosophila Insulin-Producing Cells (IPCs) are the main production site of the *Drosophila* insulin-like peptides or dilps which have key roles in regulating growth, development, reproduction, lifespan and metabolism. To better understand the signaling pathways and transcriptional networks that are active in the IPCs we queried publicly available transcriptome data of over 180 highly inbred fly lines for dilp expression and used dilp expression as the input for a Genome-wide association study (GWAS). This resulted in the identification of variants in 125 genes that were associated with variation in dilp expression. The function of 57 of these genes in the IPCs was tested using an RNAi-based approach. We found that IPC specific depletion of most genes resulted in differences in expression of one or more of the dilps. We then elaborated further on one of the candidate genes with the strongest effect on dilp expression, Homothorax, a transcription factor known for its role in eye development. We found that Homothorax and its binding partner Extradenticle are involved in regulating *dilp2*, *-3* and *-5* expression and that genetic depletion of both transcription factors shows phenotypes associated with reduced insulin signaling. Furthermore, we provide evidence that other transcription factors involved in eye development are also functional in the IPCs. In conclusion, we showed that this expression level-based GWAS approach identified genetic regulators implicated in IPC function and dilp expression.

Introduction

Insulin/IGF signaling (IIS) plays a key role during growth and development but also in regulating reproduction, lifespan, stress resistance and metabolic homeostasis[3,89,172,173]. This central role is conserved across metazoans and beyond[174]. Eight Dilps (*Drosophila* insulin-like peptides) have been described. Dilp1-6 are predicted to bind a single *Drosophila* Insulin Receptor (InR) that results in the activation of the conserved insulin signaling pathway[77]. Dilp7, produced by a group of sexually dimorphic neurons in the ventral nerve cord, acts on the receptor LGR4, and Dilp8, produced in the imaginal discs, binds the receptor LGR3[3,83,84,168,175]. Dilp1-3 and 5 are produced in the brain by 14 larval and 16 adult insulin-producing cells (IPCs)[129]. Dilp1 is expressed only after pupation and regulates lifespan and metabolism epistatic to Dilp2[100,116,119]. The function of Dilp4 is unknown and it is expressed only in the embryonic and larval midgut[3]. Dilp6 is more structurally homologous to mammalian insulin-like growth factors and is produced in the fat body and glia. It plays an important role in the non-autonomous regulation of the IPCs [99,116]. Of all Dilps, those most homologous to mammalian insulin and also best studied are Dilp2, -3 and -5, produced in and secreted by the IPCs[90]. In adults, Dilp2 seems to be exclusively expressed in the IPCs, while Dilp3 and Dilp5 are additionally expressed in midgut muscle and in the principal cells of the Malphigian tubules in adults, respectively[125,176].

The Dilps produced by the IPCs have been shown to exhibit independent roles while also being partially redundant in their function. Δ*dilp2* mutants, for example, show increased *dilp3* and *dilp5* levels as a compensatory effect. Nonetheless, Δ*dilp2* mutants still have elevated haemolymph trehalose levels not seen in any other single *dilp* mutant. In addition, loss of *dilp2* has a prominent effect on growth, that is, however, further exacerbated by additional loss of *dilp3* and *dilp5*[89,127,177]. Additional evidence for this independent regulation comes from the numerous non-autonomous signals deriving from other tissues that have been shown to regulate the expression or secretion of only a subset of Dilps[178]. Some of these signals are larval or adult-specific. The fat body plays a key role in this as it relays the nutritional status of the organism to the IPCs to adjust systemic IIS levels[135]. Low nutritional protein levels, for example, induce the release of Eiger from the fat body into the haemolymph. Eiger binds to and activates its receptor Grindelwald on the IPCs to inhibit *dilp2* and *dilp5* expression but not *dilp3*[142]. Neuronal input differentially affects *dilp* expression. Knockdown of the octopamine receptor OAMB yielded increased *dilp3* expression while leaving *dilp2* and *dilp5* levels unaltered, while knockdown of the 5-HT1a results in increased *dilp2* and *dilp5* and unaltered *dilp3*[147]. The autonomous factors acting in the IPCs that mediate this independent regulation remain poorly understood.

To identify novel cell-autonomous regulators of IPC *dilp* expression we use the *Drosophila* Genome Reference Panel (DGRP) consisting of more than 200 highly inbred fly lines whose genome has been sequenced. The DGRP has been used effectively to identify genes that determine many quantitative phenotypes such as starvation resistance[179], aggressive behavior[180], mushroom body morphology[181], cocaine consumption[182] and lifespan[183]. While previous studies relied on quantitative behavioural or physiological parameters, we used the variation in *dilp2*, *-3* and *-5* mRNA levels of 183 (DGRP)[179,184] fly lines to identify genetic variants that correlate with *dilp2*, *-3* and *-5* expression levels. Genes associated with these variants were validated for a role in determining *dilp2*, *-3* and *-5* levels in the IPCs with an RNAi-based approach. We then elaborated on the function of one of the strongest candidate genes, the homeobox transcription factor Homothorax (Hth). We show that it is an important regulator of *dilp2*, *-3* and *–5* expression in the IPCs and is necessary for normal systemic insulin signaling and regulates adult size and developmental timing. We then show that the Hth interactor Extradenticle (Exd) is equally required in the adult but not in the larval IPCs. Based on the previously established roles of Ey, Dac and now Hth/Exd in regulating both IPC and eye field development we then hypothesized that other members of the Retinal Determination Gene Network (RDGN) could also be active in the IPCs[149,150,185]. We show that most RDGN genes are expressed in the IPCs and that they are mostly necessary for normal *dilp2 and -5* expression. These results point to the existence of a reshuffled RDGN being active in the IPC to control insulin expression.

Results

GWAS for *dilp2*,*3* and *5* expression levels reveals sexually dimorphic putative regulators.

We collected whole body expression data for *dilp2*, *-3 and -5* of 183 DGRP lines[184] (Figure 11 A; Supplemental File 1). Considerable interline variation is observed and significantly higher read counts were detected in males compared to females for the three *dilps*. We then ran a GWA analysis for each of the *dilps* and identified a total of 271 variants that were associated with at least one *dilp* and in one sex (Figure 11 B). These 271 variants correspond to 125 unique genes (Figure 11 B). Based on RNA-seq data of the adult IPCs to ensure that genes were expressed in the adult stage (Supplemental file 2) and RNAi line availability we assessed the role of 57 candidate genes (Supplemental File 3) in regulating *dilp2 -3* and -5 expression in the IPCs using qPCR. This showed that all but 7 had significant effects on *dilp2*, *-3* or *-5* regulation in one or both sexes (Figure 11 C).

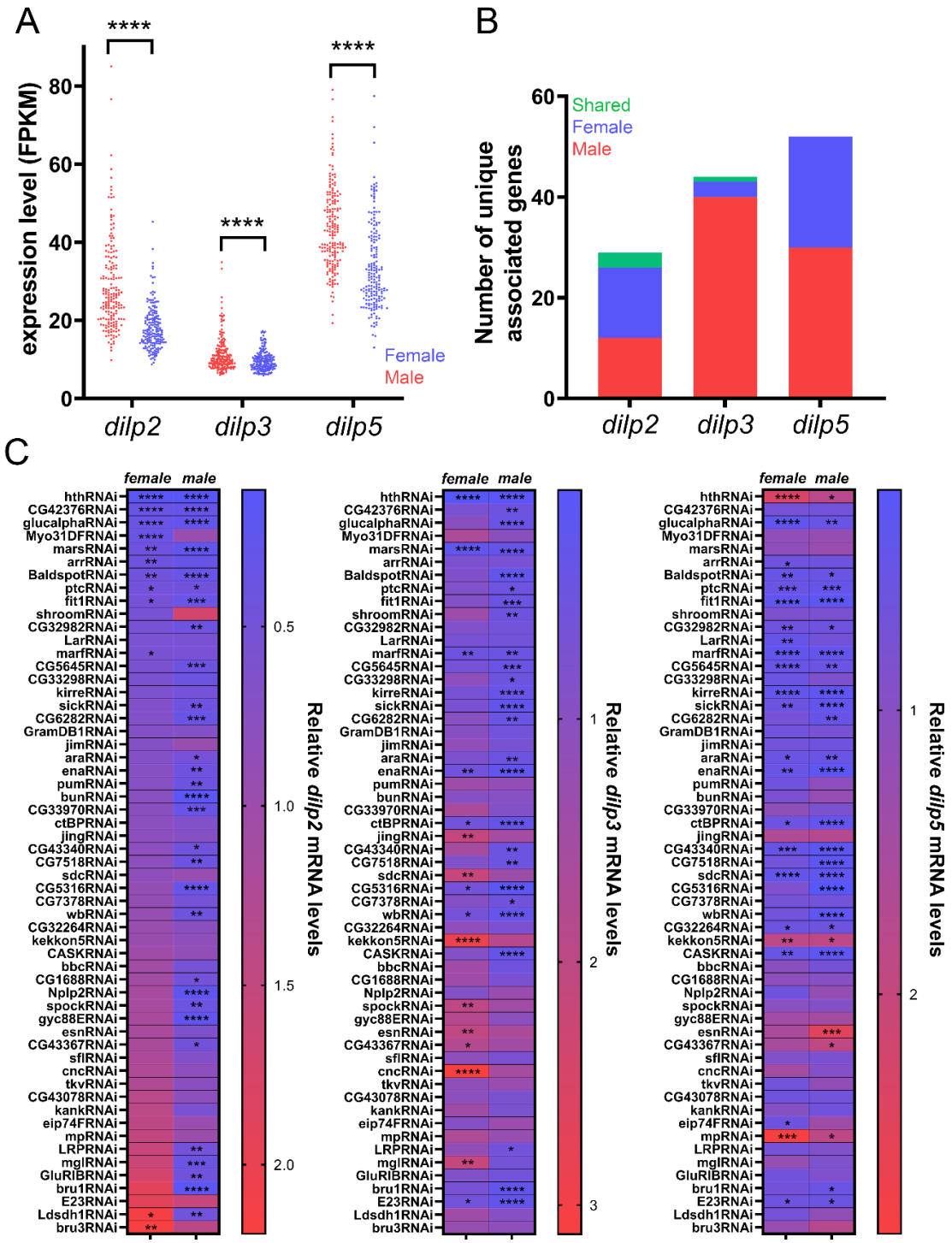


Figure 11: DGRP Screen and validation. (A) Mean *dilp2*, *-3 and dilp5* FPKM levels were collected from Huang et al.[184] and are significantly higher for males than females for *dilp2* ($p < 0.0001$, unpaired Welch's T-test, n=183), *dilp3* ($p < 0.0001$, unpaired Welch's T-test, n = 182) and *dilp5* ($p < 0.0001$, student's T-test n = 182). (B) Number of unique genes associated for each dilp/sex (male dilp2: 12, female dilp2: 14, shared dilp2: 3, male dilp3: 40, female dilp3: 3, shared dilp3: 2, male dilp5: 30, female dilp5: 22 and shared dilp5: 0). (C) Mean qPCR levels for *dilp2*, *-3 and dilp5* upon RNAi-mediated knockdown using *Dilp2-GAL4^{R};btlGAL80* (values are means of the relative mRNA quantity of n=3 consisting of 10-15 fly heads each). Results of female *dilp2* levels were used to sort the RNAi lines (low to high). (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, Dunnett's multiple comparisons compared to an *mCherry^{RNAi}* control line, which itself had no effect on *dilp* expression (Figure S1))

Most genes that were identified in the GWAS were identified in only one sex suggesting a highly sexually dimorphic regulation of *dilp* expression (Figure 11 B). However, when comparing changes in expression of the *dilp*s upon knockdown of the 57 candidate genes in the IPCs we noticed that the *dilp* levels in both sexes showed significant correlation (Figure 12 A-C). To better understand this apparent contradiction with the GWAS results, we classified per *dilp* each candidate gene as sexually dimorphic if RNAi mediated depletion in the IPCs resulted in a significant reduction or increase of *dilp* mRNA levels in one sex only, or if it had an opposite effect on *dilp* levels in both sexes. The latter possibility was only seen for *Ldsdh1* where knockdown resulted in increased *dilp2* levels in females and decreased *dilp2* levels in males (Figure 12 A'). Overall, we found that out of 57 candidate genes, knockdown of 34 genes resulted in changes in *dilp2* expression of which 79% (27) were sexually dimorphic (Figure 12 A'). Changes in expression of *dilp3* were seen for 34 out of 57 candidate genes with 76% sexually dimorphic (Figure 12 B'). Finally, 30 out of 57 genes resulted in altered expression of *dilp5* with 30% being sexually dimorphic (Figure 12 C'). Of note, the genes that result in changes in *dilp* expression upon knockdown are not the same for each *dilp*, consistent with the hypothesis that there is significant independent regulation of transcription for the *dilps*. In conclusion, we identify significant sexually dimorphic effects for regulation of *dilp2* and *dilp3* and less for *dilp5*.

Homothorax knockdown reduces *dilp2* and *-3* mRNA expression and Dilp2, -3 and -5 protein levels

The strongest decline in both *dilp2* and *dilp3* levels in males and females was noted for the transcription factor Homothorax. IPC-specific downregulation resulted in significant downregulation of both *dilp2* and *dilp3* mRNA levels while *dilp5* mRNA levels were significantly increased (Figure 13D, F, H)). First, we validated the expression of Hth within the IPCs of both the L3 and adult stage by using antibodies directed against Hth (Figure 13 A – B). We then elaborated on the effect of Hth on *dilp* expression by characterising the effect on Dilp protein levels by using specific antibodies against Dilp2, -3 and -5[123]. When compared to sibling controls this revealed a strong reduction of Dilp2 -3 and -5 immunoreactivity within the IPC somata (Figure 13 C-G'). Dilp2-GAL4 was used to knock down Hth levels. Since Dilp2-GAL4 is controlled by *dilp2* regulatory sequences, it is possible that Hth knockdown could affect GAL4 activity. To verify this, we quantified IPC GFP levels of *Dilp2-GAL4> nlsGFP;hth^{RNAi}* and found a reduction of GAL4 activity (Figure S2 A-C). However, when we used a second GAL4 line (*OK107-GAL4*, inserted in *eyeless*, and not regulated by Hth (Figure S3)[186]), the observed reductions in *dilp2* mRNA levels in both males and females were recapitulated showing that the effect of *hth* knockdown on *Dilp2-GAL4* activity has a negligible impact on the observed phenotypes (Figure S2 D-E).

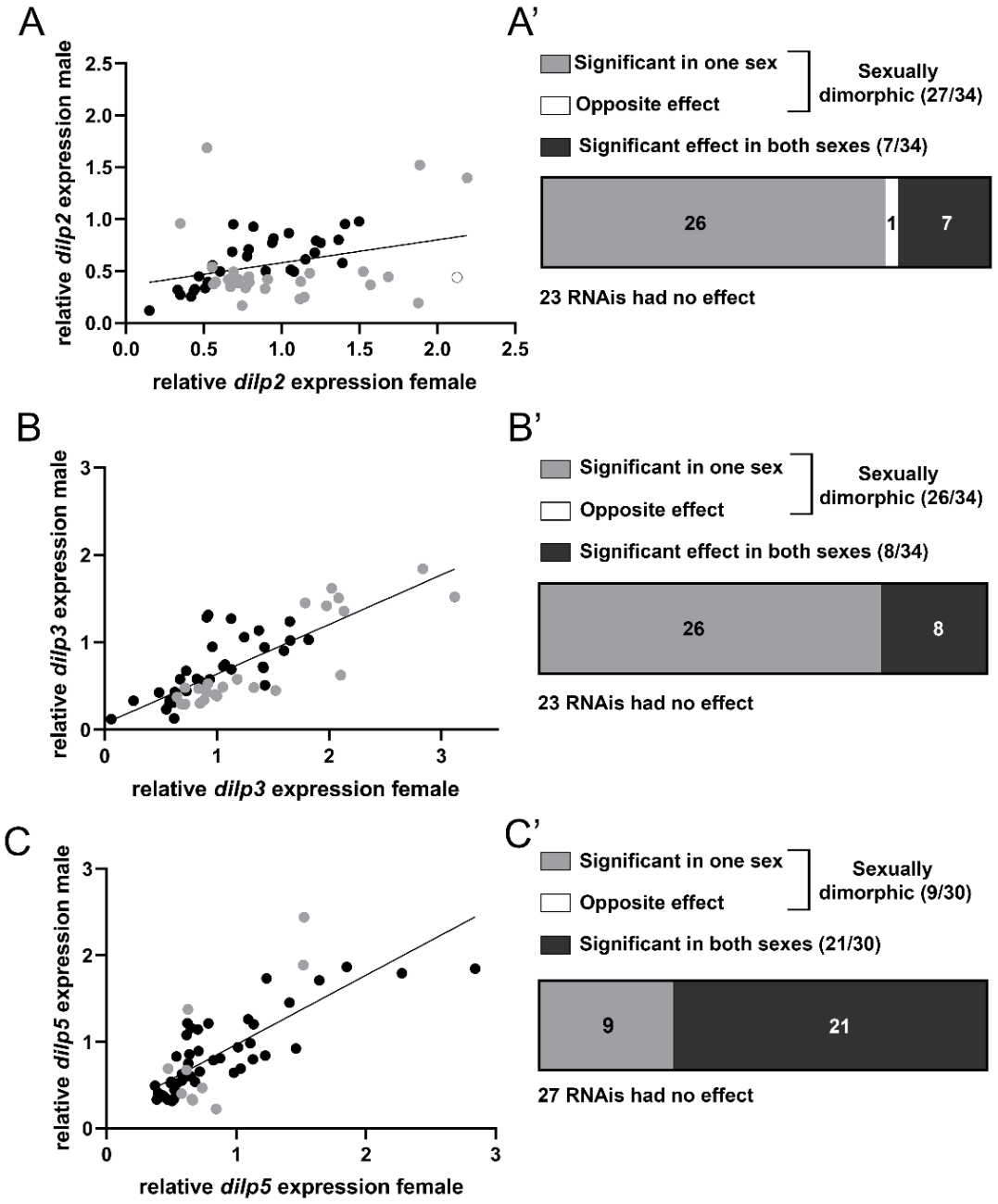


Figure 12: Knockdown of candidate genes in the IPCs produces correlated effects on *dilp* expression between the sexes. (A) The mean relative quantity of *dilp2* upon RNAi-mediated knockdown in the IPCs between female and male flies is mildly correlated (Spearman r = 0.3467, p = 0.0081). (A') Out of 57 tested genes, 27 had sexually dimorphic effects on *dilp2* expression, one (*Ldsdh1*) having opposite effects on *dilp2* expression. 30 genes had no effect (23) or the same effect (7) on *dilp2* levels. (B) Relative dilp3 between the sexes are highly correlated. (Spearman r = 0.7955, p < 0.0001) (B') 26 genes had a sexually dimorphic effect on *dilp3* expression, 31 had no effect (23) or the same effect (8). Relative quantity levels of *dilp5* are correlated (Spearman r= 0.7699, p < 0.0001). (C') 9 genes had a sexually dimorphic effect on *dilp5* levels while 48 genes had no effect (27) or the same effect (21) on *dilp5* expression.

However, when we used a second GAL4 line (*OK107-GAL4*, inserted in *eyeless*, and not regulated by Hth (Figure S3)[186]), the observed reductions in *dilp2* mRNA levels in both males and females were recapitulated showing that the effect of *hth* knockdown on *Dilp2-GAL4* activity has a negligible impact on the observed phenotypes (Figure S2 D-E).

Knockdown of the Hth co-factor, Extradenticle, induces a similar reduction in *dilp* expression and Dilp protein levels as observed upon Hth depletion.

During development, Hth physically interacts with the homeobox transcription factor Exd. Exd forms a heterodimer with Hth and is necessary for nuclear localization of Exd[187]. An α-Exd antibody detects protein expression in both L3 and adult IPCs (figure 14 A - B). We therefore hypothesized that Exd, together with its co-factor Hth, regulates *dilp* expression in the IPCs. IPC-specific knockdown of *exd* strongly reduced expression of *dilp2 and -3* while significantly increasing *dilp5* mRNA levels (Figure 14 D, F, H). Dilp protein levels were also reduced in the adult stage (Figure 14 C-G'). We then asked whether the known dependence of Exd on Hth for its nuclear localization is also at play in the IPCs. RNAi experiments in the IPCs showed that the nuclear localization of Exd in the IPCs is dependent on normal Hth expression in the IPCs in both larval and adult IPCs but the converse is not true (Figure S 4 A-D). Hth and Exd are thus equally required for normal *dilp* expression and Dilp protein levels.

Hth and Exd both regulate adult systemic insulin signaling but Exd is dispensable for larval stage-associated phenotypes.

IIS and *dilp2* in particular play key roles in the regulation of growth and development during the larval stage[89,90]. We therefore hypothesized that IPC-specific knockdown of *hth* and *exd* knockdown would result in decreased adult size[89,90]. *hth* but not *exd* knockdown results in decreased adult weight, head width and wing area (Figure 15 A-C). Developmental timing, as measured by time until eclosion, is also increased only upon *hth* depletion (Figure 15 D). Adult fat content was increased in *hth* knockdown only (Figure 15 E). Oxidative stress resistance, induced by the addition of paraquat in the food, was increased in both *exd* and *hth* knockdown (Figure 15 F). Lastly, female fecundity is reduced in both *exd* and *hth* knockdown flies (Figure 15 G). Based on the reduction in fecundity and the increased lifespan under oxidative stress conditions we conclude that Hth and Exd are essential for normal systemic insulin signaling in adults.

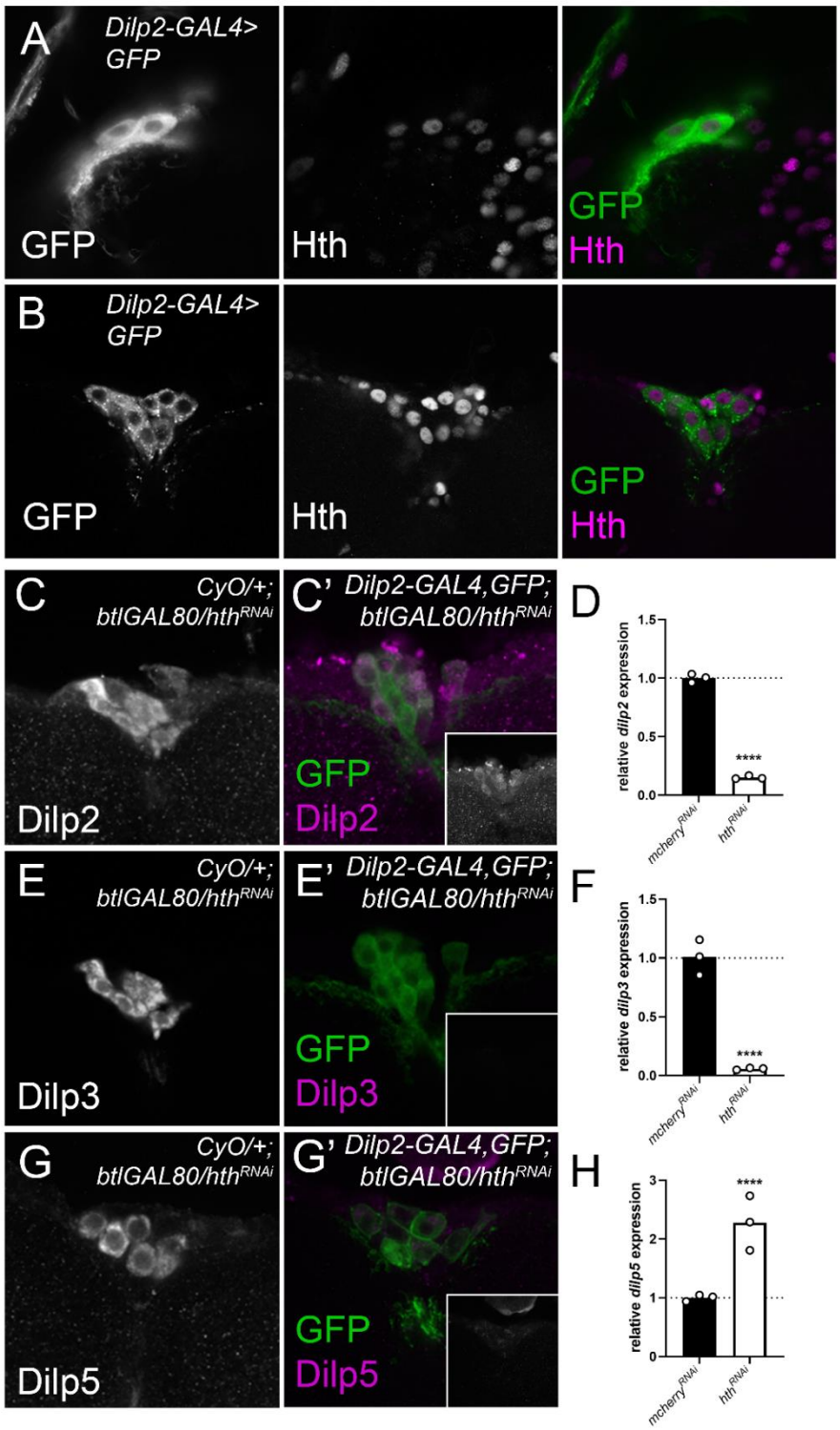


Figure 13: Hth is expressed in larval and adult IPCs and is necessary for normal Dilp protein levels. Antibody staining detects Hth in L3 (A) and adult (B) nuclei in IPCs that are marked by CD8-GFP expression. Knockdown of *hth* reduces Dilp2 (C-C'), Dilp3 (E-E') and Dilp5 (G-G') protein expression in the IPCs compared to balanced age-controlled siblings (left) reared in identical environmental conditions. Insets (C', E', G') show Dilp2, -3 and -5 only (grayscale). Brains from knockdown and balanced siblings were dissected and stained in the same tube. Expression of CD8-GFP was used to distinguish control and test genotypes, and to label the IPCs. Images were acquired with identical confocal settings and representative examples are shown. Expression of *dilp2* (D) and *dilp3* (F) is significantly reduced and *dilp5* (H) is significantly increased upon Hth depletion in the IPCs (*dilp2*, *-3* and *-5* qPCR results are added to facilitate comparison with Dilp protein results, and correspond to the Hth-related results shown in Figure 11 C).

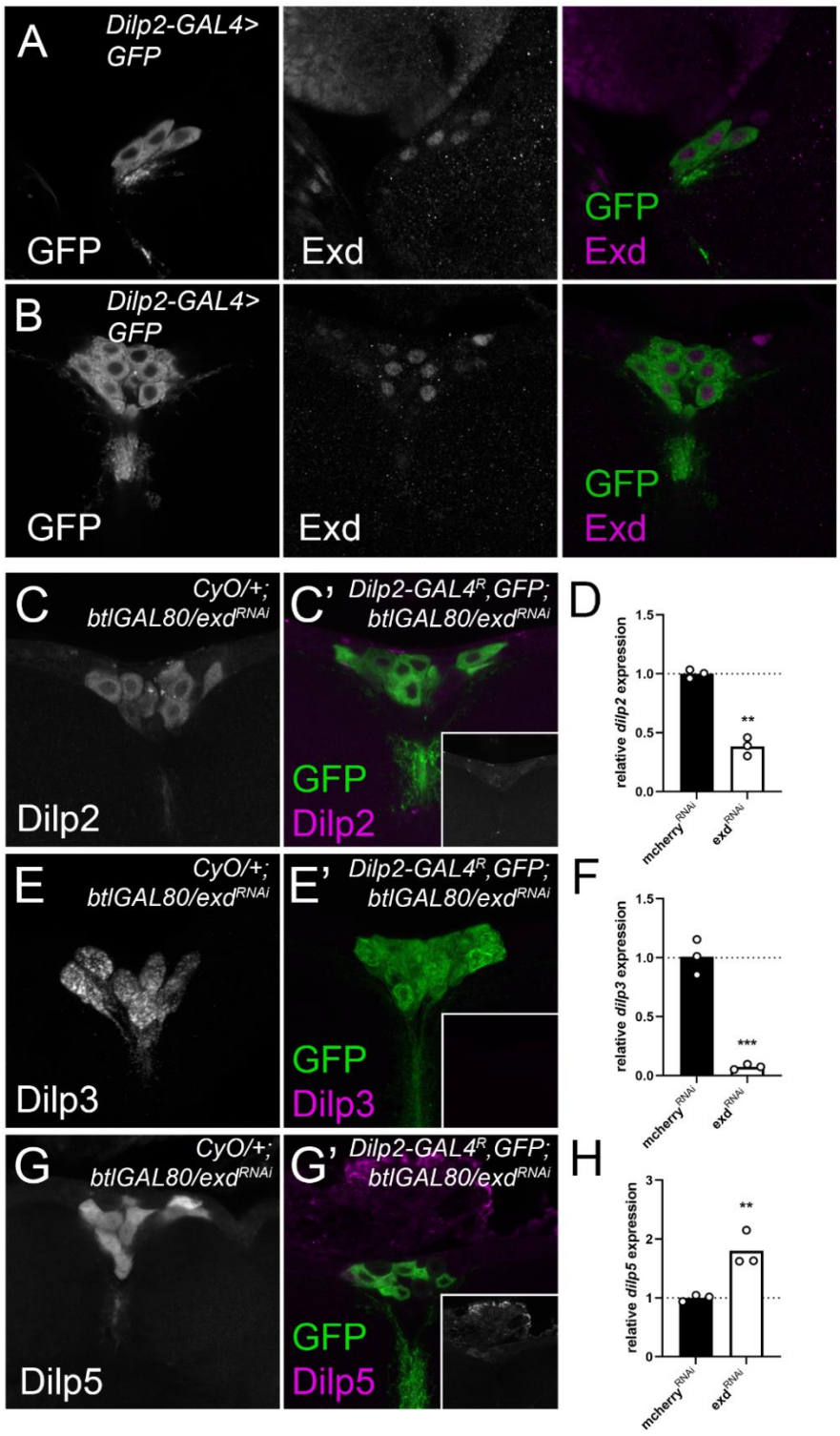


Figure 14: The Hth interactor Exd, is expressed in the IPCs and is equally required for *dilp* expression. (A-B) Antibodies against Exd detect expression in L3 and adult IPC nuclei, marked by *dilp2-GAL4*^{R} driven *UAS-CD8GFP expression*. (C-C') Dilp2, (E-E') Dilp3 and Dilp5 (G-G') protein levels are reduced in adult IPCs upon Exd depletion in the IPCs compared to balanced age-controlled siblings. Insets (C', E', G') show Dilp2, -3 and -5 only (grayscale). Brains from knockdown and balanced siblings were dissected and stained in the same tube. Expression of CD8-GFP was used to distinguish control and test genotypes, and to label the IPCs. Knockdown of *exd* in the IPCs reduced expression of (D) *dilp2* and (E) *dilp3* significantly while *dilp5* is upregulated. Flies were reared at controlled densities. RNA was collected from 10-15 female fly heads. After normalization, *dilp* expression was normalized to a control *dilp2-GAL4R; btlGAL80* driving *UAS-mCherry*^{RNAi}.

The size and developmental timing phenotypes, on the other hand, show that Exd is dispensable for normal larval systemic insulin signaling while Hth is necessary during the larval stage. We next wondered if overexpressing Hth and Exd in the IPCs is sufficient to increase *dilp* expression. No significant effects on *dilp2*, *-3* and *-5* mRNA levels were observed (Figure S5 A). Overexpressing Hth or Exd also left adult weight unaffected (Figure S5 B). We next checked if Hth and Exd are sensitive to different nutritional conditions or if they function independently of nutritional status. Flies were placed on starvation, sucrose only or protein rich media for 72h and the presence of Hth/Exd in IPC nuclei was assessed (Figure S6 A-C). This revealed that there were no appreciable differences in nuclear localization of Hth/Exd. Moreover, flies depleted of *hth* or *exd* in the IPCs in these three different nutritional conditions show similar reductions in *dilp2* and *-3* mRNA levels and increased *dilp5* levels as observed for flies on control food (Figure S6 D-G). This shows that Hth-Exd play a role in the constitutive regulation of *dilp* mRNA levels and are not influenced by the nutritional status of the fly. Furthermore, the presence of multiple Hth binding sites in promotor-enhancer regions of *dilp2*, *-3* and *-5* genes suggests that they may be directly regulated by Hth (Figure S7 A-C).

Hth and Exd have different effects on dilp regulation in larvae.

To address the different effects of Hth and Exd adult size, we looked at dilp levels in the larval stage. This revealed that already at the late L2 stage (68h After Egg Laying (AEL)) Dilp2 protein levels are reduced by *hth* depletion but not upon *exd* depletion (Figure S8 A-B). Later during development, at mid L3 (92h AEL), Dilp2, -3 and -5 protein levels are reduced in both *hth* and *exd* knockdown (Figure S9A -F). Later in larval development at 100h AEL, prior to wandering, immunohistochemistry revealed that Dilp2 protein levels are reduced in both Hth and Exd depletion (Figure 16 A, C), while Dilp5 levels are unaffected by e*xd*^{RNAi} but reduced by *hth*^{RNAi}. Furthermore, mRNA levels of *dilp2* and *-5* are reduced only for *Hth*^{RNAi} but not *Exd*^{RNAi} at 100h AEL (Figure 16 E-F). Dilp2, the main regulator of growth, is thus severely reduced at all checked timepoints upon Hth depletion while the effect of Exd depletion on Dilp2 is more subtle where Dilp2 protein levels are reduced at late L3, but mRNA levels are not significantly decreased which is suggestive of increased Dilp2 secretion. In conclusion, we show that Hth is necessary for Dilp regulation from at least late L2 onwards while the larval role of Exd is more complex.

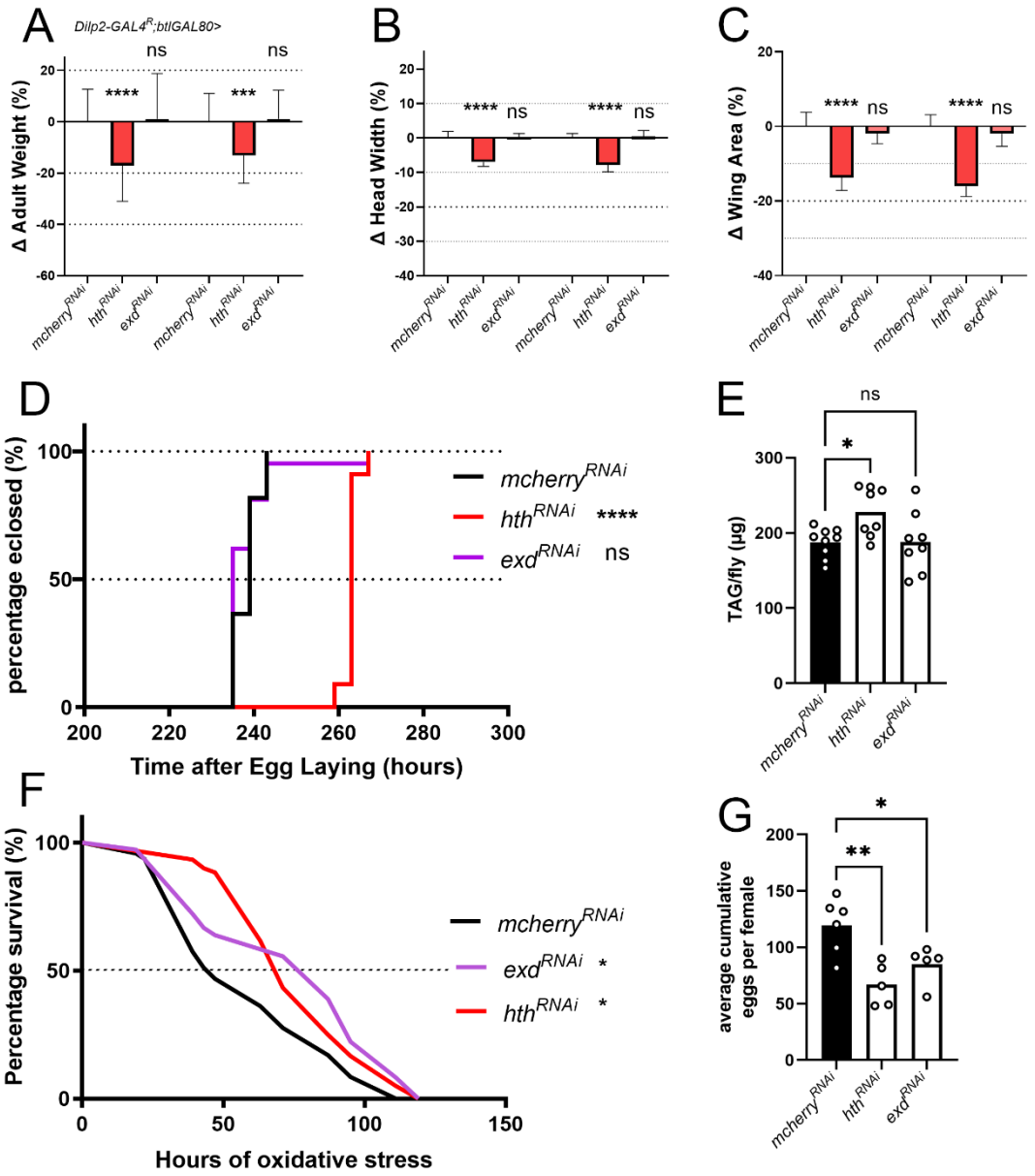


Figure 15: Depletion of Hth and Exd in the IPCs results in phenotypes associated with reduced systemic insulin signaling. (A-C) Knockdown of *hth* but not of *exd* results in decreased adult weight, head width and wing area (one-way ANOVA with post-hoc Dunnett's test comparing to *mCherry*^{RNAi} controls; ns, not significant, *** p<0.001, **** p < 0.0001). Embryos were collected on yeasted apple-juice plates for 3 hours and allowed to develop to early L1 larvae when they were transferred to food vials at a fixed density (20 larvae per vial). Flies were anaesthetized with chloroform 4-5 days after eclosion for measurements. (D) Knockdown of *hth* in the IPCs but not of *hth* results in a developmental delay (one-way ANOVA with post-hoc Dunnett's test comparing to *mCherry*^{RNAi} controls; ns, not significant **** p < 0.0001). (E) Total TAG of female flies is increased upon Hth knockdown but not in case of Exd knockdown (one-way ANOVA with post-hoc Dunnett's test comparing to *mCherry*^{RNAi} controls; ns, not significant * p < 0.05). (F) Oxidative stress resistance is increased in flies expressing *hth*^{RNAi} and *exd*^{RNAi}. Flies were reared on normal food and aged to 3 days old and placed at controlled densities on food containing 20mM paraquat. Dead flies were scored regularly (* p <0.05,** p<0.01, ns, not significant, log-rank test with Bonferroni correction for multiple testing). (G) IPC-specific depletion of hth or exd resulted in decreased female fecundity. Every data point is the average (4-5 females per vial) cumulative number of eggs layed for 14 days (one-way ANOVA with post-hoc Dunnett's test comparing to *mCherry*^{RNAi} controls * p <0.05,** p<0.01).

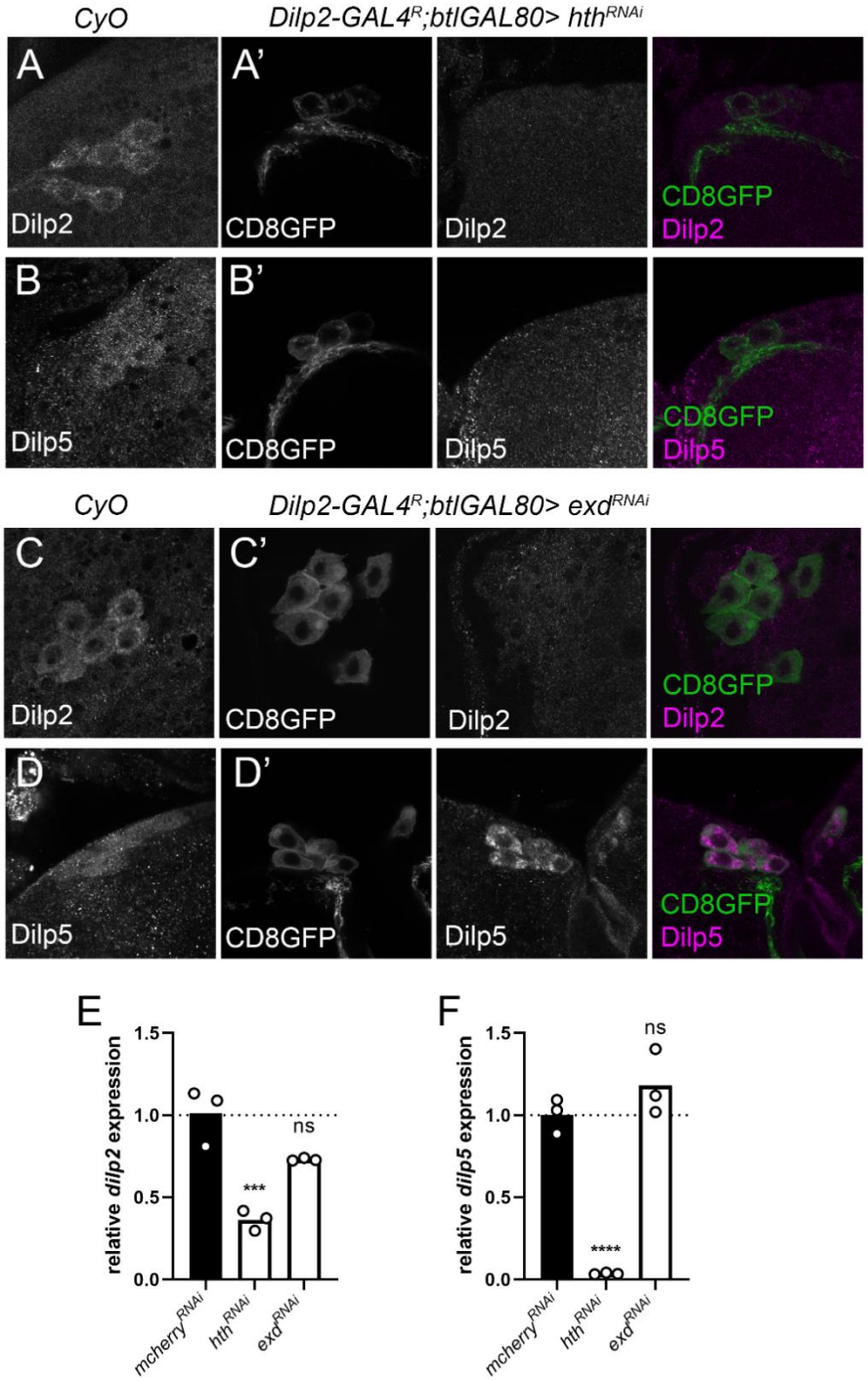


Figure 16: Hth and Exd have differing effects on Dilp protein and mRNA levels in late L3. (A-A', B-B') representative images showing reduced Dilp2 and Dilp5 protein levels in late L3 (100h after egg laying) larvae genetically depleted of *hth* compared to balanced sibling controls (CyO, left panels). (C-C', D-D') IPC-specific *exd* knockdown results in decreased Dilp2 and unaltered Dilp5 protein levels in IPC somata. (E) *dilp2* mRNA levels are significantly reduced upon *hth* depletion and not significantly reduced upon *exd* depletion. (F) *dilp5* mRNA levels are severely reduced *hth* knockdown larvae and are unaltered in *exd* knockdown. (ns, not significant, ***p<0.001, ****p<0.0001), one-way ANOVA with post-hoc Dunnett's test to correct for multiple comparisons).

Retinal Determination Gene Network is active in the IPCs to regulate *dilp* expression

In the mammalian pancreas, the Hth/Exd homologs Meis1/PBX1/2 function upstream of PAX6/Dach1/2 where they are required to induce *pax6* expression[188,189]. We thus hypothesized that Hth would have a similar regulatory relationship with Ey in the IPCs. Depletion of *hth* in the IPCs did not result in differences in Ey levels in the IPCs, while Exd expression was no longer detectable in the IPC nuclei, showing that Hth is required for normal Exd localization and/or expression, but not for normal Ey expression (Figure S3 and S4).

Hth/Exd are functional in the retinal determination gene network (RDGN), necessary for eye development[185]. We and others had already described a role for other RDGN members: Eyeless (Ey) and Dachshund (Dac) in regulating the IPCs. Therefore, we hypothesized that a version of this regulatory network controls *dilp* expression in the IPCs[149,150] To test this, we first queried the transcriptome of both larval and adult IPCs (Cao et al[103], and this study) for expression of 11 different RDGN members. Of these 11 TF, all were expressed to some degree in the adult IPCs, with *hth*, *gro*, *optix*, *dac*, *eya*, *ey* and *dan* showing expression in both stages. (Figure 17 A). These results were validated with antibody staining, transgenic GAL4 reporter lines and tagged versions where available (Figure 17 B - W). Only Eya showed no detectable expression in the larval or adult IPCs but the overall staining quality for Eya was poor (Figure 17 P -Q) and while RNA-seq results for the larval stage suggested that only a subset of RDGN TF were expressed in L3 IPCs, all but Eya are clearly detectable at the protein level using immunohistochemical reagents while Eyg, Toe, Eya, Dan, and So are not detected in the adult stage (Figure 17 M, O, Q, U, W). This shows that most RDGN transcription factors are expressed in the IPCs. To establish a role for these transcription factors in the IPCs, we genetically depleted these TFs in the IPCs using *Dilp2-GAL4$^{215\text{-}1\text{-}1\text{-}1}$* and their respective RNAi lines[190]. Except for *dac* and *toe*, knockdown of these TFs resulted in a significant reduction in *dilp2* and *dilp5* expression compared to D*ilp2-GAL4$^{215\text{-}1\text{-}1\text{-}1}$* flies driving a control RNAi under identical environmental conditions (Figure 17 X - Z). Knockdown of *dac* reduced *dilp5*, and increased *dilp2* levels, while knockdown of *toe* only reduced *dilp5* expression. Furthermore, only the depletion of *tsh* and *ey* affected *dilp3* levels (Figure 17 S). Since reduced *dilp2* and -*5* levels could be due to a reduction in the number of IPCs, we verified this and found that the number of IPCs are unaltered (Figure S10). Overall, these results suggest that the RDGN is re-used in the regulation of IIS in the IPCs.

In summary, these data show that a repurposed RDGN is used in the IPCs. Knockdown of *hth* and *exd* affects all three *dilps*. On the other hand, most RDGN TFs regulate either *dilp2* and *dilp5*, while *toe* regulates only *dilp5* and *tsh* regulates all three. Our data provide an interesting

example of how regulatory networks are re-used in other developmental contexts to achieve different transcriptional outcomes.

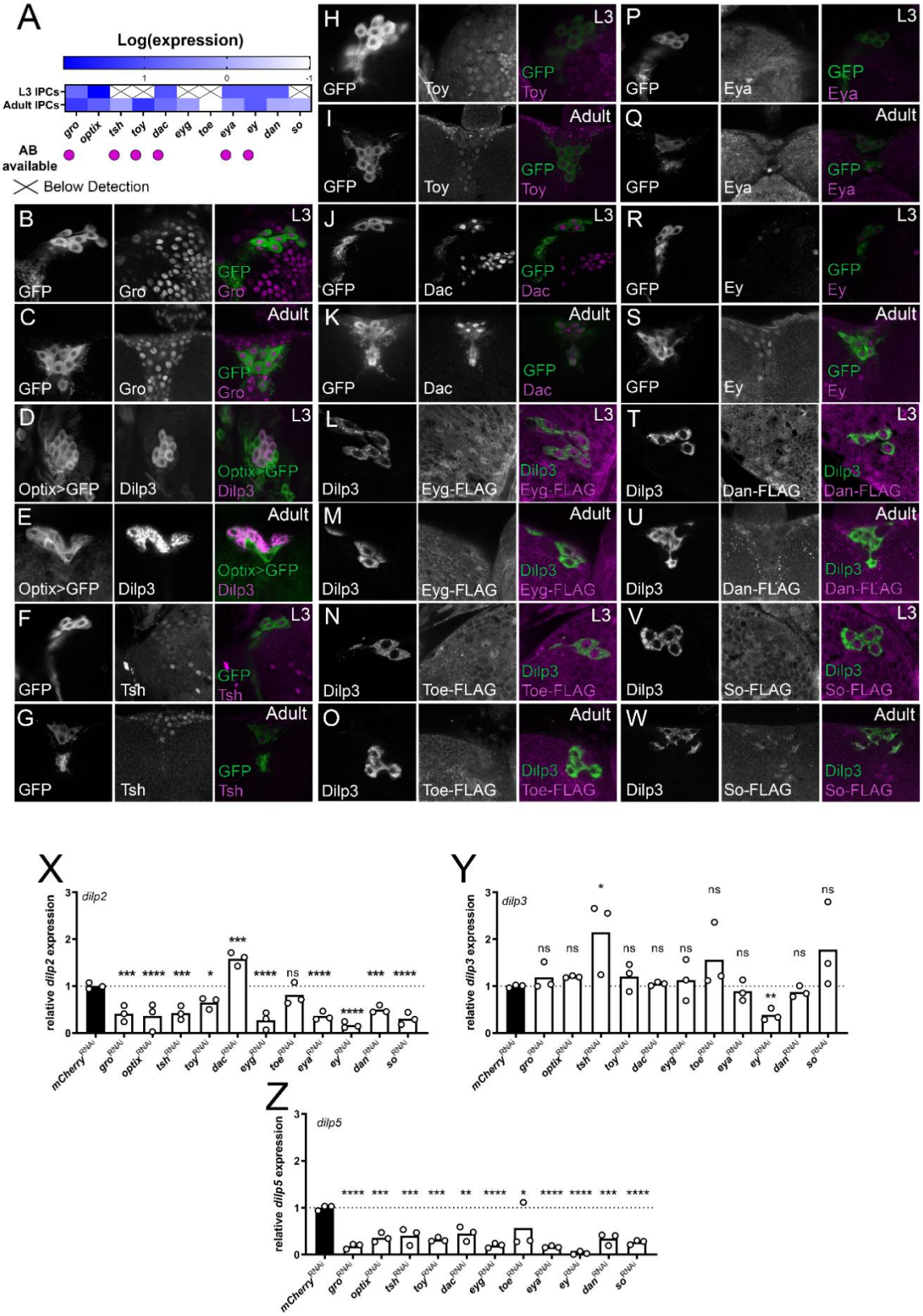


Figure 17: Most RDGN transcription factors are expressed in the IPCs and are required for normal adult *dilp* expression. (A). Additional RDGN transcription factors are expressed in larval and/or adult IPCs (Larval transcriptome data is taken from Cao et al., (2014)[103] while adult transcriptome data was generated for this study. (B-W) Gro, Tsh, Toy, Dac and Ey are readily detected in L3 and adult IPCs, only Eyg exhibited no detectable protein expression in larval or adult IPCs. Adult *Dilp2-GAL4^{R} > CD8GFP* flies were reared to sexual maturity and brains were dissected and stained with available ABs directed against RDGN TFs. Expression of CD8GFP was used to label IPCs. For Optix, an Optix-GAL4 driver drove expression of the CD8GFP and was localized with an antibody against Dilp3, which labelled the IPCs. For Eyg, Toe, Dan and So, flies carrying a FLAG-GFP tagged version of these TFs were stained with an anti-FLAG antibody and anti-Dilp3 (X -Z) Genetic knockdown of all RDGN TFs in the IPCs using *Dilp2-GAL4$^{215\text{-}1\text{-}1\text{-}1}$* except *dac* and *toe* resulted in significant reduction of *dilp2* expression (one-way ANOVA with post-hoc Dunnett's test comparing to *mCherry^{RNAi}* controls, n = 3), *tsh* depletion in the IPCs increases dilp3 expression (one-way ANOVA with post-hoc Dunnett's test comparing to *mCherry^{RNAi}* controls, n = 3) and all were required for dilp5 expression (one-way ANOVA with post-hoc Dunnett's test comparing to *mCherry^{RNAi}* controls, n = 3). Flies were reared at controlled densities and RNA was collected from 10 female fly heads of flies controlled for age and diet.

Discussion

The IPCs are key regulators of growth and metabolism in the fly but the regulatory networks governing their development and function remain poorly understood. In this study, we use the variation in *dilp* expression across DGRP lines to identify genes that are necessary for normal IPC-autonomous *dilp* regulation. There was very little overlap between the sexes for the identified candidate genes (3 genes for *dilp2*, 2 genes for *dilp3* and none for *dilp5*). Furthermore, only one gene was shared between the dilps (*Glucalpha* for *dilp3* and *dilp5* in females). The results from our RNAi based approach to validate the role of these putative regulators revealed that depletion of most genes had sexually dimorphic effects, yet also that the effects on *dilp* expression between the sexes were correlated (Figure 11 C, Figure 12 A-C'). Additionally, genes implicated in the regulation of one *dilp* often also had effects on the other dilps. Variants in Hth, for example, were identified to correlate with *dilp5* levels in males but not females. Depletion of Hth in the IPCs produced a pronounced effect on *dilp5* levels and additionally resulted in a strong reduction of both *dilp2* and *dilp3* levels in both sexes (Figure 11 C). Several possible explanations for this apparent discrepancy exist. One explanation is that our RNAi-based approach might have obscured the effect of the specific variants which might have sex-specific and/or *dilp*-specific effects. The compensatory effects previously noted in *dilp* regulation might additionally add another layer of complexity onto our results[89,120]. Furthermore, the identified genes might play an additional role in the non-autonomous regulation of *dilp* expression by acting in other tissues to regulate *dilp* expression[178]. Lastly, it is possible that identified genes are associated with multiple *dilps* and in both sexes but that they did not reach the significance threshold due to insufficient statistical power.

Our approach revealed that Hth is required in the adult IPCs to regulate *dilp* expression (Figure 11 C). IPC-specifc depletion of Hth resulted in almost complete absence of *dilp2* and *dilp3* mRNA in both sexes (~10% of control) while doubling *dilp5* mRNA levels. Immunohistochemistry confirmed the strong reduction of Dilp2 and Dilp3 on the protein level but an increase on Dilp5 protein levels was not seen. The latter could be due to increased Dilp5 secretion or a failure to produce properly matured Dilp5 protein (Figure 13 C-E). The increase in *dilp5* mRNA levels is reminiscent of what is seen in *Δdilp2-3* deletion mutants which show upregulated *dilp5* mRNA levels, presumably to partially compensate for loss of *dilp2* and *-3*[89]*.*

Further phenotypic characterization revealed that Hth but not Exd is essential for both normal development and growth at the larval stage (Figure 15 A-D). As both larval growth and developmental timing are disturbed upon *hth* depletion, this suggests that IIS levels are reduced before and after attainment of critical weight[191,192]. Exd seems dispensable for both growth and developmental timing although IPC Dilp2, -3 and 5 protein levels are reduced at mid L3. By late L3, however, *dilp2* and-*5* mRNA levels are not significantly altered (Figure 16 E-F) and Dilp2 protein levels are reduced in the IPCs (Figure 16 D). This suggests that Dilp2 secretion might be increased thereby minimizing growth defects. Regulation of *dilp2*, *-3* and *–5* in larvae by Hth thus seems to be partially independent of Exd and depends on the specific timepoint during larval development. During larval development several peaks in the titre of the steroid hormone, Ecdysone, mark the transitions between larval stages and more subtly in L3, marking critical weight and the transition towards wandering[192].Recent work from our lab showed that the larval IPCs are Ecdysone responsive[123]. This raises the possibility that the interaction between Hth-Exd might itself be Ecdysone sensitive, explaining their effects in larval *dilp* regulation.

Further analysis revealed that Hth and Exd had similar effects on regulating *dilp* expression in adults. Depleting Exd and Hth in the IPCs decreased female fecundity and increased oxidative stress resistance consistent with a reduction in systemic insulin signaling (Figure 15 F-G). By contrast, total TAG in adult flies is increased only when *hth* is knocked down. Recent work showed that *dilp2* knockdown in the IPCs in the larval stage only is sufficient to increase total TAG levels in adult flies[193]. Consistently, we find that Exd knockdown does not significantly alter *dilp2* mRNA levels in late larvae (Figure 16 E)

In the mammalian pancreas the fly homolog of Ey, Pax6, is required to establish the cellular fate of endocrine cell types and to promote expression of their respective hormone gene products. To do this, Pax6 cooperates with Dach1/2, which is also observed in the Ey/Dac regulation of *dilp5* in *Drosophila*[150]. Upstream of Pax6, the Meis gene family members

Meis1/2/Prep1/2 and Pbx1/2, which are homologs of *Drosophila* Hth and Exd, are required to promote pancreatic *pax6* expression[188]. Depletion of Hth in the IPCs, however, leaves Ey expression unaffected while resulting in the absence of Exd in IPC nuclei (Figure S5 and S7). This shows that while the TFs are identical between flies and humans, their regulatory relationships are not. Similarly, differences in regulatory relationships are also seen during eye development where Meis1/2 regulate *pax6* in the mouse lens. In *Drosophila*, however, Hth has no direct regulatory relationship with *ey* or its paralog *toy*[194,195]. Key TFs are thus conserved in *Drosophila/*mammalian eye development and IPC/pancreas development while the regulatory relationships between the TFs have diverged during evolution.

We further show that the genes of the RDGN are active in the IPCs to regulate insulin signaling, mostly via control of *dilp2* and *dilp5* expression. Of all tested TFs only Ey and Hth/Exd are necessary for promoting *dilp3* expression while Tsh depletion resulted in an upregulation of *dilp3* levels (Figure 17 Y). Other RDGN TFs were not required to regulate *dilp3*, and most regulated *dilp2* and *dilp5*. In literature, co-regulation between *dilp2* and *-5* is commonly reported, despite them having distinct functions[127]. However, given the complex regulatory relationships between the RDGN TFs in the developing eye, it is surprising to see that, in general, their regulatory relationship is largely positive, promoting expression of these two *dilps* (Figure 18 A-B). One explanation is that since a couple of the RDGN TFs are already known to be expressed at the neuroblast stage[97], depleting these TFs during development might result in a failure to properly establish IPC cell fate. Additionally, a more complex regulatory architecture may exist that is relevant in other physiological contexts to which the *dilps* are sensitive, such as in changing dietary conditions, aging, or oxidative stress[196]. Of note are the opposing effects of Dac and Ey on *dilp2* expression, while still physically interacting to promote *dilp5*. This implies the presence of *dilp*-specific regulatory relationships. We do, however, not provide evidence for the direct regulation of the *dilps*, but we hypothesize, based on the strong effects of Hth/Exd on *dilp* expression and on the presence of Hth binding sites within *dilp2*, *-3* and *-5* enhancers that they may directly regulate the *dilps*, particularly *dilp2* and *dilp3*, while others may contribute via indirect interaction or regulation of additional factors (Figure 18 B).

Interestingly, a study by Okamoto *et al.*[150], which originally identified Dac as a co-regulator of *dilp5* together with Ey, also knocked down a number of other RDGN TFs (*optix*, *eya*, *eyg*, *tio*, *so*, *toy*). They find small and often not significant effects on *dilp* levels when compared to our results. We speculate that differences between our results and theirs lay in the age of the flies, the different *dilp2*-GAL4 line that was used or the differences in dietary conditions. Another study based on type 2 diabetes risk loci in humans by Peiris *et al.*[197], showed that IPC-specific knockdown of *Optix* resulted in decreased circulating Dilp2 levels.

Our results suggest that this reduction in Dilp2 release might be indirect due to the reduction of *dilp2* mRNA levels upon *Optix* depletion.

To conclude, our approach identified many novel regulators of *dilp* expression with a prominent role for Hth/Exd in regulating *dilp* expression in both sexes. In addition, we show that the RDGN is repurposed to regulate the development and function of the IPCs. The RDGN is thus a paradigm of how evolution uses the finite genetic tools available to establish the structural and functional diversity observed in biology.

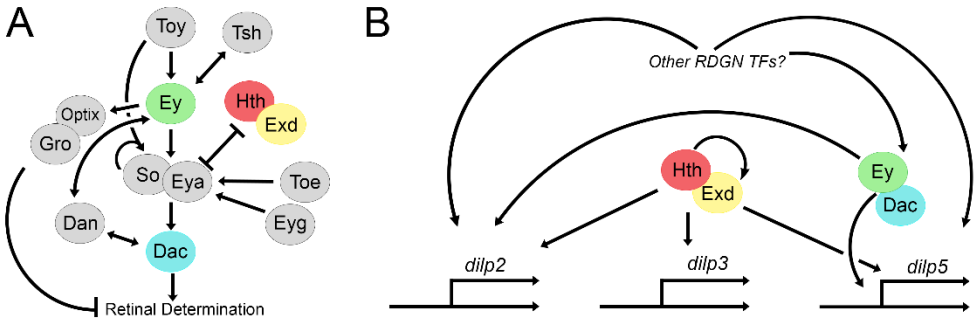


Figure 18: A reshuffled RDGN regulates *dilp* expression in the IPCs. (A) Schematic showing the regulatory relationships of the RDGN in the developing eye (modified from Kumar[185] and Bürgy-Roukala *et al.*[198]). (B) Speculative model of the RDGN in *dilp* regulation. Hth and Exd as a pair directly or indirectly regulate *dilp2*, *-3 and -5* expression, while Ey and Dac directly regulate *dilp5* epxression[150] and via other factors *dilp2* and *dilp3*. Other RDGN TFs directly or indirectly regulate *dilp2* and *dilp5* expression.

Methods

Fly Stocks and Transgenes used in this Study

Drosophila stocks were reared at 25°C on Nutrifly (Genesee Scientific) for the initial DGRP validation screen. Subsequent experiments were performed using standard yeast fly food recipe (See Table 1, below). Stocks were ordered from the Bloomington *Drosophila* Stock Center (BDSC), Vienna *Drosophila* RNAi Center (VDRC) or from the Zurich Orfeome Project (FlyORF). The following GAL4 lines were used: *Dilp2-GAL4R; btl-GAL80* (*BDSC# 37516*, described in[90]), *OK107-GAL4* [199]and *Dilp2-GAL4*$^{215\text{-}1\text{-}1\text{-}1}$ [130].To visualize IPCs, *Dilp2-GAL4* lines controlled expression of *UAS-CD8GFP (BDSC# 5137)* or *UAS-nlsGFP (BDSC# 4776)*. A list of the RNAi lines used in the DGRP screen are included in supplemental file 1.

RNAi lines for the RDGN follow-up experiments include *hth-RNAi (BDSC# 34637)*, *ey-RNAi (BDSC# 32486)*, *dac-RNAi(BDSC# 35022)*, *exd-RNAi #1 (BDSC# 34897)*, *tsh-RNAi (BDSC# 35030)*, *optix-RNAi (BDSC# 55306)*, *toy-RNAi (BDSC# 29346)*, *toe-RNAi (BDSC# 50660)*, *dan-RNAi (BDSC# 51501)*, *eyg-RNAi (BDSC# 26226)*, *gro-RNAi (BDSC# 35759)*, *so-RNAi*

(BDSC# 35028) and *eya-RNAi (BDSC# 35725)*. Overexpression lines include *UAS-Hth* (F000241) and *UAS-Exd* (F004513).

Flyfood recipe

Ingredient	Amount
Water	15.2 L
Cornmeal	1000 g
Yeast	180 g
Agar	112.5 g
Dextrose	270 g
Molasses	600 mL
Propionic Acid	150 mL
10%Hydroxybenzoate	22.5 g
99.8% Ethanol	225 mL

Candidate gene identification

GWA analyses on *dilp* values were performed using the DGRP2 webtool (http://dgrp2.gnets.ncsu.edu/) [179,200]. Mean FPKM values of *dilp2*, *-3* and *-5* were used as the input and genes associated with variants that reached a nominal significance threshold of p < 10^-3 in the mixed effects model in one of the sexes were considered for further validation. Input files are provided in supplemental file 2.

Transcriptome data

IPCs were fluorescently labelled with *Dilp2-GAL4; UAS-mCD8GFP*. GFP-positive cells were manually sorted according to Nagoshi et al[201]. In short, 100 brains were dissected from adult flies, 3-5 days after eclosion into ice-cold dissecting solution (9.9 mM HEPES-KOH buffer, 137 mM NaCl, 5.4 mM KCl, 0.17 mM NaH2PO4, 0.22 mM KH2PO4, 3.3 mM glucose, 43.8 mM sucrose, pH 7.4) containing 50 µM d(−)-2-amino-5-phosphonovaleric acid (AP5), 20 µM 6,7-dinitroquinoxaline-2,3-dione (DNQX), 0.1 µM tetrodotoxin (TTX), and immediately transferred into modified SMactive medium (5 mM Bis-Tris, 50 µM AP5, 20 µM DNQX, 0.1 µM TTX) on ice. Brains were digested with l-cysteine-activated papain (50 units ml−1 in dissecting saline) for 20 min at 25 °C. Digestion was quenched with a fivefold volume of the medium, and brains

were washed twice with the chilled medium. Brains were triturated with a flame-rounded 1,000-µl pipette tip with filter followed by a flame-rounded 200-µl pipette tip. Single, GFP-positive cells were selected, then reselected a further 2 times until approximately 100 fluorescent cells were pooled. Total RNA was extracted using a Nucleospin RNA micro kit (Macherey-Nagel). cDNA synthesis and adapter ligation were performed with the SMARTer cDNA synthesis kit (Takara Bio). 51bp paired-end reads were generated on an Illumina HiSeq2000 sequencer. Resulting reads were submitted to the European Nucleotide Archive, Accession Number PRJEB50165.

Quality assessment and read trimming was performed with fastp (v. 0.23.1)[202] with default settings. Reads were then pseudo-aligned to the reference *Drosophila* genome (BDGP6.32) and TPM values determined using salmon (v. 1.5.0)[203] and summarized to the gene level using tximport (v. 1.18.0)[204]

RNA extraction and qPCR

For the validation of candidates from the DGRP, offspring was collected after eclosion and maintained at controlled densities of 20 total flies, aged to 3-5 days and snap frozen on dry ice. 10 to 15 fly heads were used per replicate for RNA extraction. For subsequent qPCR experiments concerning the RDGN transcription factors, flies were aged until 12-15 days old and only female flies were used for qPCR experiments. For larval qPCR, staged larvae (see below) were collected and all organs except the brain were dissected out. Each replicate is 5-6 larvae.

Total RNA was isolated using TRI™ reagent (Invitrogen) and reverse transcription was performed on 1µg RNA using Transcriptor First Strand Synthesis kit (Roche). Primer sequences used are listed in Table S1. qPCR was performed on a Step-One-Plus using the SYBR Green detection system. Transcript levels were normalized to *rp49* transcript levels (initial screen) or using the geometric means of *rpS13* & *rp49* (transcription factor experiments). Mean ΔCt values were statistically compared as described below. Relative quantitation of transcript levels to control genotypes were calculated using the ΔΔCt method and plotted with GraphPad Prism software to visualize expression differences.

Primers used in this study

Gene		Sequence (5' → 3')
rps13	Forward	GGG TCT GAA GCC CGA CAT T
	Reverse	GGC GAC GGC CTT CTTT GAT
rp49	Forward	GCC CAA GGG TAT CGA CAA CA
	Reverse	CTT GCG CTT CTT GGA GGA GA
rpl3	Forward	AAG GAT GAC GCC AGC AAG CCA GTC
	Reverse	TAG CCG ACA GCA CCG ACC ACA ATC
dilp2	Forward	GCG AGG AGT ATA ATC CCG TGA T
	Reverse	GGA TTG AGG GCG TCC AGA
dilp3	Forward	GCA ATG ACC AAG AGA ACT TTG GA
	Reverse	GCA GGG AAC GGT CTT CGA
dilp5	Forward	GAG GCA CCT TGG GCC TAT TC
	Reverse	CAT GTG GTG AGA TTC GGA GCT

Immunohistochemistry

Adult and larval brains were dissected in 1x PBS prior to 30 minute fixation with 4% formaldehyde. Primary antibodies were incubated overnight at 4°C in PAXD (PBS containing 1% BSA, 0.3% Triton X-100, 0.3% deoxycholate). Antibodies used in this study were mouse α-GFP (University of Iowa, Developmental Studies Hybridoma Bank (DSHB) 8H11; 1:100), rabbit α-GFP (Life Technologies, Ref#: A6455; 1:1000), mouse anti-FLAG (Sigma, F1804, 1:50), rabbit α-Hth (Dr. Richard Mann, 1:200), mouse α-Ey (DSHB E4H8, 1:5), rabbit α-Exd (Dr. Richard Mann, 1:200), mouse α-Exd (DSHB EXD B11M, 1:5), mouse α-Dac (DSHB mAbdac2-3, 1:20), α-Gro (DSHB anti-Gro, 1:10), rat α-Tsh (Dr. John Bell, 1:25), rat α-Toy (1:50), mouse α-Eya (DSHB eya10H6, 1:100), rat α-Dilp2 (Dr. Pierre Leopold, 1:1000), rabbit α-Dilp3 and rabbit α-Dilp5.

Rabbit antibodies directed against the Dilp3 partial peptide sequence DEVLRYCAAKPRT and against the Dilp5 peptide sequence RRDFRGVVDSCCRKS were generated as a service by Thermo Fischer Scientific Inc. For immunostaining, α-Dilp3 antibodies were used at a dilution of 1:100 and α-Dilp5 antibodies used at a dilution of 1:400. Secondary antibodies used include goat α-mouse FITC & Cy3, goat α-rabbit FITC Cy3 & Cy5, goat α-rat Cy3 (Jackson Immunoresearch), donkey α-rabbit Alexa-594, donkey α-rat Alexa-594, donkey α-mouse Alexa-488 & Alexa-594. All secondary antibodies were used at a dilution of 1:200. Samples were mounted in Vectashield® Mounting Medium (Vector Laboratories Cat #H-1000).

Immunohistochemistry images were taken with an Olympus FluoView FV1000 confocal microscope and processed using ImageJ64 (1.6.0_65; FIJI) [205,206] and Photoshop software.

Oxidative stress resistance with Paraquat

Female flies were collected and aged until 12- 15 days old, then transferred onto medium (1% agar, 10% Sucrose) containing 20mM Paraquat (Sigma 75365-73-0). Deaths were scored regularly and plotted in GraphPad Prism. Median lifespan was compared using a log-rank test with post-hoc Bonferonni correction for multiple testing.

TAG determination

Groups of five 4-5 day old females flies, with ovaries and gut removed, were collected and total TAG was measured using the sulfo-phospho-vanillin method modified from[207]. In brief, flies were homogenised in chloroform:methanol (2:1) and incubated overnight at 37°C with agitation. Samples were then placed in a 69°C heatblock and chloroform:methanol was allowed to evaporate. Once fully evaporated, sulfuric acid (95-98%) is added and samples were placed on the heatblock for 10 minutes. Vanillin reagent (0.6% Vanillin, 34% phosphoric acid) is then added and samples are briefly mixed and incubated in the dark for 15 minutes. Samples are transferred in clear bottom plates and measured at 540 nm on a Synergy H1 spectrophotometer (BioTek). A standard curve made with Triolein was used to determine TAG content in samples.

Developmental staging and time to eclosion

Parental genotypes (female *Dilp2-GAL4^{R};btlGAL80* and male *UAS-hth^{RNAi}*, *UAS-exd^{RNAi}*, or *UAS-mcherry^{RNAi}*) were allowed to lay eggs on yeasted apple juice plates (1% agar, 10% Sucrose, 20% apple juice plates) for two to three hours. Embryos were allowed to develop to early L1 larvae at 25*C, collected and transferred to food vials at a defined larval density (20 larvae per vial). Larvae were then collected at specific timepoints for immunohistochemistry (68h AEL, 92h AEL and 100h AEL) or qPCR. (100h AEL).

Time until eclosion was scored by regularly assessing the number of eclosed flies. Time to eclosion was plotted in GraphPad Prism and average time until eclosion per genotype across vials (n=4) was compared with a One-way ANOVA with post-hoc Dunnett's test.

Determination of Adult Size Phenotypes

Adult flies 4-5 days after eclosion were collected from controlled conditions as described above, anesthetized using chloroform to prevent dessication, and weighed on a Mettler Toledo XS204 scale (Wet Weight; d = 0.1 mg). To measure wing size, a wing of chloroform anesthetized flies was removed and placed on a dissection slide. Wing length (L) was

measured as a straight line from the alula opening to where the L3 vein meets the distal edge and wing width (W) was measured as a straight line starting from where the L5 vein meets the posterior edge to the anterior edge of the wing, perpendicular to the wing length line[208]. Total wing area (A) was then approximated by using the formula for an ellipse where A = π * (L/2) * (W/2). Head width was measured by carefully removing heads from the thoraces of chloroform anaesthetized flies and placed with its posterior side down on a slide. Head width is then measured as a straight line from the furthest point of the left eye from the midline to the furthest point on the right eye. Images of wings and heads were taken with an Olympus and analyzed with Cell^D software

Mean weights between experimental and control flies were compared using a One-way ANOVA with post-hoc Dunnett's test with GraphPad software, and the relative change in adult weight between experimental and control genotypes were plotted.

Female fecundity

Eclosed females were collected and allowed to mate for 48 hours with sibling males in a 4 female to 1 male sex ratio. Females were then collected and placed in groups of 4-5 in separate vials. Flies were transferred onto fresh food daily and the number of eggs was counted for 14 days. Mean number of eggs laid was compared with a One-way ANOVA with post-hoc Dunnett's test

Nutritional challenges

For Hth and Exd nuclear localization experiments *Dilp2-GAL4^{R};btlGAL80>CD8-GFP* flies were reared on normal food and female flies (24-48h) post eclosion were placed on three different diets (Starvation food: 1% agar, 1.2% glacial acetic acid; Sucrose only food: Starvation food supplemented with 15% sucrose; High protein: Starvation food supplemented with 40% yeast paste) for 72 hours followed by IHC. For nutritional dependency experiments *Dilp2-GAL4^{R};btlGAL80> mcherry^{RNAi}*, *Dilp2-GAL4^{R};btlGAL80> hth^{RNAi}*, *Dilp2-GAL4^{R};btlGAL80> exd^{RNAi}* flies were placed on the three different nutritional conditions for 72h followed by qPCR.

Statistics

All statistics were conducted using the GraphPad Prism software. All statistics were performed on raw data after normalization, where applicable (i.e. qPCR). For assays comparing means, unpaired T-tests for pairwise comparisons or one-way analysis of variance (ANOVA) tests for multiple comparisons were used. To test whether variance significantly differed between samples, an F-test was performed for pairwise comparisons and both Brown-Forsythe and Bartlett's Tests performed for multiple comparisons. If the standard deviations between samples were significantly different, a T-test with Welch's correction (Welch's T-test) was

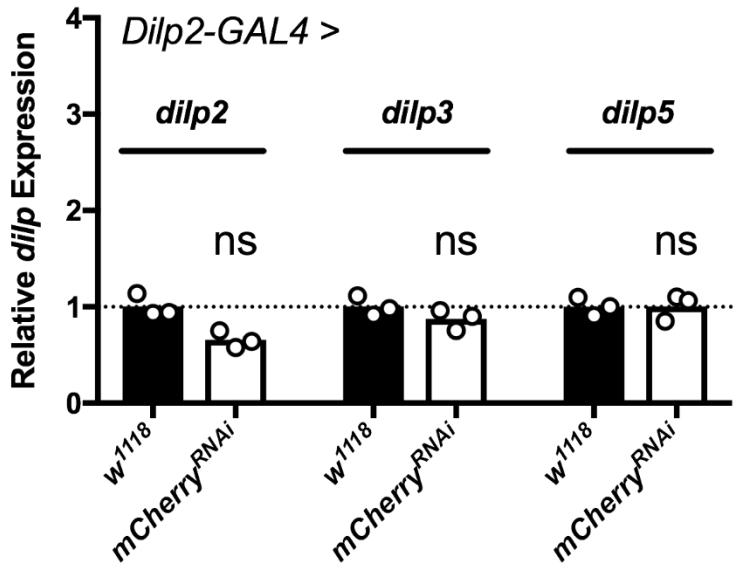
performed for multiple comparisons, whereas the Geisser-Greenhouse correction was always applied on all one-way ANOVA analyses, as sphericity of the data was not assumed. When comparing means to a single control sample, Dunnett's multiple comparisons test was used. For lifespan analysis under oxidative stress, median lifespan was compared with a log-rank test, with Bonferroni correction for multiple testing. For all statistical analyses, we assumed a significance level of 0.05.

Hth motif identification

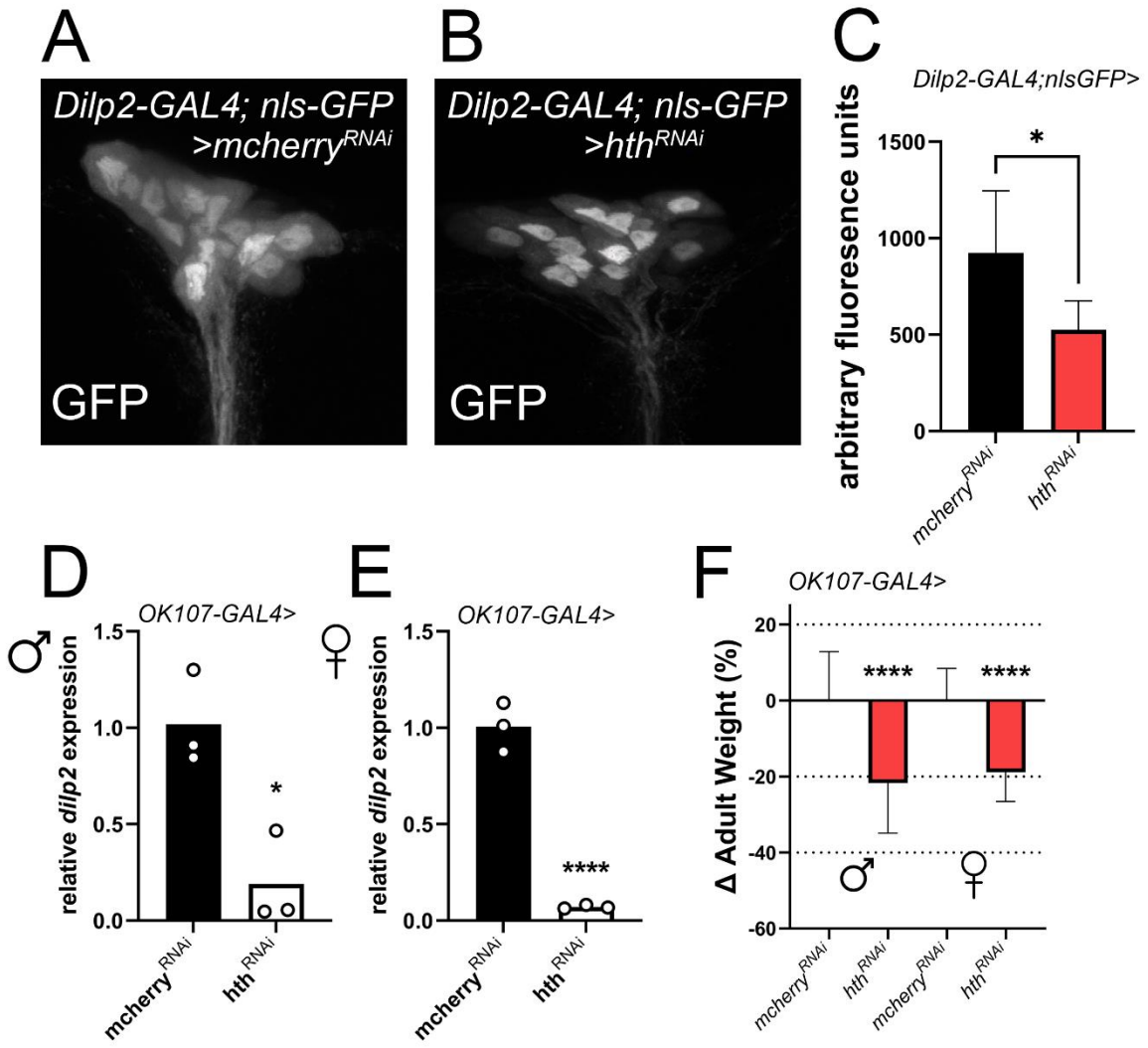
The Hth binding motif was identified in the JASPAR database[209] and FIMO[210] (Find Individual Motif Occurrences) from the MEME suite and was used to scan *dilp2*, *-3* and *-5* enhancers for putative Hth binding sites ($p<0.001$, both strands were scanned).

Supplementary information

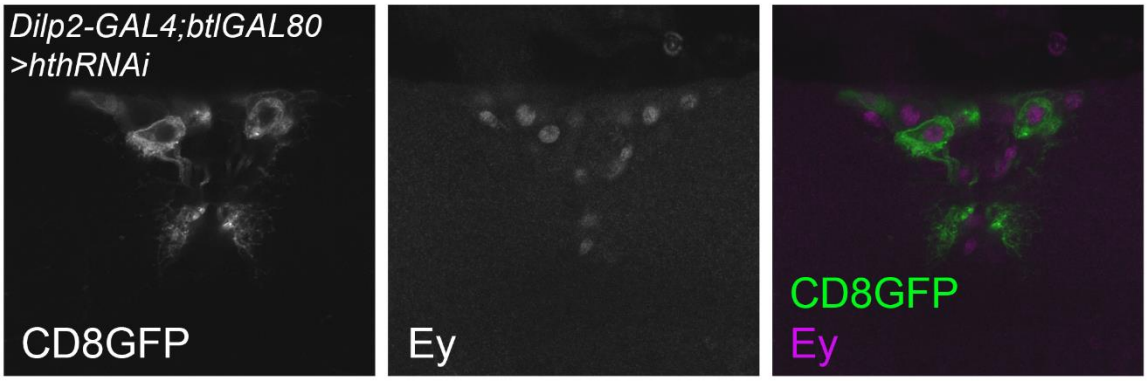
Supplementary tables can be found online at (https://doi.org/10.1371/journal.pgen.1010380).



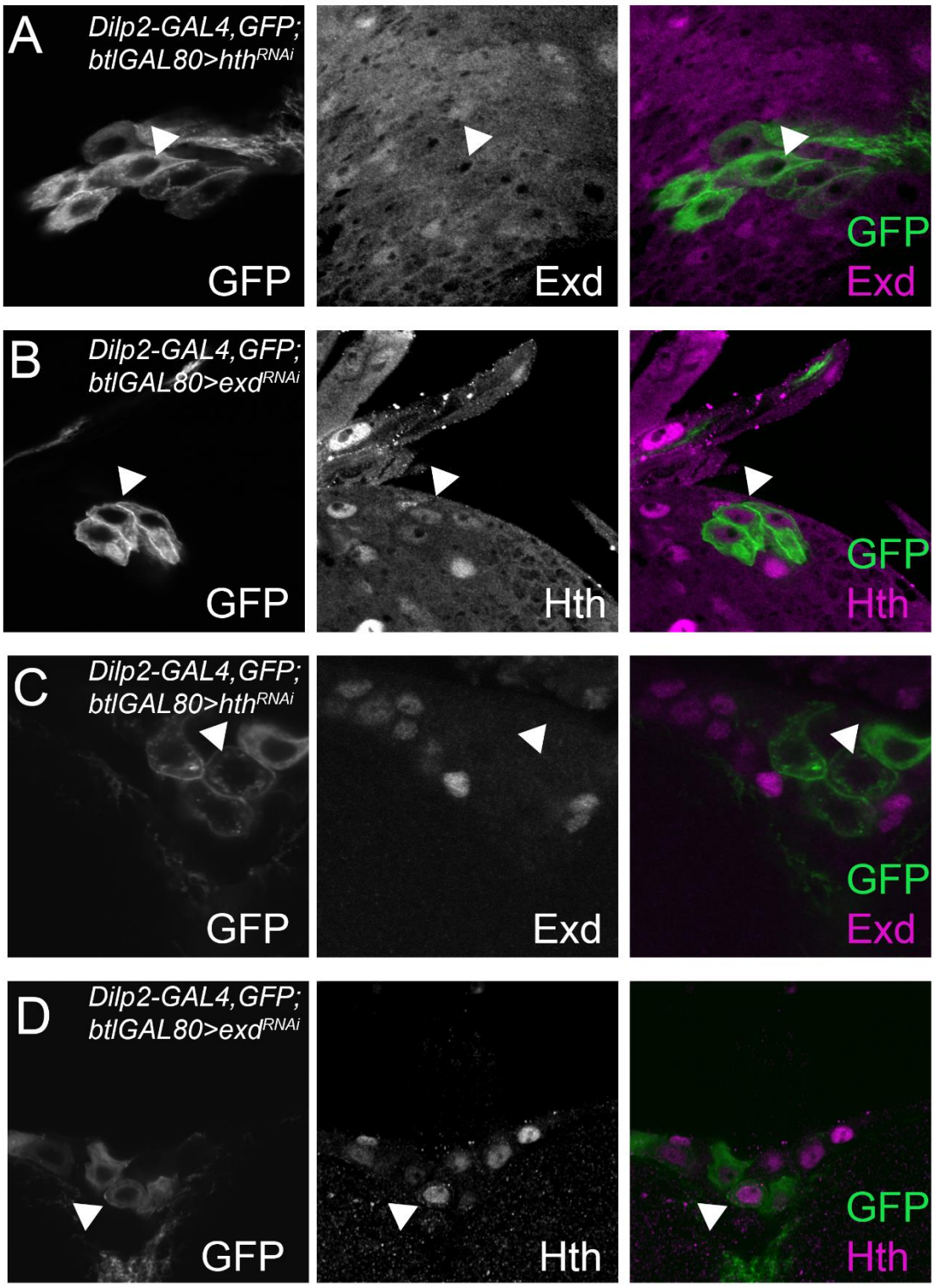
Supplemental Figure 1: Expressing an *mCherry*^{RNAi} in the IPCs does not affect *dilp* expression. *dilp2*, *-3* and *-5* mRNA levels are not significantly different in *Dilp2-GAL4*^{R} > *mCherry*^{RNAi} compared to *Dilp2-GAL4*^{R} /+ fly heads (unpaired two tailed Student's t-test $p > 0.05$).



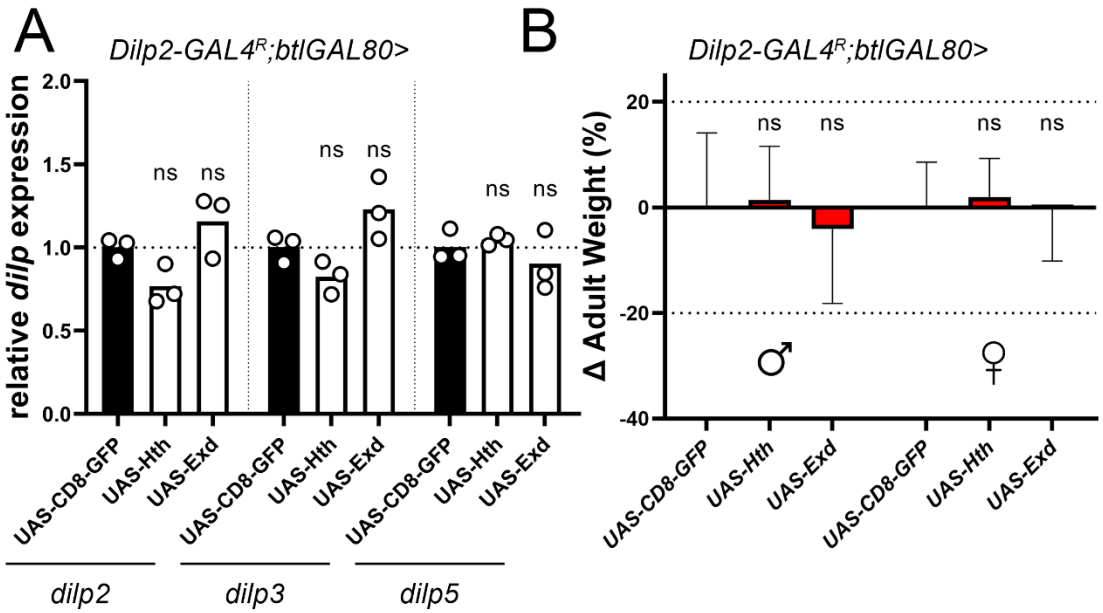
Supplemental Figure 2. Hth knockdown in the IPCs reduces *Dilp2-GAL4* activity but identical effects are seen with second GAL4 line. Representative images of adult female IPCs upon knockdown of mcherry (A) or Hth (B), IPCs are visualised by nuclear-GFP (nls-GFP) expression. Images were taken at identical confocal settings. (C) GFP fluorescence in the IPCs is significantly but mildly reduced upon hth depletion (Student's t-test, * $p<0.05$). Male (D) and female (E) *dilp2* expression is significantly reduced in *OK107-GAL4>hth^{RNAi}* flies compared to flies expressing an *mcherry^{RNAi}* control (Student's t-test, * $p<0.05$, **** $p<0.0001$). Adult weight is significantly reduced in *OK107-GAL4>hth^{RNAi}* male and female flies (Student's t-test, **** $p<0.0001$).



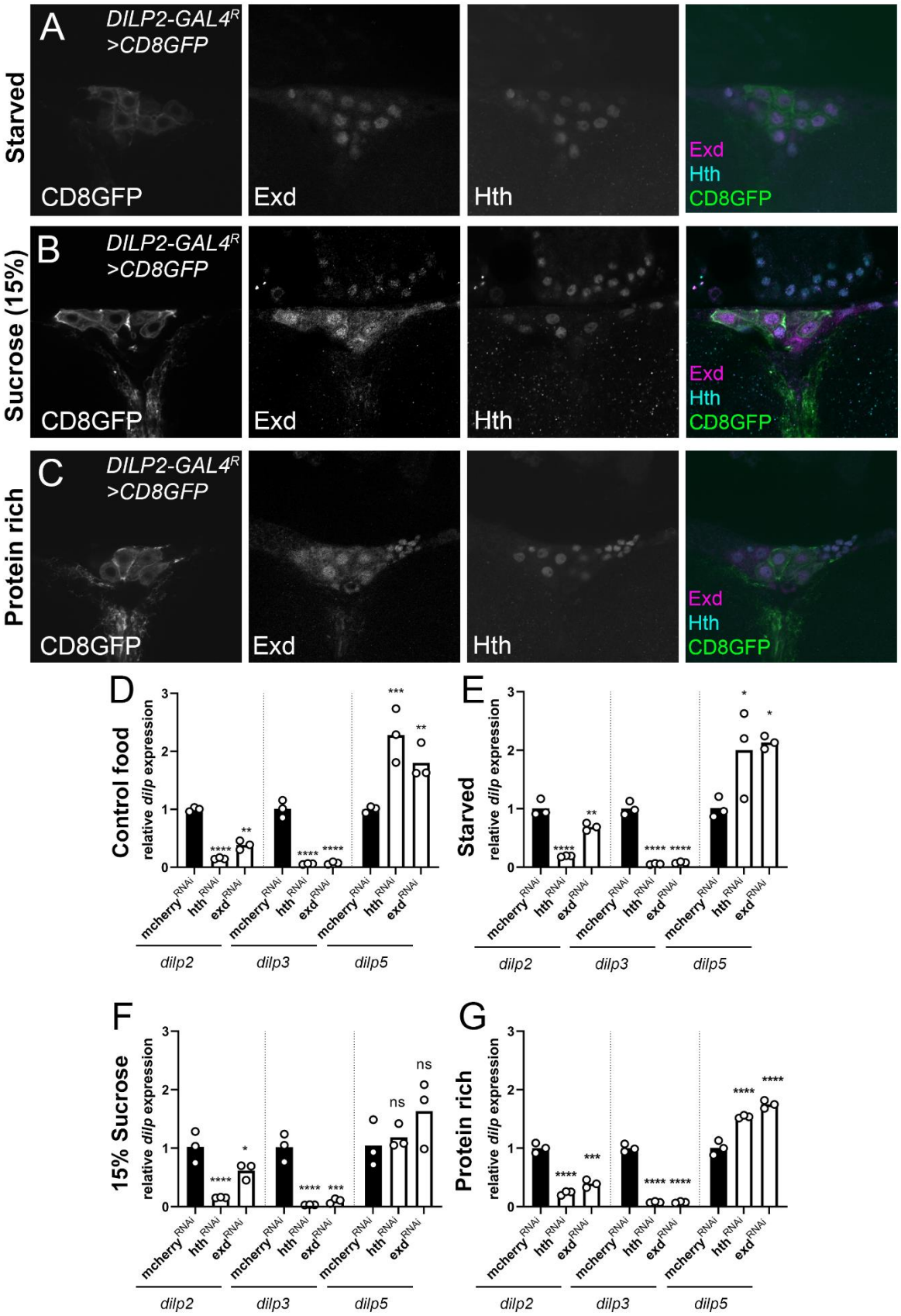
Supplemental Figure 3. Hth does not regulate Ey in in the IPCs. Genetic depletion of *hth* in the IPCs does not affect Ey expression in the IPCs. IPCs are marked with CD8-GFP driven by *Dilp2-GAL4*.



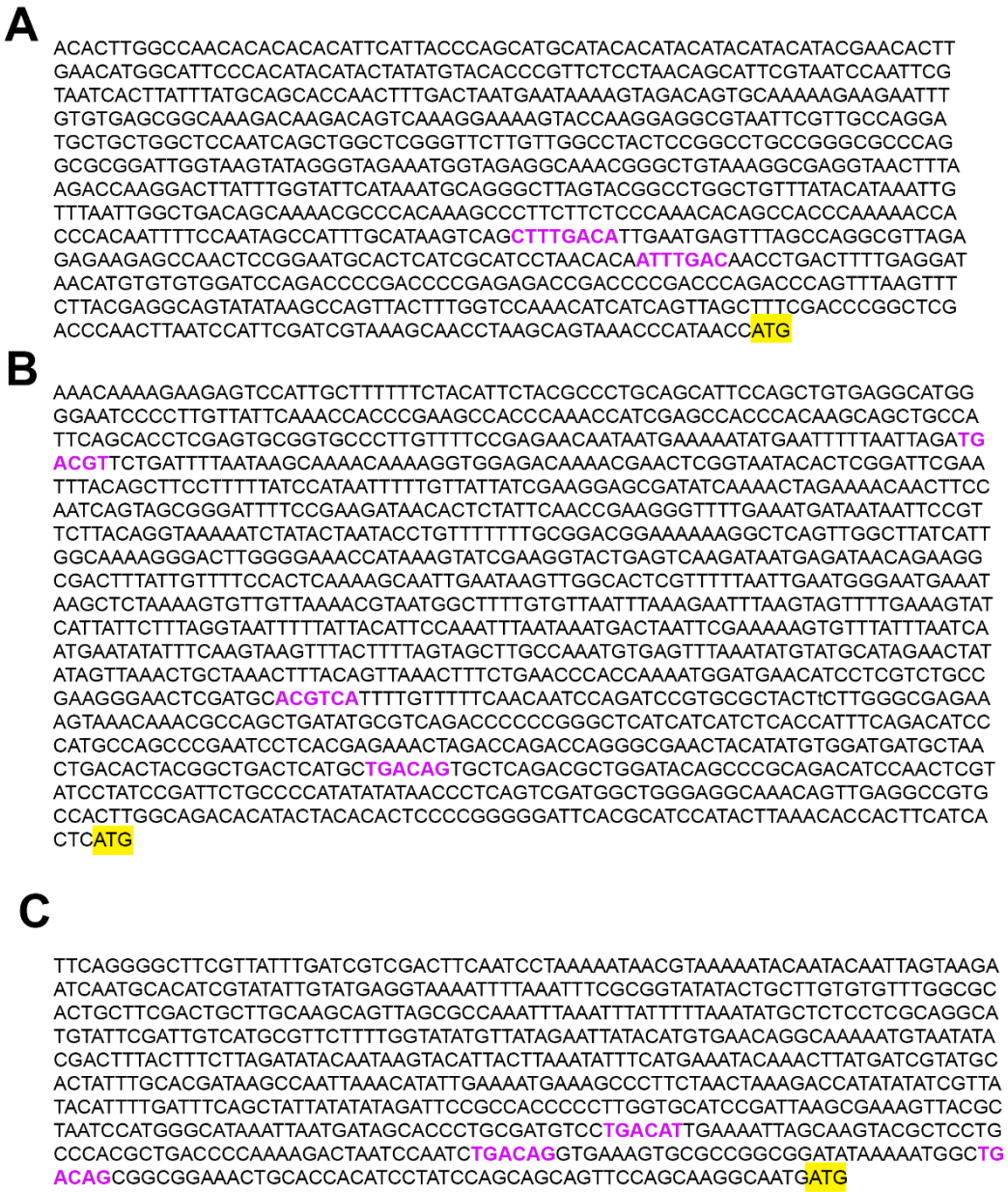
Supplemental Figure 4. Hth expression in the IPCs is essential for nuclear localization of Exd in both larval and adult stages but not vice versa. (A) Knockdown of *hth* results in the absence of Exd in larval IPC nuclei (arrow). (B) Knockdown of *exd* in larval IPCs does not affect Hth localization. (C) IPC-specific knockdown of *hth* results in the absence of Exd in IPC nuclei. (D) Hth localization is unaffected upon *exd* knockdown.



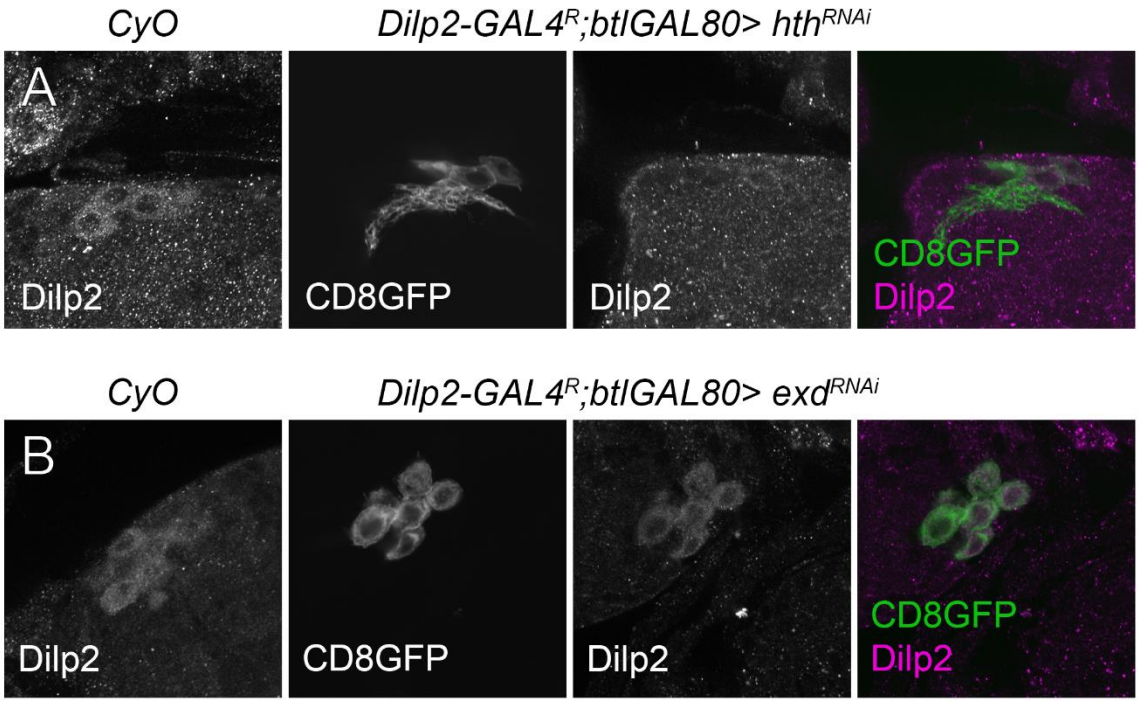
Supplemental Figure 5. Overexpression of Hth or Exd does not alter *dilp* expression or adult weight. (A) Male or female flies overexpressing Hth or Exd in the IPCs do not weigh significantly more or less than control flies overexpressing GFP. (B) Overexpression of Hth or Exd does not alter *dilp2*, *-3 or -5* mRNA levels (ns, not significant, one-way ANOVA with post-hoc Dunnett's test to correct for multiple comparisons).


A
DILP2-GAL4R >CD8GFP
Starved
CD8GFP
Exd
Hth
Exd
Hth
CD8GFP
B
DILP2-GAL4R >CD8GFP
Sucrose (15%)
CD8GFP
Exd
Hth
Exd
Hth
CD8GFP
C
DILP2-GAL4R >CD8GFP
Protein rich
CD8GFP
Exd
Hth
Exd
Hth
CD8GFP
D
Control food
relative dilp expression
E
Starved
relative dilp expression
F
15% Sucrose
relative dilp expression
G
Protein rich
relative dilp expression
mcherry RNAi
hth RNAi
exd RNAi
dilp2
dilp3
dilp5

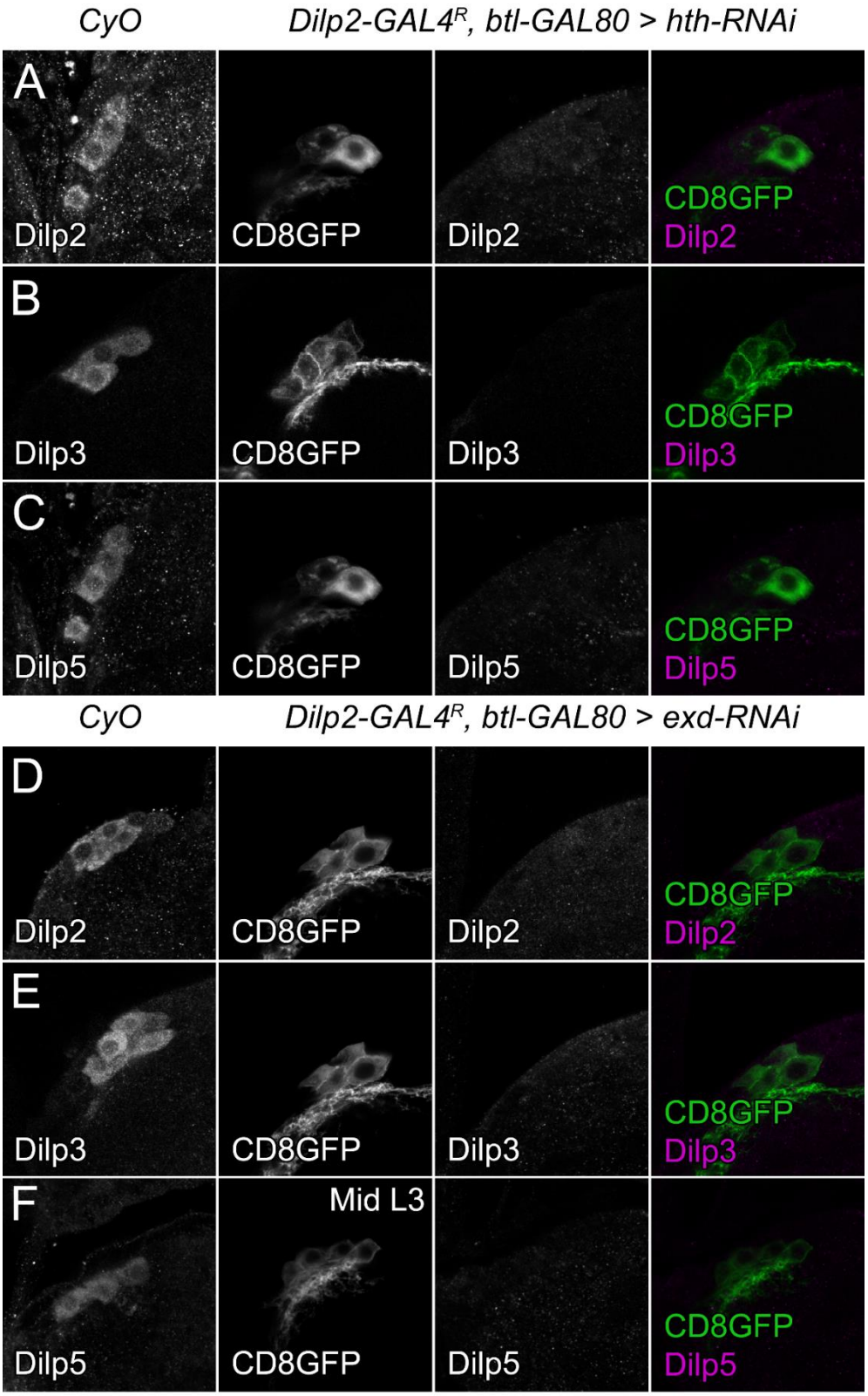
Supplemental Figure 6. Nuclear localization of Hth and Exd is not altered in different nutritional conditions and has nutrition-independent effects on *dilp* expression. Nuclear localization of Hth and Exd in the IPCs is unaltered in starved (A), sucrose only (B) and protein rich (C) conditions. Flies were reared on control food and 24–48-hour old females were placed on different diets for 72 hours and subsequently dissected and stained. Female flies expressing hth^{RNAi} or exd^{RNAi} driven by *dilp2-GAL4^{R};btlGAL80* were subjected to the same conditions as described above and collected for qPCR. Hth and Exd depletion had identical effects on *dilp2*, *-3* and *-5* expression levels in starved (E), sucrose only (F) and protein rich (G) as in control food (D). (ns, not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, Dunnett's multiple comparisons).



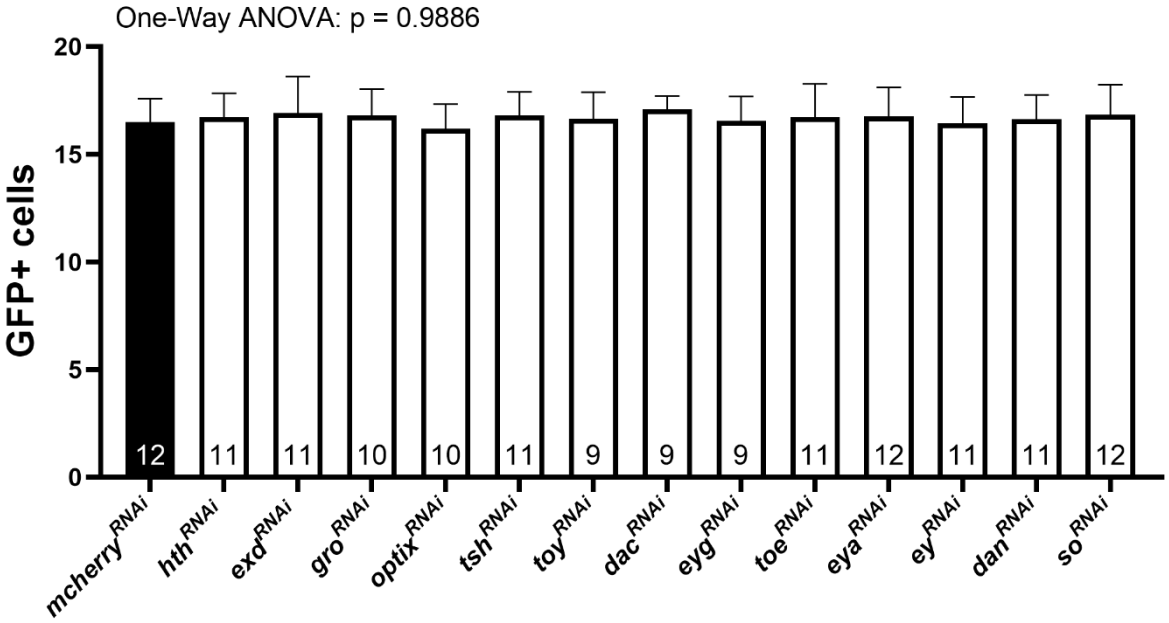
Supplemental Figure 7. Putative Hth binding sites within *dilp2*, *-3* and *-5* enhancers. 2, 3 and 3 Hth motifs (**purple**) were identified in *dilp2*, *-3* and *-5* enhancers, respectively. Hth binding motif from the JASPAR database was used to scan in *dilp2*, *-3* and *-5* enhancers using the FIMO from the MEME suite ($p<0.001$, both strands). Start codon is highlighted in yellow.



Supplemental Figure 8. Hth but not Exd depletion affects Dilp2 protein levels in the IPCs in late L2 stage (68h after egg laying). (A-B) Representative images of Dilp2 protein levels in *Dilp2-GAL4;btlGAL80> hth^{RNAi}* and *Dilp2-GAL4;btlGAL80> exd^{RNAi}* compared to control siblings (Cyo, left) reared in identical conditions. Flies were reared and age-controlled then brains from knockdown and balanced sibling flies were dissected and stained in the same tube. Expression of CD8GFP was used to differentiate control and test genotypes, and to label the IPCs. Images were acquired using identical confocal settings.



Supplemental Figure 9. **Hth and Exd depletion reduces Dilp2, -3 and -5 protein levels in mid-L3 (92h AEL).** Representative images of Dilp2, 3 and -5 protein levels in (A-C) *Dilp2-GAL4;btlGAL80> hth^{RNAi}* and (D-F) *Dilp2-GAL4;btlGAL80> exd^{RNAi}* compared to control siblings (Cyo, left) reared in identical conditions. Flies were reared and age-controlled. Brains from knockdown and balanced sibling flies were dissected and stained in the same tube. Expression of CD8GFP was used to differentiate control and test genotypes, and label IPCs. Images were collected



Supplemental Figure 10. Knockdown of RDGN transcription factors does not affect IPC cell number. Female flies expressing RNAis against RDGN transcription factors using *Dilp2-GAL4;nls-GFP* were dissected 4-6 days post-eclosion, fixed and mounted. The number of IPCs was determined by manually counting GFP+ nuclei. Numbers of brains analysed per genotype are shown inside the bars (ns, not significant, one-way ANOVA with post-hoc Dunnett's test to correct for multiple comparisons). Images were acquired using identical confocal settings.

4 The *TCF7L2* homolog, *pangolin*, is essential for *Drosophila* insulin-producing cell development.

Winant M[1], Callaerts P[1]

1 Laboratory of Behavioral and Developmental Genetics, Department of Human Genetics, KU Leuven—University of Leuven, Leuven, Belgium,

Abstract

Drosophila Insulin-Producing Cells (IPCs) are a group of neuroendocrine cells situated in the pars intercerebralis of the brain that produce *Drosophila* insulin-like peptides 2, 3 and 5. The IPCs are the functional equivalent of the pancreatic beta cells. Previous work has identified a few homologous transcription factors to act in IPCs and pancreatic beta cells, suggesting that the regulation of development and function of these cells is conserved. Multiple genome-wide association studies have independently found that variants in *TCF7L2* are strongly associated with increased type 2 diabetes risk and several studies show that it is involved in beta cell functioning. We here show that Pangolin, the fly homolog of TCF7L2, is involved in development and regulation of IPCs. IPC-specific expression of a dominant-negative form of Pangolin results in smaller flies that are developmentally delayed. This is the result of reduced insulin-like peptide expression in the IPCs. The IPCs additionally show prominent morphological defects in neurite complexity and cell size. We go on to show that Pangolin acts upstream of Dimmed, a transcription factor that promotes a prosecretory cell fate, of Eyeless which regulates IPC morphology and controls insulin-like peptide transcription, and of Homothorax which in the larval stage is also involved in insulin-like peptide transcriptional control. Our work identifies an additional regulator of the IPCs that is conserved in the mammalian pancreatic beta cell, lending credence to the hypothesis that both cell types are evolutionarily related.

Introduction

One in every ten adults suffers from diabetes, a number that is predicted to rise in the foreseeable future[211]. Type 2 diabetes (T2D) accounts for most diabetes cases and has a complex etiology. In healthy individuals the feedback loop between the insulin-producing beta-cells in the pancreas and the insulin-sensitive tissues results in normal glucose homeostasis[212]. When insulin resistance develops in the peripheral tissues, this causes increased insulin production and secretion by the beta cells[213]. Failure of this compensatory mechanism results in T2D.

T2D diabetes has a strong genetic component and large scale GWAS studies have identified a plethora of genetic variants that are significantly associated with development of type 2 diabetes[214]. The most prominent identified locus is *TCF7L2* which has been robustly and independently replicated in various genetically distinct populations[215–217]. It encodes for a transcription factor with a central role in canonical Wnt signaling[218–220]. Wnt signaling is activated by the binding of Wnt ligands to Frizzled receptors, leading to the stabilization and nuclear translocation of the transcriptional co-activator β-catenin[221]. β-catenin can then bind members of the TCF/LEF family of transcription factors causing them to switch from being a transcriptional repressor to an activator[221]. The exact function of Tcf7l2 in the etiology of T2D has been the subject of intense scientific inquiry. Several studies in mice revealed a role for TCF7L2 in fat tissue, the liver and the intestinal L-cells where its dysfunction/dysregulation contributes to T2D[222–224]. TCF7L2 was also shown to be essential in the beta cells for the regulation of beta cell mass and insulin content[225]. In contrast, others have shown that Tcf7l2 in murine beta cells is dispensable for metabolic phenotypes[223]. In another non-mammalian model organism, zebrafish carrying just a single functional *tcf7l2* copy present with significant pancreatic dysmorphology and dysregulated glucose homeostasis[226], showing that the role of TCF7L2 in metabolism is widespread amongst vertebrates.

In *Drosophila*, the functional equivalent of the beta cells are the Insulin-Producing Cells (IPCs), a group of neuroendocrine cells that produce *Drosophila* insulin-like peptides 2, 3 and -5 (Dilp2, 3 and -5)[98]. These Dilps are homologous to mammalian insulin/IGFs and regulate growth and metabolism[3,89,90,92]. The IPCs exhibit, despite being neurons, remarkable similarities to beta cells. They are connected via their neuronal processes to AKH-producing cells of the corpora cardiaca, the equivalent of the pancreatic glucagon-producing alpha cells, and can autonomously sense circulating glucose levels[90,130]. Genetic ablation of the IPCs results in hyperglycemia, reminiscent of type 1 diabetes[90]. *Drosophila* models of T2D have been developed through the induction of peripheral insulin resistance either genetically, e.g. insulin receptor hypomorphic mutants, or through high-sugar and high-fat diets[91,227,228].

Work from our lab and others has identified several transcription factors that are essential for IPC function and *dilp* regulation and have a conserved role in beta cells[149,150,229]. PAX6, for example, directly regulates insulin transcription and additionally instructs endocrine cell development while its homolog Eyeless (Ey) directly regulates *dilp5* and is necessary for normal IPC development[149,230,231]. Additionally, we have shown that Homothorax (Hth) and its co-factor Extradenticle (Exd) are involved in the regulation of *dilp2* and *-3*[229]. The conservation of several of these TFs suggests that there are similar transcriptional networks at play in the IPCs and the beta cells.

Despite the prominent role of TCF7L2 in diabetes and in beta cell functioning, it is not known whether its homolog Pangolin (Pan) has an equivalent role in the IPCs. Here, we show that Pan activity is essential for normal IPC development in *Drosophila* and that it is essential for normal Dimmed (Dimm) expression while also impairing the normal expression of *dilp2*, *-3*, and *-5* in the IPCs. Pangolin thus plays a key role in establishing IPC cell fate during development, a role that is conserved in the mammalian beta cells.

Results

We used available transcriptomic data[103,229,232] and a GAL4 inserted in the *pangolin* locus and thus controlled by endogenous *pan* regulatory sequences to demonstrate that Pangolin is expressed in both larval and adult IPCs (Figure 19 A - B).To establish a role for Pan in IPC development and function we initially used transgenic RNAi constructs. We observed high variability of outcomes, an observation not uncommon for knockdown of transcription factors. Therefore, instead we overexpressed a dominant negative form of Pan (*UAS-Pan*^{DN}), which lacks the Arm binding domain and therefore acts as a constitutive transcriptional repressor, in the IPCs. This results in flies that were visibly smaller and developmentally delayed (Figure 20 A-B). Quantification of adult weight and wing area showed that expressing dominant-negative Pan in the IPCs significantly perturbed normal development in both males and females (Figure 20 C-D). This delayed eclosion of ~28h is due to an increase in the duration of larval developmental duration as time until pupariation is equally increased, showing that the duration of metamorphosis remains unaffected (Figure S11 A). Similar reductions in adult weight were seen with two additional independent *Dilp2-GAL4* lines (Figure S11 B - C).

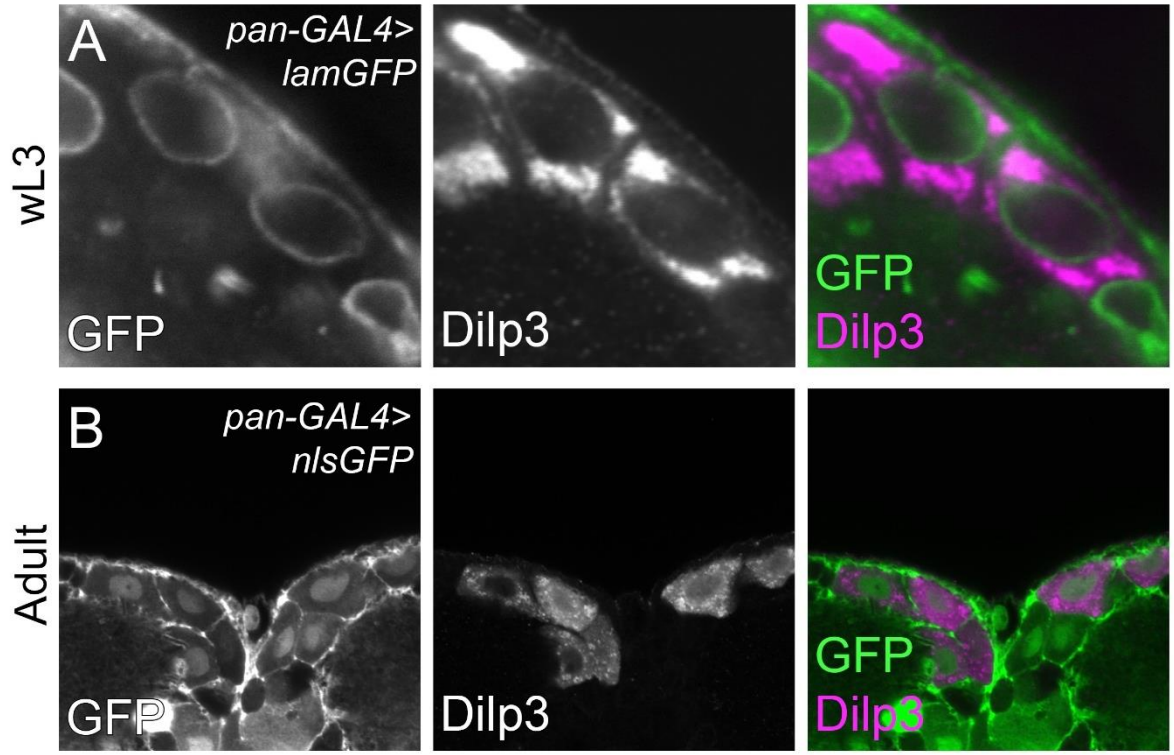


Figure 19. Pangolin is expressed in the IPCs. *Pan-Gal4* drives expression in the IPCs, as shown with double-labeling with anti-GFP (green) and anti-Dilp3 (magenta) in both late larval (A) and adult IPCs (B).

Size and developmental timing phenotypes are associated with reduced IIS during the larval stage[233]. We performed qPCR on larvae at different timepoints during larval development and measured *dilp2*, *-3* and *-5* mRNA levels to quantify IIS levels. This revealed that *dilp2* and *-5* expression, which are dynamically regulated in the larval stage, show no such dynamic patterns, and remain relatively constant and reduced compared to control larvae (Figure 21 A and C). Interestingly, *dilp3* expression, which is significantly lower than *dilp2* and *-5* and only peaks in expression during the wandering third instar larval stage (L3) and during pupation, is slightly upregulated for most of larval development. This is consistent with previous results showing that precocious *dilp3* upregulation occurs in *dilp5* mutants[99]. Expression of *dilp3*, however, remains significantly lower during the wandering L3 stage (Figure 21 B). Expression of *dilp1* is increased in first instar larvae (L1) compared to controls and is significantly higher during L3 compared to controls (Figure S13 A). A compensatory upregulation of *dilp6* is seen in late second instar larval stage (L2) and persists during L3. The programmed peak in *dilp6* expression during the non-feeding wandering stage (~112h AEL) is delayed (Figure S13 B). The normal *dilp* expression profile during larval development is thus perturbed upon Pan^{DN} expression in the IPCs. This results in the precocious upregulation of *dilp1* and *-3*, which are normally both expressed in the IPCs during late L3. Similarly, there is an early upregulation of *dilp6*, presumably to compensate for low IIS activity from the IPCs. Low IIS levels during

development are therefore responsible for the developmental delay and reduction in adult size.

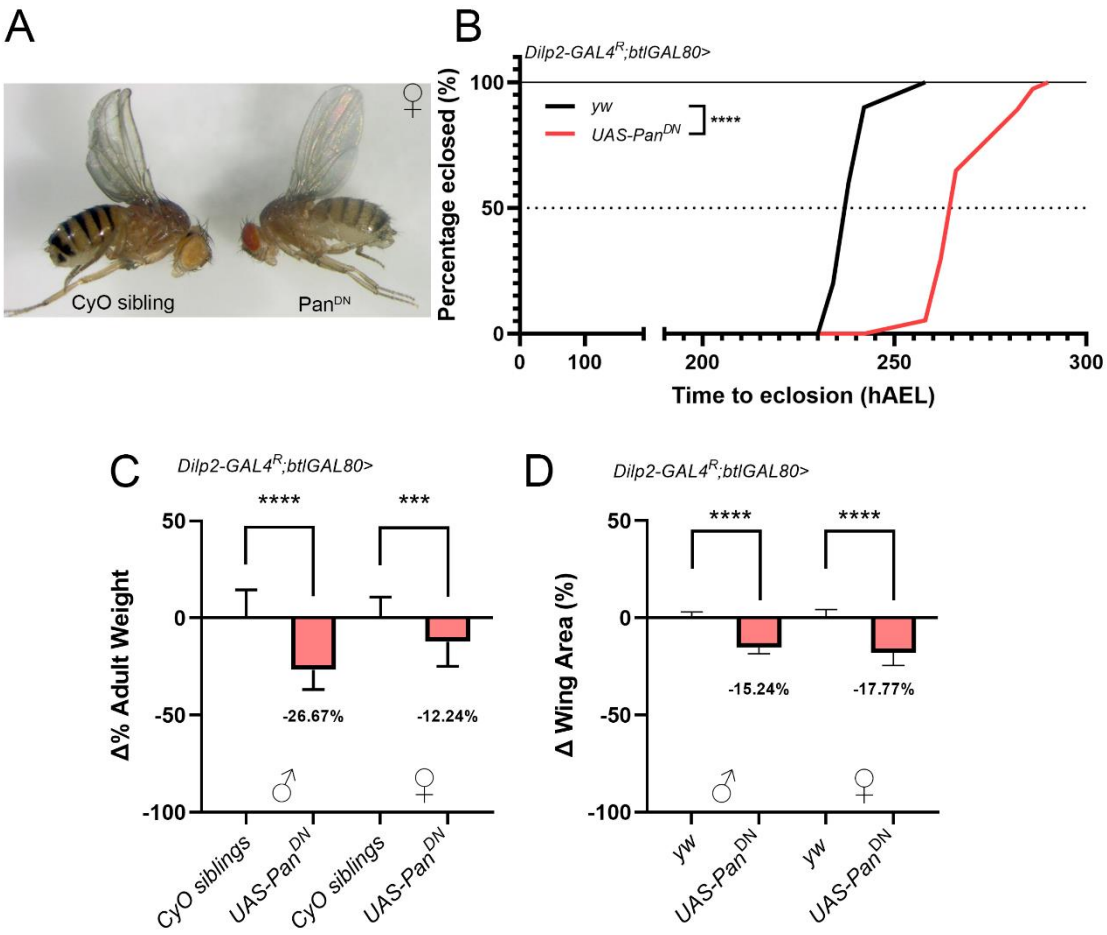


Figure 20. Pangolin regulates *Drosophila* growth and development in the IPCs. (A) Overexpression of UAS-Pan^{DN} in the IPC using Dilp2-GAL4^{R} (right) results in visibly smaller flies compared to control sibling controls (left). (B) UAS-Pan^{DN} overexpression in the IPCs results in delayed adult eclosion ($p<0.0001$, unpaired t-test; n=4 *yw/w; Dilp2-GAL4*, *UAS CD8-GFP/+;btlGAL80/+*, n=6 *yw/w*; *Dilp2-GAL4*, UAS-*CD8-GFP* /+;*UAS-Pan^{DN}/btlGAL80* with each group= 8-12 female flies). (C) IPC-specific UAS-Pan^{DN} expression significantly reduces adult weight compared to sibling controls in both males and females (unpaired t-test: n=18 males $p<0.001$; n=23 females $p<0.0001$). (D) Wing area of males (left) and females (right) was significantly reduced compared to control flies (unpaired t-test, n=6 and n=11 males $p<0.0001$; n=13 and n=14 females $p<0.0001$).

We then evaluated the effects of developmental expression of Pan^{DN} overexpression in the adult stage. This resulted in severe reductions of *dilp2*, *-3* and *-5* mRNA levels (Figure 21 D, G and J) and the complete absence of detectable Dilp2, -3 and -5 protein in IPC somata (Figure 21 E-F, H-I and K -L). Restricting the expression of Pan^{DN} to the adult stage using a temperature sensitive GAL80^{ts} reduced *dilp5* mRNA levels by half while leaving *dilp2* and *–3* unaltered (Figure 22 A- C). This shows that Pan is mostly necessary during the larval/pupal stage but also plays an adult specific role in the maintenance of normal *dilp5* expression.

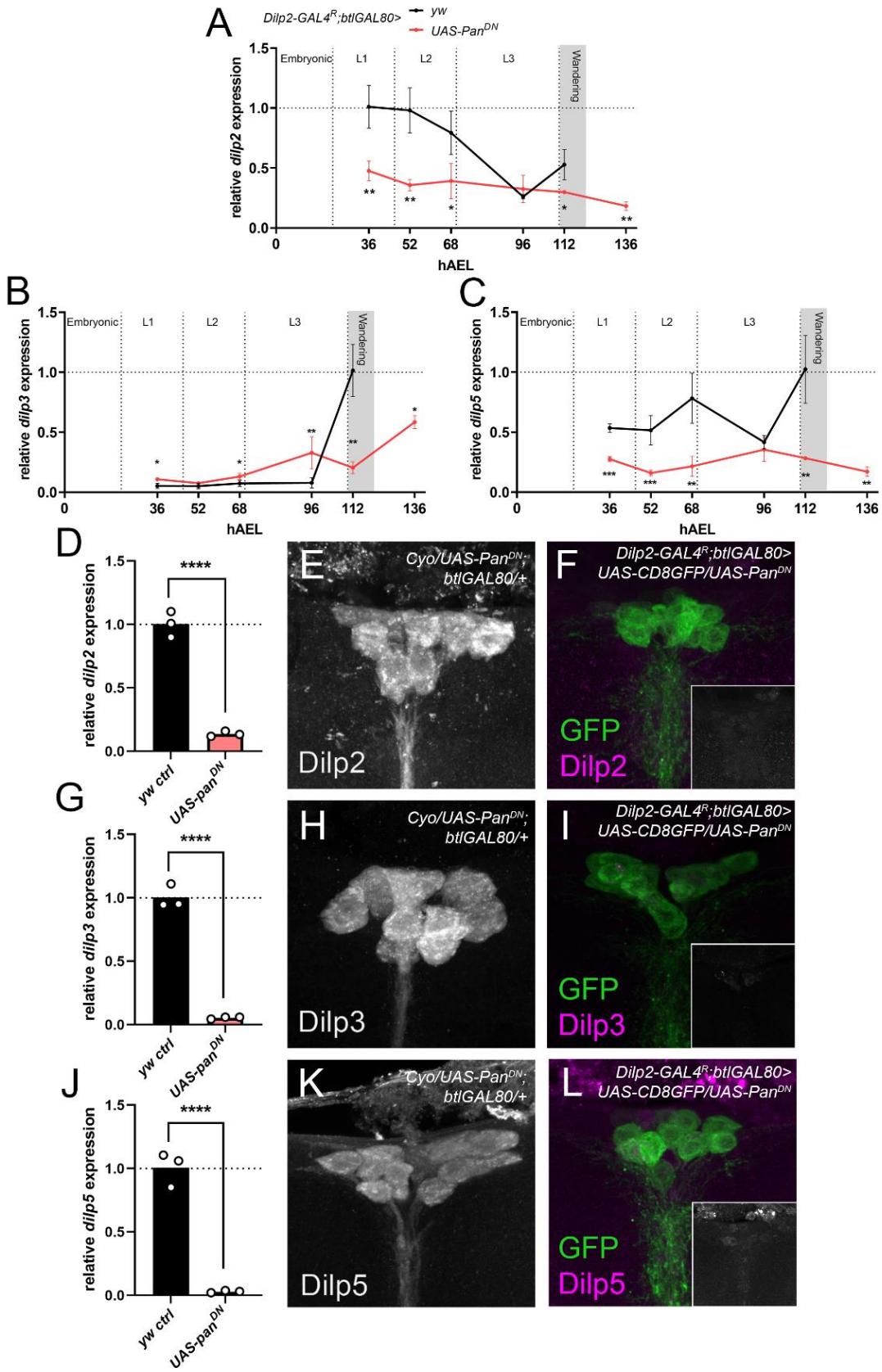

A
Dilp2-GAL4R;btlGAL80>
yw
UAS-PanDN
relative dilp2 expression
Embryonic
L1
L2
L3
Wandering
hAEL
B
relative dilp3 expression
C
relative dilp5 expression
D
E
Cyo/UAS-PanDN;
btlGAL80/+
F
Dilp2-GAL4R;btlGAL80>
UAS-CD8GFP/UAS-PanDN
GFP
Dilp2
G
H
Dilp3
I
J
K
Dilp5
L
yw ctrl
UAS-panDN

Figure 21. Pangolin controls *dilp2*, *-3 and -5* mRNA levels during larval development and is essential for normal adult *dilp* expression. (A-C) Timed control and Pan^{DN} larvae (*yw/w; Dilp2-GAL4*, *UAS-CD8-GFP/+; btlGAL80/+* and *yw/w; Dilp2-GAL4*, *UAS-CD8-GFP/+; btlGAL80/UAS-Pan^{DN})* were collected at 36, 52, 68, 96, 112, and for Pan^{DN} larvae at 136 hAEL for qPCR. Unpaired t-tests were used to compare mRNA levels at each timepoint between both genotypes and for the 136 hAEL Pan^{DN} timepoint with 112h hAEL control. Relative quantities compared to the highest value for each dilp in the control larvae are reported. (*p<.05, **p<0.01, ***p<0.001, ****p<0.0001). The timing of developmental checkpoints of control larvae are shown. (D) *dilp2* (G) *dilp3* and (J) *dilp5* mRNA levels are severely reduced upon IPC-specific *pan^{DN}* expression compared to *yw/+*; *Dilp2-GAL4*, *CD8-GFP/+;btlGAL80/+* control flies in the adult stage (p<0.0001). Representative image of Dilp2 (E - F) Dilp3 (H - I) and Dilp5 (K – L) showing that Dilp protein levels are reduced in adult IPCs as a result of Pan^{DN} expression compared to balanced age-controlled siblings (*CyO/UAS-pan^{DN}; btlGAL80/+)*. Insets (F, I, L) show Dilp2, -3 and -5 only (grayscale). Brains from knockdown and balanced siblings were dissected and stained in the same tube and CD8-GFP was used to differentiate between control and test genotypes and to label experimental IPCs. Only female flies were used.

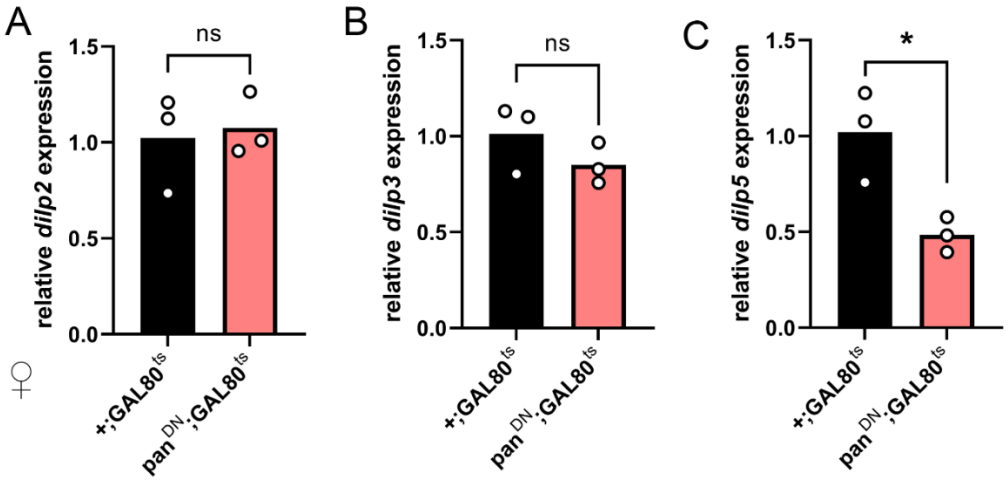


Figure 22. Limiting Pan^{DN} expression to the adult stage has minor effects on *dilp* expression. mRNA level of *dilp2* (A) and *dilp3* (B) are unaffected by adult restricted Pan^{DN} expression while *dilp5* (C) is significantly reduced. Temporal restriction was obtained using Gal80^{ts}, Flies were reared at 18°C and shifted to 29°C for 3 days prior to RNA collection. (Unpaired t-test, ns not significant, *p<0.05).

Morphological examination of the IPCs showed that they present with considerable defects such as reduced complexity in the lateral branching of their neurites and reduced median bundle width (Figure 23 A - B). Both phenotypes are also seen in *eyeless* (*ey*) mutants and we thus hypothesized that *ey* expression in the IPCs might be perturbed[149]. Indeed, Ey which is normally present in adult and larval IPCs can no longer be detected in IPC nuclei upon Pan^{DN} expression (Figure 23 C-H and Figure S14 A - B)[229]. Dachshund (Dac) which cooperates with Ey to regulate *dilp5* expression remains unaffected and is readily detectable in larval and adult IPCs (Figure S15 A - D). We further evaluated the role of Hth, a transcription factor for which we recently determined a role in *dilp2* and *-3* regulation[229]. Larval Hth levels in the IPCs were reduced upon Pan^{DN} expression while adult Hth levels were unaffected, consistent with the absence of effects on *dilp2* and *-3* expression when Pan^{DN} is restricted to the adult stage (Figure S16 A - D).

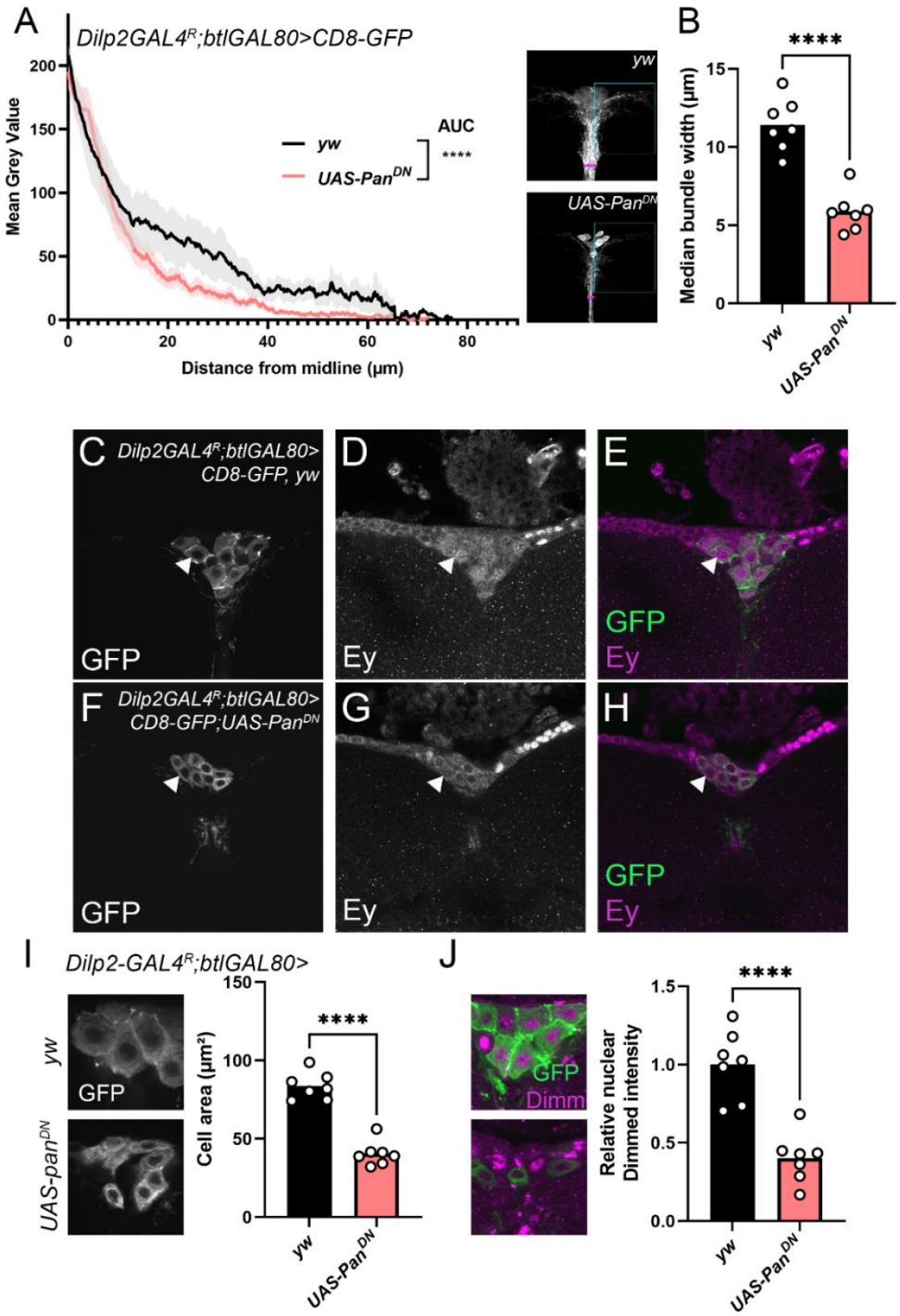


Figure 23. Pan controls IPC morphology and key developmental IPC transcription factors Dimmed and Ey. (A) Lateral neurite complexity is reduced upon Pan^{DN} overexpression as measured by mean grey value of CD8-GFP. Example images of *yw; Dilp2-GAL4*, *CD8-GFP/+; btlGAL80/+* (top) and yw*; Dilp2-GAL4*, *CD8-GFP/UAS-Pan^{DN}; btlGAL80/+* (bottom), cyan box indicates the area where grey values were measured. AUC values were compared using an unpaired t-test ($p<0.0001$) (B) Median bundle is reduced. Magenta line in example images indicates where bundle thickness was measured (unpaired t-tests, ****$p<0.0001$, n =7). Ey is detectable in adult IPC nuclei (C – E) but is absent upon Pan^{DN} expression (D - G). Test and control genotypes were stained in the same tube and one optic lobe was removed from control brain to differentiate them. Only female flies were used. (I) Expression of Pan^{DN} in IPCs significantly reduces IPC cell size (unpaired t-test, $p<0.0001$ each data point is average cell area of 5 IPCs from a single brain). Representative examples are shown on the left at identical magnification. (J) Nuclear Dimm expression is reduced upon Pan^{DN} expression (unpaired t-test, $p<0.0001$, each data point (n=7) is the average Dimm intensity of 5 IPCs from a single brain.

IPC-specific Pan^{DN} also results in considerably smaller IPC cell bodies (Figure 23 I). The transcription factor Dimm, which acts as a scaling factor for secretory cell fate, is known to be expressed in the IPCs and determines cell size in other neuropeptide producing cells[113,114]. We therefore hypothesized that perturbing Pan in the IPCs also impacts Dimm expression. IHC revealed that Dimm protein levels were indeed significantly reduced compared to control conditions (Figure 23 J). Dimm expression in larval IPCs was also reduced but this effect was more variable (Figure S14 C - D). We additionally evaluated the effect of restricting Pan^{DN} to the adult stage only, as compared to developmental misexpression. This does not alter nuclear Dimm in the IPCs, indicating a developmental role (Figure S16 A).

Besides clear effects in weight and developmental timing we set out to determine additional phenotypic consequences. Given that IPC *dilp* levels approach zero upon Pan^{DN} expression, we hypothesized that this would phenocopy what is seen upon ablation of the IPCs[92]. Indeed, starvation and oxidative stress resistance were increased to a similar extent (Figure S17 A and C). Egg-laying as a measure for fecundity was similarly reduced (Figure S7 B). Total lipid content, however, was unaffected by Pan^{DN} (Figure S17 D). This further demonstrates that perturbing Pan in the IPCs results in physiological phenotypes associated with reduced systemic IIS except for lipid content.

The repressive function of Pangolin depends on multiple co-factors including Groucho (Gro), SoxNeuro (SoxN) and D (Dichaete). It has been shown that their overexpression in other tissues is sufficient to suppress Wnt signaling[234–236]. Overexpression of these co-factors in the IPCs resulted in significant reductions of *dilp2*, *-3* and *-5* expression except for overexpression of D on *dilp5* expression. For the latter, lower expression was observed albeit not statistically significant (Figure S18 A - C). Overexpression of SoxN or D also reduced IPC cell body size and adult weight (Figure S18 D - E), thus fully phenocopying Pan^{DN}. Overexpression of repressor co-factors thus shows that the effects of Pan^{DN} are mediated by its repressor activity and that inactivation of this repressor function is essential for proper IPC development.

We thus show that Pan acts upstream of at least three critical IPC transcription factors (Ey, Dimm and Hth) to instruct IPC development and is necessary for normal *dilp* expression (Figure 24).

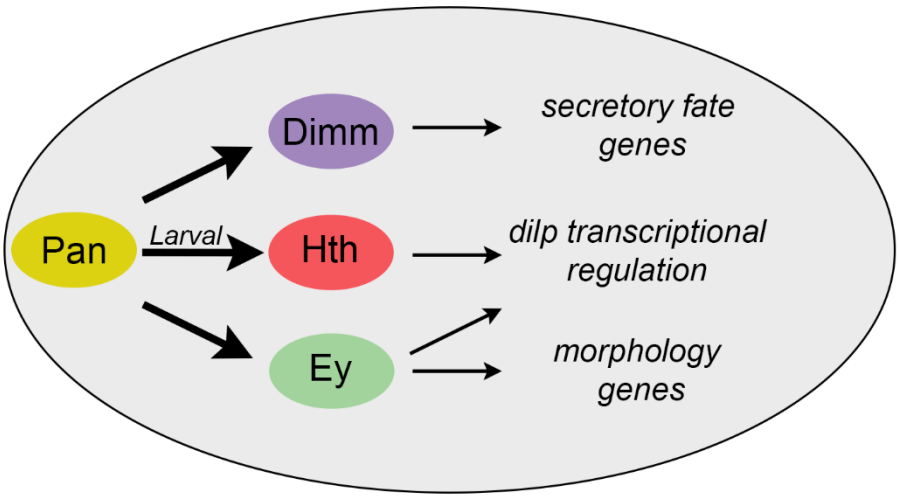


Figure 24. Model of Pangolin function in the IPCs. Pangolin acts upstream of Dimm which controls a prosecretory cell fate. Pangolin also acts upstream of Ey which is involved in IPC morphology and the direct regulation of *dilp5*. In the larval stage, Pan is additionally upstream of Hth in that larval stage and is itself involved in *dilp2* and *-3* regulation.

Discussion

Our work has shown that the TCF7L2 homolog, Pan, is essential for IPC development and differentiation. Upon expression of a dominant-negative form of Pan in the IPCs we found that this results in major defects in *dilp* expression and IPCs morphology. We show that Pan acts upstream of Dimm which acts as a scaling factor in neurosecretory cells, promoting formation of Large Core Dense Vesicles and increased soma size[111]. Dimm is a member of the ATONAL family of basic Helix-Loop-Helix (bHLH) transcription factors and shares highest sequence similarity with mammalian MIST1[113,237]. MIST1 is found in the pancreas but is restricted to the acinar cells which produce digestive enzymes where MIST1 promotes the expansion of endoplasmic reticulum, assembly of secretory granules and proper Ca^{2+}signaling[238–240]. Other members of the ATONAL family include NEUROD1 which is essential for beta cell maturation and regulation of insulin transcription[29]. *Tcf7l2*$^{-/-}$ mice present with highly reduced NeuroD1 expression in beta cells[241]. Interestingly, previous work has shown that Wnt3a can stimulate *Ins2* production in the murine hypothalamus via the induction of NeuroD1[242]. There thus seems to be a conserved link between Wnt signaling and insulin expression via NeuroD1/Dimm. The IPCs are located in the pars intercerebralis (PI) of the brain, which has been proposed to be similar to the vertebrate hypothalamus on structural, developmental and functional levels[96].

We further show that Pan acts upstream of Ey which has been shown to be essential for proper IPC morphology and additionally directly regulates *dilp5* expression[149]. The mammalian homolog of Ey, PAX6, plays a similar role in the beta cells. Although no interactions between TCF7L2 and PAX6 have been identified in mammalian beta cells, a more detailed analysis

might reveal the existence of such interactions. The findings that *dilp* expression levels of the IPCs are mildly reduced by Pan^{DN} during the larval stage and show only a modest developmental delay contrasts with the total absence of *dilp* mRNA levels in the adult stage. This suggests that Pan is necessary during the larval stage but is additionally critical during metamorphosis, a period where there is extensive remodeling of the IPCs[103]. The exact functional consequences of altered IPC neurite morphology remain poorly understood, but a reasonable assumption is that this would result in reduced or altered connectivity with neuronal partners such as DN1 clock neurons or GABAergic partners[140,243].

The effects of Pan^{DN} on adult *dilp* expression are stronger than on larval *dilp* expression. We previously showed that Hth is essential for adult *dilp2* and *-3* expression[229]. However, Hth was unaffected in the adult upon expression of Pan^{DN}, indicating that the downregulation of *dilp* expression does not depend on Hth in this case. In the larval stage, there consistently was an absence of nuclear Hth in the IPCs which implies that *hth* expression depends on Pan in the larval stage but not in the adult stage. Differences between larval and adult IPCs, for example, with regard to their ability to autonomously sense glucose levels have been noted before[126,190]. These differences in regulation and function show that there are underlying differences in transcriptional control between larval and adult IPCs.

The similarities between human beta cells and *Drosophila* IPCs can be leveraged by researchers to screen *Drosophila* homologs of genes associated with T2D to determine if they are involved in IPC and therefore perhaps in beta cell functioning[153]. Such a screen in *Drosophila* has been performed and was able to identify and validate *BCL11A* as a negative regulator of insulin secretion in mice[197]. Interestingly, this same study found no effect on circulating Dilp2 levels upon RNAi-mediated *pan* knockdown in the IPCs.

We further showed that overexpression of Gro, SoxN, and D, co-factors of the Pan repressor complex, are sufficient to phenocopy the effects of Pan^{DN}. SoxN and D are members of the SoxB family of proteins which play key roles in the maintenance of an undifferentiated neuronal stem cell-like state in vertebrates and invertebrates [244,245]. SOX9 plays this role in the maintenance of pancreatic progenitors[246]. Based on these observations, we propose that the association of variants in *TCF7L2* with reduced beta cell mass could be caused by aberrant maintenance of undifferentiated progenitors[225].

In conclusion, we show that inactivation of the Pan repressor complex is essential for IPC development and differentiation by regulating key transcription factors such as Ey, Dimm and larval Hth (Figure 24).

Methods

Fly stocks and husbandry

All experiments were performed on Nutrifly (Genesee Scientific) at 25°C with a 12h/12h day night cycle unless otherwise specified. IPC-specific manipulations were performed with *Dilp2-GAL4^{R};btlGAL80* originally described in[90,123] and *Dilp2-GAL4$^{Janelia\ 247}$* and *Dilp2-GAL4$^{Park\ 130}$*. A *pan-GAL4* (BL #93683) was used to determine the spatiotemporal expression pattern of Pan. *UAS-Pan^{DN}* (BL #4784) was used on its own or together with *tub-GAL80^{ts}* to temporally control its activity. y-w- background was used as controls for *UAS-Pan^{DN}*. *UAS-CD8-GFP* (Bl #5137), *UAS-nls-GFP* (BL # 4776) and *UAS-lamin-GFP* (BL #7378) were driven by GAL4 lines to visualize cells.

Immunohistochemistry

Larval and adult brains were dissected in cold 1X PBS, fixed for 20 minutes in 4% formaldehyde, and washed in PBS prior to 30'-60' blocking in PAXD (PBS containing 1% BSA, 0.3% Triton X-100, 0.3% deoxycholate). Brains were incubated overnight at 4°C in PAXD with primary antibodies, washed and incubated overnight at 4°C in PAXD with secondary antibodies. Samples were then washed and mounted in Vectashield Mounting Medium (Vector Laboratories Cat#H-1000). Images were taken with an Olympus FluoView FV1000 confocal microscope and processed using ImageJ64 and Photoshop software[205].

Antibodies used in this study include, mouse α-GFP (Roche 11814460001, 1:100), rat α-Dilp2 (1:100), rabbit α-Dilp3 (1:100), rabbit α-Dilp5 (1:400), guinea pig anti-Dimm (Dr. Paul Taghert, 1:400), mouse α-Ey (DSHB, 1:5), rabbit α-Hth (Dr. Richard Mann, 1:200) and mouse α-Dac (DSHB mAbdac2-3, 1:10). Secondary antibodies include donkey α-mouse Alexa-488, goat α-guinea pig Cy3, donkey α-mouse Alexa-594, donkey α-rat Alexa-594, donkey α-rabbit Alexa-594, goat α-rabbit Cy5. All secondary antibodies were from Jackson Immunoresearch and were used at 1:200 concentration.

Rat antibodies directed against the Dilp2 partial peptide sequence TRQRQGIVERC, rabbit antibodies directed against the Dilp3 partial peptide sequence DEVLRYCAAKPRT and against the Dilp5 peptide sequence RRDFRGVVDSCCRKS were generated as a service by Thermo Fischer Scientific Inc and have been previously characterized[123].

Quantification of Immunohistochemical images

All analyses were performed using ImageJ (version 1.8.0_112)[205]. IPC cell size, as visualized by membranal GFP expression (CD8-GFP), was measured using the freehand selection tool. 5 IPCs were measured per brain and a total of 7 brains were imaged. Absolute area values

are reported. Neurite complexity was measured on max Z-stacks of IPCs using the plot profile tool starting from the midline to the furthest visible neurite. Statistics were performed on the area under curve values. Dimmed intensity was determined in IPC nuclei as measured by mean intensity of nuclear Dimmed staining. IPCs were marked by CD8GFP expression and intensity values were relativized to control genotypes.

Size phenotypes

Adult flies 3–6 days after eclosion were anesthetized using chloroform to prevent desiccation and weighed on a Mettler Toledo XS204 scale (Wet Weight; d = 0.1 mg). Balanced siblings (*CyO*) were used as internal controls where possible. Otherwise, control and experimental genotypes were reared at controlled densities. To determine wing area, a single wing per fly was carefully removed using a fine dissection scissor and placed on a dissection slide. Wing length (L) was measured as a straight line from where the L3 vein meets the distal edge to the alula opening and wing width (W) was measured as a straight line starting from where the L5 vein meets the posterior edge to the anterior edge of the wing, perpendicular to L[208]. Total wing area (A) was then approximated by using the formula for an ellipse where A = π * (L/2) * (W/2). Wings were visualized with an SZX12 Olympus microscope with CX30 camera and images were analyzed with Cell^D software.

RNA extraction and qPCR

Adult flies were collected shortly after eclosion and maintained at low densities until 3-6 days old and snap-frozen in triplicate (10-15 heads and 3-5 bodies per biological replicate). Timed larvae were collected on dry ice (10-15 L1 larvae per biological replicate, 5-7 L2 larvae and 3-4 L3 larvae). For adults, RNA was extracted using the TRI Reagent (Invitrogen) as per protocol, larval RNA was extracted using Qiazol lysis reagent which is optimized for fatty tissues. 1 µg of RNA was used for cDNA synthesis with the SensiFAST cDNA synthesis kit (Meridian Bioscience, Bioline). qPCR was performed on a QuantStudio™ 12K Flex Real-Time or QuantStudio™ 6 Flex Real-Time PCR systems (Thermofisher Scientific) using the SYBR Green detection system (FastGene 2x IC Green mix-low ROX (Nippon Genetics)). Values were normalized using the geometric mean of *LaminC* and *Rp49* for adults and to *Rp49* in larvae. The ΔΔCt method was used to quantify transcripts relative to the controls.

Primers used were:

Gene Name	(5'-3') Forward	(5'-3') Reverse
Dilp2	AGCAAGCCTTTGTCCTTCATCTC	ACACCATACTCAGCACCTCGT
Dilp3	GCAATGACCAAGAGAACTTTGGA	GCAGGGAACGGTCTTCGA
Dilp5	GCTCCGAATCTCACCACATGAA	GGAAAAGGAACACGATTTGCG
LaminC	TCCACCCAACAATCTGGTGAT	GCAACATCCTCTTTGTCGGC
Rp49	GCCCAAGGGTATCGCAACA	CTTGCGCTTCTTGGAGGAGA
Dilp1	CAATGGCAATGGTCACGCCG	AGAGTATCCTGCCCCGTCC
Dilp6	TGCTAGTCCTGGCCACCTTGTTCG	GCTTCCCGAAACTGTTGGGAAATAC

Female fecundity

Recently eclosed females were collected and allowed to mate for 48 hours with sibling males in a 4 female to 1 male sex ratio. Females were then collected and placed in groups of 4–5 in separate vials. Flies were transferred onto fresh food daily and the number of eggs was counted for 21 days. Mean number of cumulative eggs laid was compared.

Survival analyses

Female flies were collected and aged until 4-7 days old, then transferred onto medium (1% agar, 10% Sucrose) containing 20mM Paraquat (Sigma 75365-73-0), an oxidative stress inducer. Deaths were scored regularly and plotted in GraphPad Prism. For starvation resistance, females were collected as described above and transferred to medium (1% agar). Deaths were scored twice daily.

TAG analysis

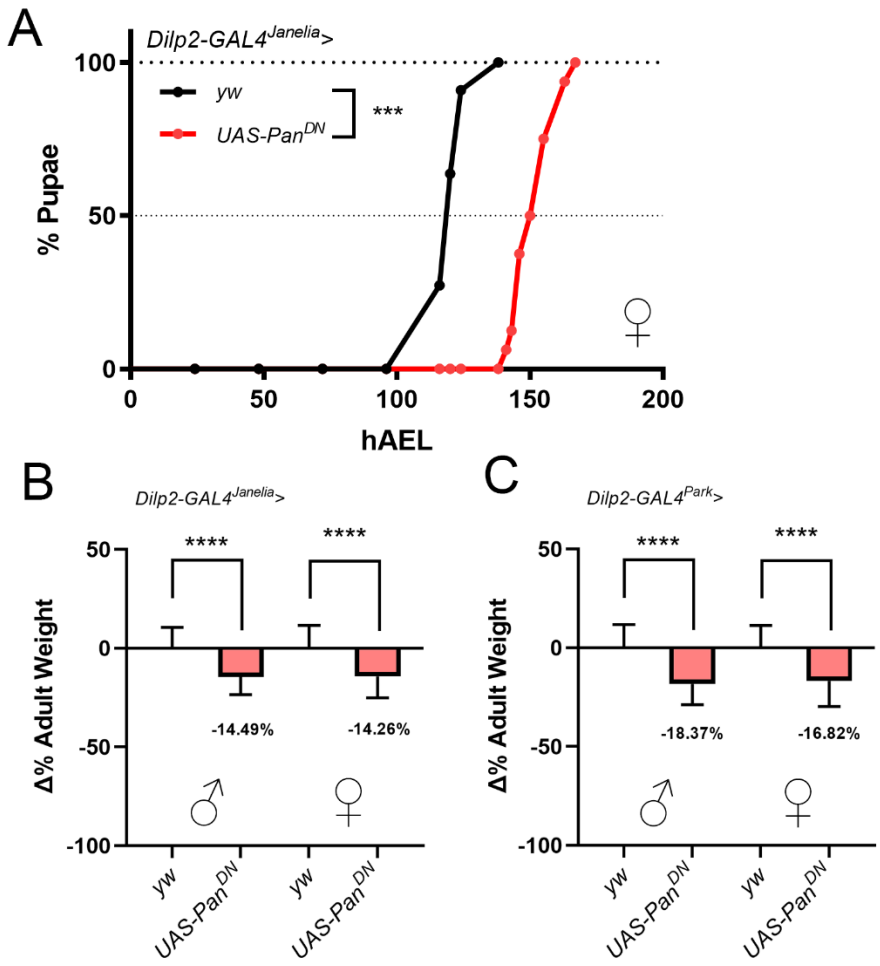
Groups of five 4–5 day old female flies, with ovaries and gut removed, were collected and snap frozen. Total TAG was determined using the sulfo-phospho-vanillin method modified from[207]. In brief, flies were homogenized in chloroform:methanol (2:1) and incubated overnight at 37°C with agitation. Samples were then placed in a 69°C heatblock to allow for chloroform:methanol evaporation. When fully evaporated, sulfuric acid (95–98%) was added and samples were placed on a heatblock for 10 minutes. Vanillin reagent (0.6% Vanillin, 34% phosphoric acid) was then added and samples were briefly mixed and incubated in the dark for 15 minutes. Samples were then transferred to clear bottom plates and measured at 540 nm on a Synergy H1 spectrophotometer (BioTek). A standard curve made with triolein was

used to determine TAG content in samples. TAG samples were normalized to the weight of the dissected flies.

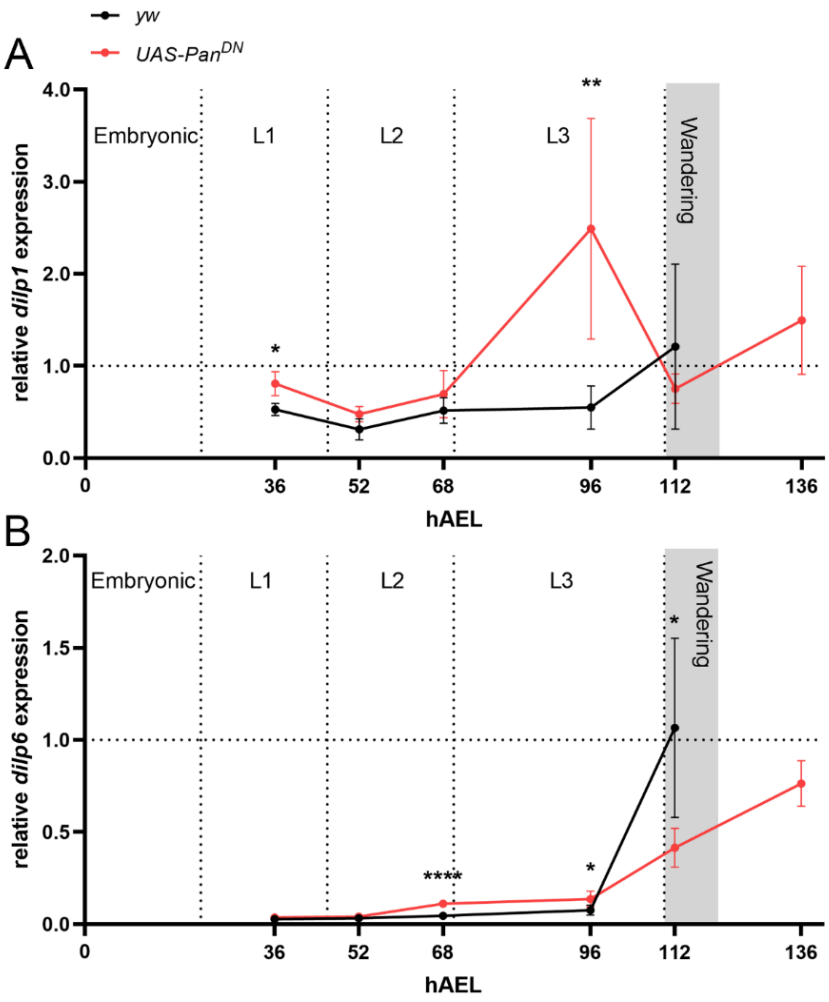
Statistics

All statistics were performed using GraphPad Prism. When comparing means of two groups a two-tailed unpaired Student's t-test was used. When more than two groups were compared a One-Way ANOVA with post-hoc Dunnett's tests was used to determine significant differences compared to the control group. For survival analyses median lifespan was compared using log-rank tests. A significance level (α) of 0.05 was used, p-values are reported as follows: not significant (ns) p>0.05, *p<.05, **p<0.01, ***p<0.001, ****p<0.0001).

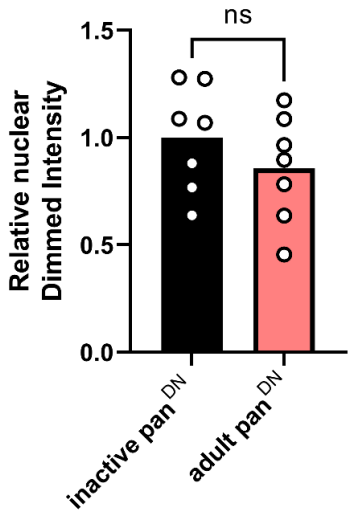
Supplementary information



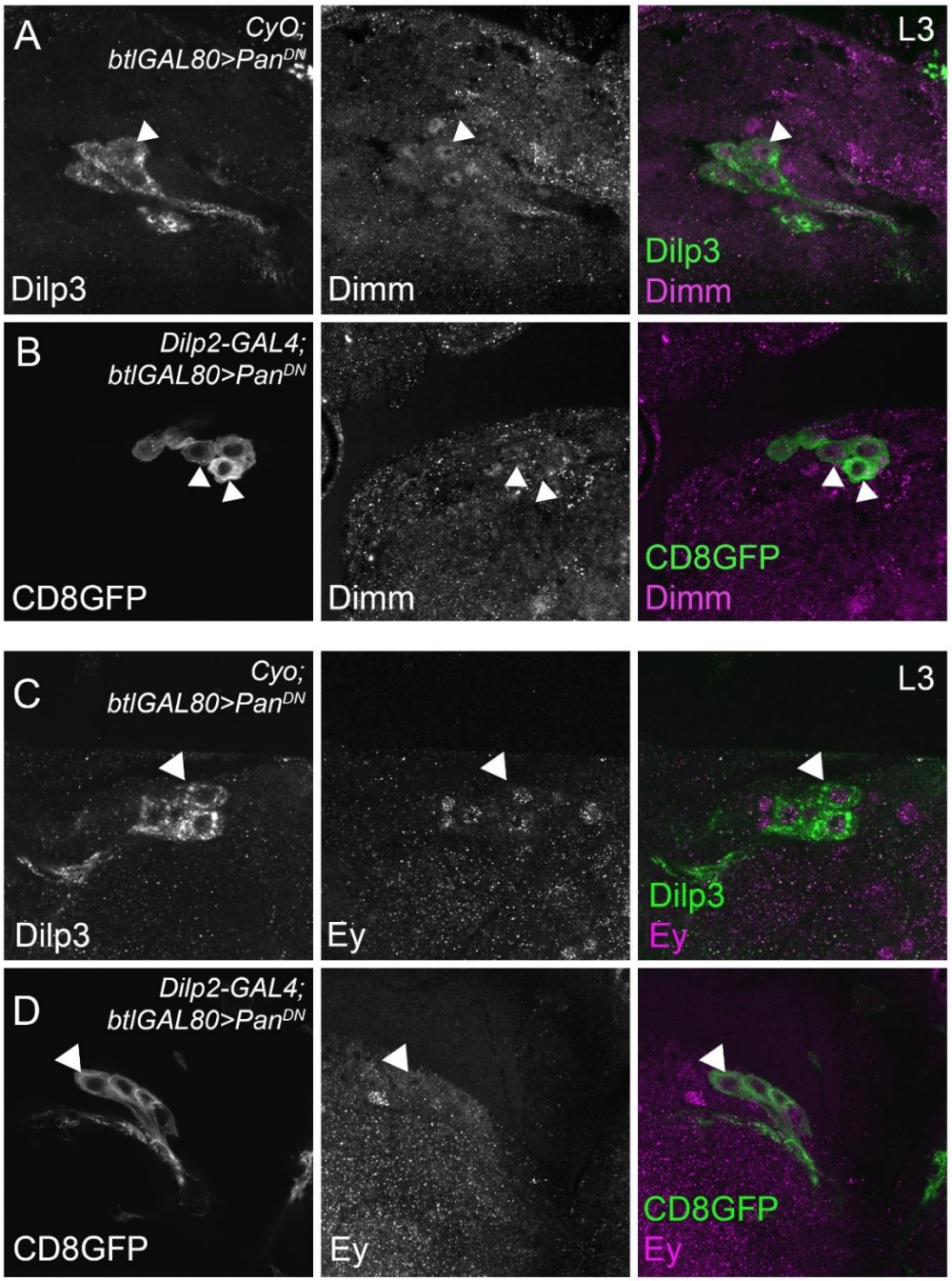
Supplemental figure 11. Developmental timing and adult weight defects upon Pan^{DN} expression in the IPCs. Dilp2-GAL4^{Janelia} -mediated Pan^{DN} expression extends time until pupariation with ~26 hours (unpaired t-test of mean time until pupariation of n = 3 vials with ~20 timed L1 larvae, p < 0.001). (B-C) Similar weight phenotypes were obtained using two different *Dilp2-GAL4* lines, *Dilp2-GAL4^{Janelia}* (control and experimental males n = 29, control females n = 23 and experimental females n = 28) and *Dilp2-GAL4^{Park}* (control males n = 23, experimental males n = 28, control females n= 19 and experimental females n =36). Unpaired t-tests, *** p<0.001 and ****p<0.0001.



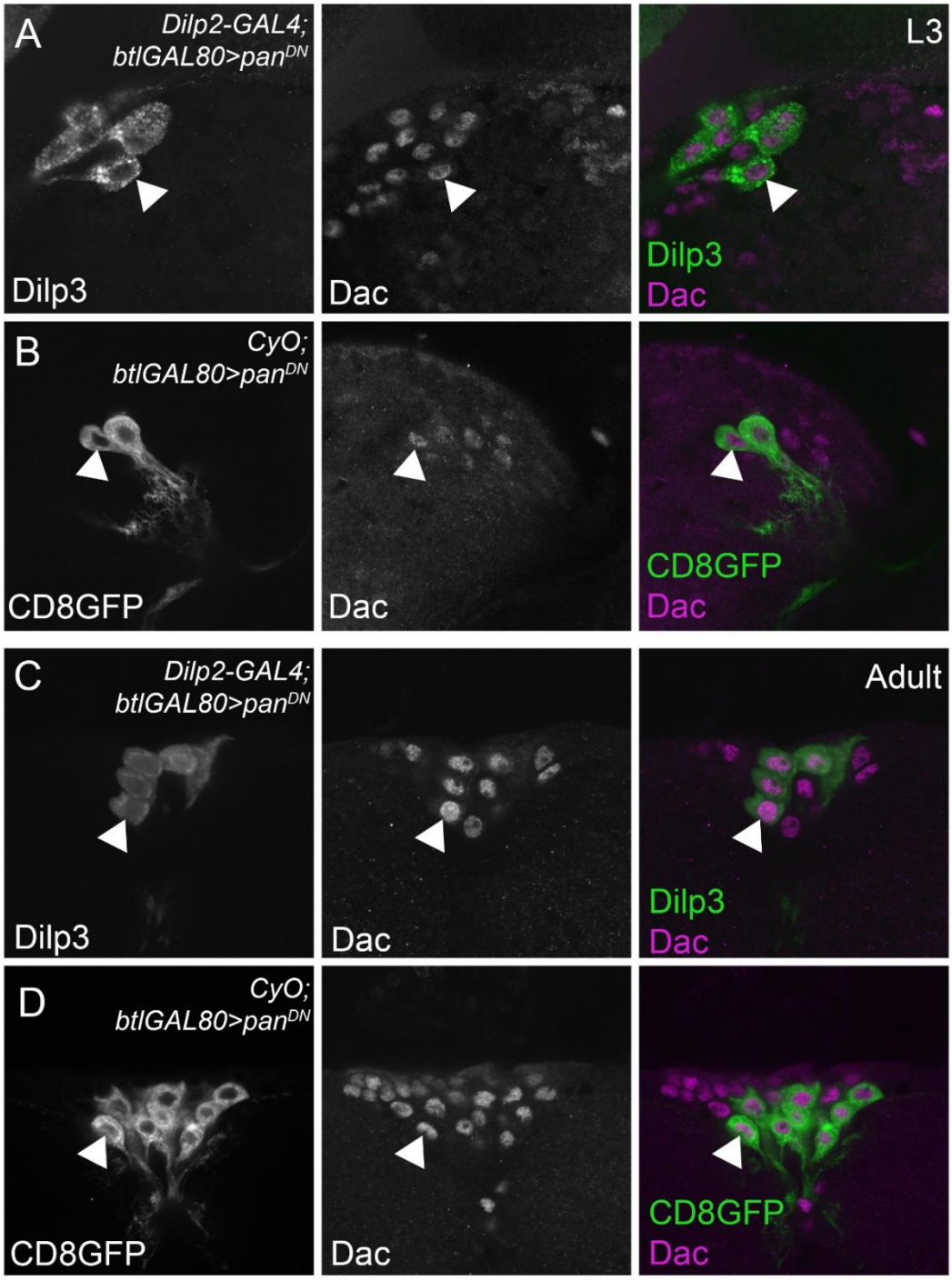
Supplemental Figure 12. Pangolin in the IPCs controls *dilp1* and *-6* expression during larval development. (A – B) Timed control and Pan^{DN} larvae (*yw/w; Dilp2-GAL4*, *UAS-CD8-GFP/+; btlGAL80/+ and yw/w; Dilp2-GAL4*,*UAS-CD8-GFP/+;btlGAL80/UAS-Pan^{DN})* were collected at 36, 52, 68, 96, 112, and for Pan^{DN} larvae at 136 hAEL for qPCR. Unpaired t-tests were used to compare mRNA levels at each timepoint between both genotypes and for the 136 hAEL Pan^{DN} timepoint with 112h hAEL control. The timing of developmental checkpoints of control larvae are shown. Relative quantities compared to the highest value for each dilp in the control larvae are reported. (Unpaired t-tests, *p<.05, **p<0.01, ****p<0.0001).



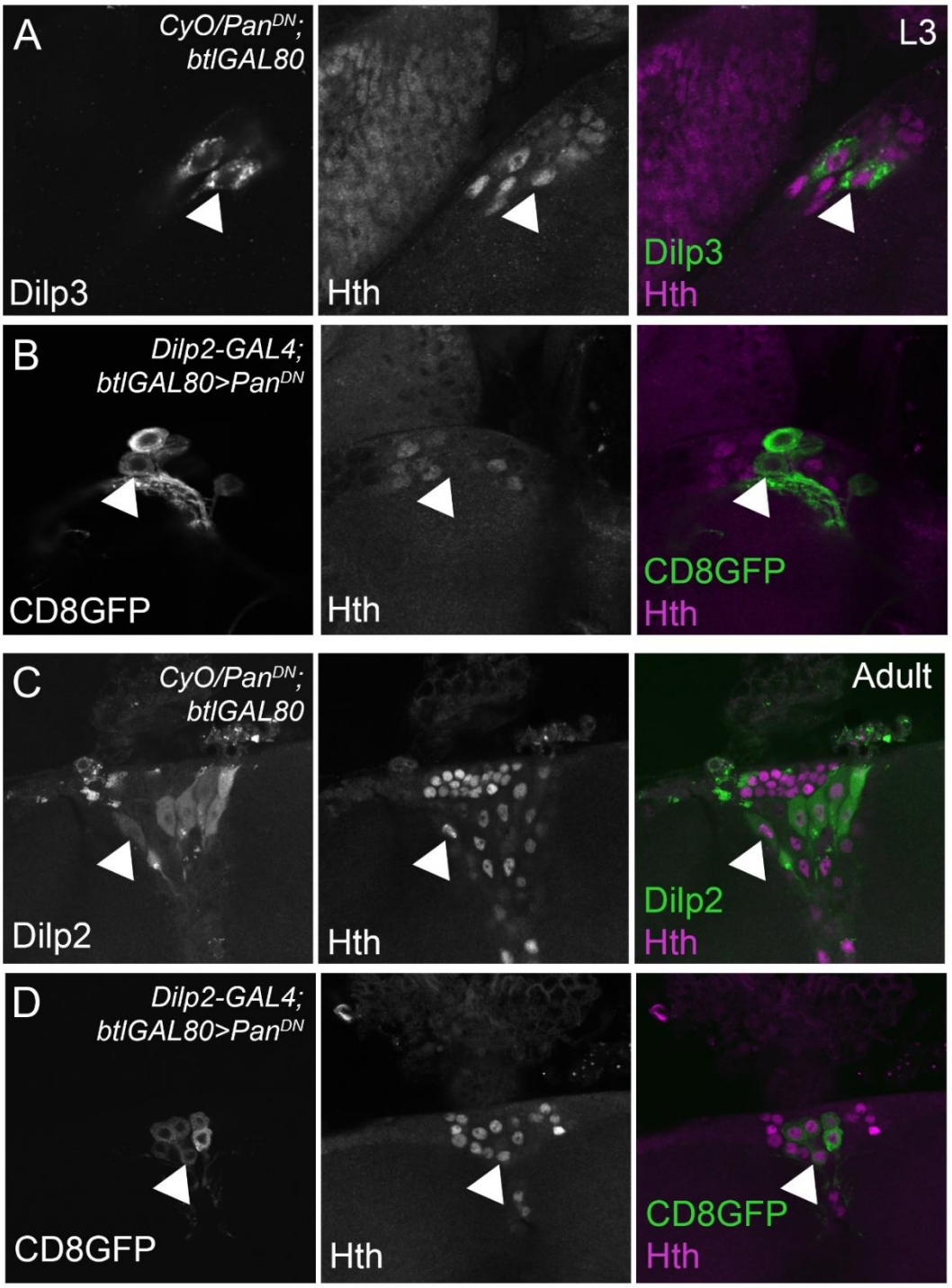
Supplemental Figure 13. Limiting Pan^{DN} activity to the adult stage leaves Dimmed levels unaffected. Dimmed nuclear intensity is unaffected. Temporal restriction was obtained using Gal80^{ts}. Unpaired t-test, ns, not significant.



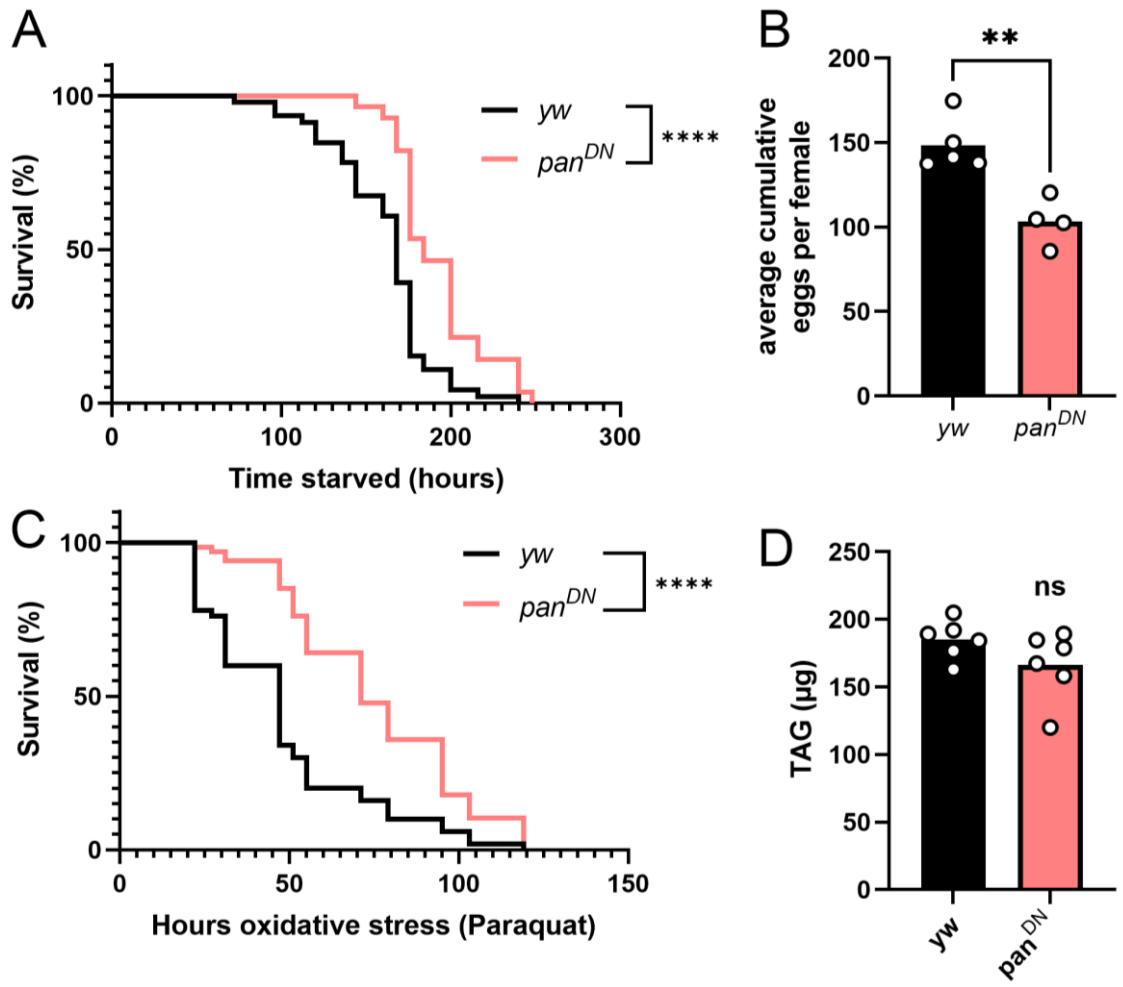
Supplemental Figure 14. Pan^{DN} reduces larval Dimm and Ey levels in the IPCs. Control Larval IPCs are Dimm^{+} (A) while Pan^{DN} overexpression results in reduced but not absent Dimm (B). Ey levels are undetectable in IPC nuclei (C -D). Test and control genotypes were stained in the same tube and were differentiated based on CD8-GFP expression. Only female flies were used. Representative images are shown, and arrows point to example nuclei.



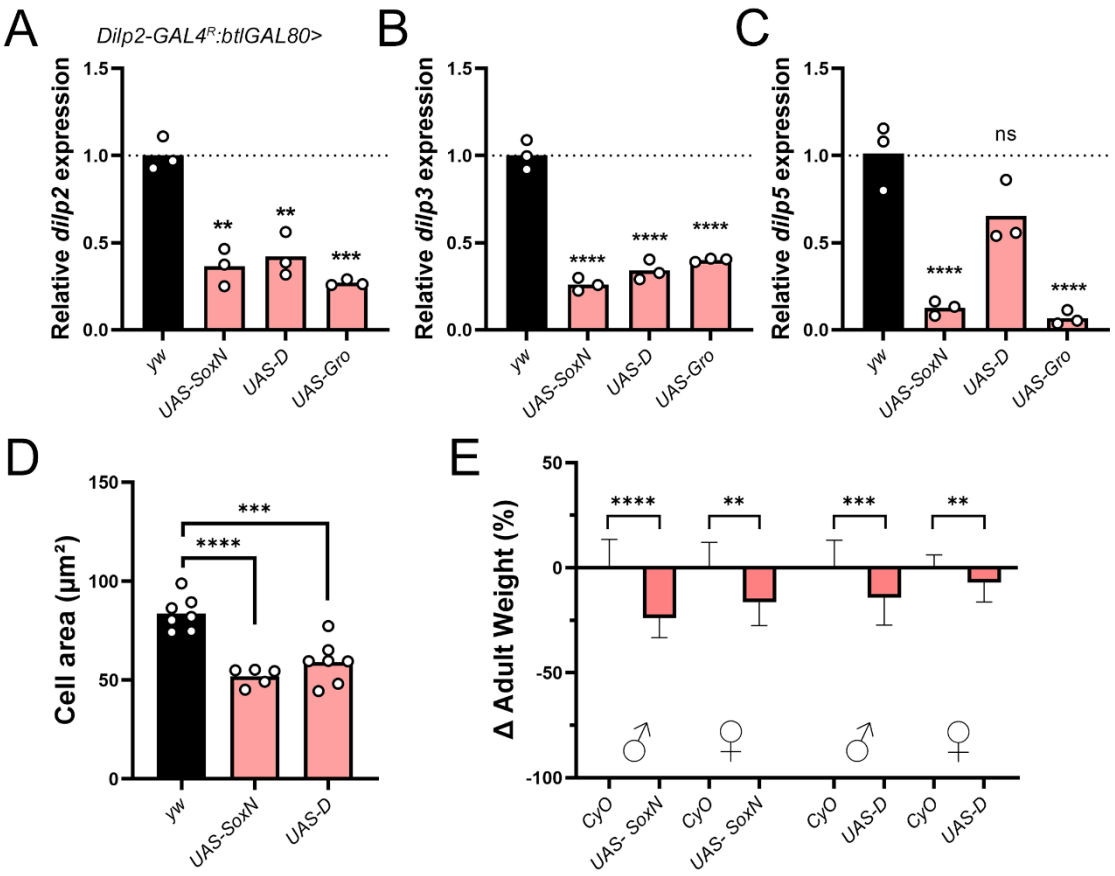
Supplemental Figure 15. The Ey binding partner Dac remains unaffected upon Pan^{DN} expression. Dac is detectable in both larval IPC nuclei upon Pan^{DN} (A - B) and in adult nuclei (C – D). Test and control genotypes were stained in the same tube and were differentiated based on CD8-GFP expression. Only female flies were used. Representative images are shown, and arrows point to example nuclei.



Supplemental Figure 16. Hth is depleted upon Pan^{DN} expression in larval IPCs only. Larval Hth levels are severely reduced by Pan^{DN} expression (A - B) while Hth remains detectable in adult control and Pan^{DN} IPC nuclei (C - E). Test and control genotypes were stained in the same tube and were differentiated based on CD8-GFP expression. Only female flies were used. Representative images are shown, and arrows point to example nuclei.



Supplemental Figure 17. Phenotypes associated with Pan^{DN} misexpression in the IPCs show that systemic IIS is reduced. (A) Pan^{DN} expression in the IPCs results in increased starvation resistance in female flies. Flies were reared on Nutrifly and were then transferred to 1% agar to induce starvation. Controls, n = 46, median lifespan = 168h; pan^{DN}, n = 28, median lifespan = 184h; Log-rank test p<0.0001. (B) Cumulative egg-laying between 2-23 days old flies is reduced (*yw*: n = 5; Pan^{DN}: n = 4, unpaired t-test p<.01, each data point is the average amount of egg laid of 4-5 flies). Oxidative stress resistance is increased in females which was induced by the presence of 20mM Paraquat in 10% sucrose, 1% agar medium. Controls, n = 50, median lifespan = 47h; Pan^{DN}, n =76, median lifespan = 71h; Log-rank test p<0.0001. (D) Total TAG of female flies is not altered (unpaired t-test, ns, not significant n = 6, with each data point the average of 5 flies)



Supplemental figure 18. Pangolin repressor complex co-factors phenocopy Pan^{DN} in the IPCs. (A-C) Overexpression of SoxN, D and Gro in the IPCs using *Dilp2-GAL4^{R};btlGAL80* significantly reduces IPC *dilp2*, *-3* and *-5* expression, except for D on *dilp5* expression where a trend but no statistical significance was observed (One-way ANOVA with post-hoc Dunnett's test compared *yw* controls (** p <0.01, *** p<0.001, **** p < 0.0001, ns: not significant). (D) IPC cell area was significantly decreased upon SoxN and Gro overexpression (One-way ANOVA with post-hoc Dunnett's test compared *yw* controls (*** p<0.001, **** p < 0.0001). (E) Adult weight is significantly reduced upon SoxN (n = 14 and 13, males and females) and D (n = 24 and 21) overexpression compared to balanced sibling controls (*CyO*, n = identical to experimental condition).

5 Deciphering the transcription factor code regulating expression of the *Drosophila dilp2*, *-3* and *-5* genes.

Winant M[1], Wilms T[1], Hens K[2], Callaerts P[1]

1 Laboratory of Behavioral and Developmental Genetics, Department of Human Genetics, KU Leuven—University of Leuven, Leuven, Belgium.
2 Centre for Functional Genomics, Department of Biological and Medical Sciences, Faculty of Health and Life Sciences, Oxford Brooks University, Headington Campus, Oxford, United Kingdom

Abstract

The insulin-producing cells in the pars intercerebralis of the *Drosophila* brain are central players in the regulation of metabolism, growth, lifespan, and fecundity. These cells produce *Drosophila* insulin-like peptides *2*, *-3* and *-5* (*dilp2*, *-3* and *-5)*. The transcript levels of *dilp2*, *-3* and *-5* vary throughout development and respond to a broad range of stimuli. Their expression depends on distinct cis-regulatory elements consistent with *dilp2*, *-3* and *-5* expression being regulated independently. However, the specific transcription factors involved remain largely unidentified to date. Here, we perform a Yeast-1-hybrid screen using the proximal promoter regions of *dilp2*, *-3* and *-5* as bait. This identified 39 direct regulators of which most were associated with only one *dilp*. We further validated the role of these TF in *dilp* regulation using transgenic approaches. We provide evidence for the direct transcriptional control of *dilp2* by Nuclear factor Y-box B, E2 transcription factor 2, Hormone Receptor 78 and Twin of eyeless (Toy); of *dilp3* by Scalloped and of *dilp5* by Achintya and Brain-specific homeobox. Several of these transcription factors also affect transcript levels of other *dilps*. We demonstrate that Toy is a direct regulator of *dilp2* expression. In summary, this study identifies the transcription factors necessary for independent *dilp* regulation, highlights the compensatory and interdependent nature of *dilp* expression and shows that a second PAX6 homolog is necessary for *dilp2* expression. Given that the vertebrate homologs of many of the *Drosophila* transcription factors have previously been implicated in pancreatic development or function, our results are also discussed in an evolutionary perspective.

Introduction

The *Drosophila* insulin-like peptides (Dilps) play major roles in the regulation of metabolism, growth, lifespan and fecundity[89–91,93]. A total of 8 *dilps* are encoded in the *Drosophila* genome. They all contain the spaced cysteine residues typical for insulin-related proteins[3]. *dilp1-6* are considered true insulin/insulin-like growth factor homologs whereas *dilp7* and *-8* are relaxin-like[248]. They have a complex spatiotemporal expression pattern which indicates that they are under different transcriptional control[3]. *dilp2*, *-3* and *-5* are all expressed in the *Drosophila* Insulin-Producing Cells (IPCs), neurosecretory cells located in the pars intercerebralis of the brain, with additional expression of *dilp2*, *-3* and *-5* being detected in other cell types. *dilp2* expression has been observed in salivary glands, imaginal discs and tracheal cells[3,123]. *dilp3* has been detected in the visceral muscles of the midgut where it is involved in intestinal stem cell proliferation[124]. *dilp5* is also expressed in the Malpighian tubules and in ovarian follicle cells[3,125].

dilp2, *-3* and *-5* show a specific temporal pattern of expression in the IPCs. Expression of *dilp2* starts late in embryonic development, followed by *dilp5* in late L1 and ending with upregulated *dilp3* during the wandering larval stage, prior to pupariation[99]. Throughout the adult stage, the IPCs express *dilp2*, *-3* and *-5*[92]. This temporal control is further regulated by the physiological status of the fly. Short Neuropeptide F, for example, stimulates *dilp2* and *-5* expression in response to sugar feeding[145]. This is but one example of a plethora of non-autonomous signals that have been identified that stimulate or inhibit *dilp2*, *-3* and *-5* transcription in a specific manner[249]. Despite the important role of Dilps in metabolic homeostasis, growth, lifespan and fecundity, the transcription factor code controlling their expression is unknown. Direct regulation has so far only been shown for *dilp5* by Eyeless (Ey) and Dachshund (Dac)[149,150]. Other transcription factors, like lame duck (Lmd), Sugarbabe (Sug) and dFOXO have been implicated as putative regulators of *dilp* expression in the IPCs[99,151,190], but for none has a direct regulation been shown. Work from our lab previously showed that Homothorax, together with Extradenticle is essential for *dilp2* and *-3* expression while other transcription factors of the Retinal Determination Gene Network are involved in *dilp2* and *-5* regulation[229]. However, also for these transcription factors it is not known whether they control *dilp* expression in a direct manner. Characterizing these direct interactions would give us a better insight into the regulatory landscape of the IPCs and is also required for further studies of the links between non-autonomous signals controlling *dilp* transcription and the relevant transcription factors.

The study of DNA-Protein interactions can be tackled in several ways. One powerful and high-throughput tool is the yeast one-hybrid (Y1H) system, which enables the analysis of protein-

DNA interactions in a eukaryotic cellular environment[250]. Here, we screened a library of 700 *D. melanogaster* transcription factors for interactions with the regulatory regions of *dilp2*, *-3* and *-5* using an established high-throughput Yeast-1 hybrid system[251,252]. This screen yielded 33 unique putative direct transcriptional regulators. We further validated and characterized the role of these transcription factors using transgenic approaches and discovered several transcription factors that are likely involved in direct *dilp* regulation in the IPCs. We also found that many of the identified direct regulators had additional effects on the expression of one or two other *dilp* genes. We previously showed that the PAX6 homolog *eyeless* directly regulates *dilp5*. The identification of a possible role for the second PAX6 homolog *twin-of-eyeless* (*toy*) prompted us to characterize this further. This revealed that not only is Toy a direct regulator of *dilp2* expression, but it also impacts IPC differentiation and function.

Results

We screened 700 *Drosophila* transcription factors that were fused to the GAL4 activation domain in yeast strains containing *dilp2*, *-3* or *-5* regulatory sequences located upstream of the transcription starting site as bait (Figure 25 and Supplemental table 1). This screen yielded 9 positive hits for *dilp2*, 27 for *dilp3* and 12 for *dilp5* regulatory elements. Out of this total of 48, 39 unique transcription factors were detected. Three transcription factors were detected in all three *dilps*: *Rfx*, *Checkpoint suppressor-like 1* (*CHES-1-like)* and *wornia* (*wor*), one was detected in both *dilp2* and *dilp5*: *engrailed* (*en*), two in both *dilp3* and *dilp5*: *odd paired* (*opa*) and *brain-specific homeobox* (*bsh)*, and none in both *dilp2* and *dilp3* (Table 1). Of the identified transcription factors, *sug* and *lmd* have previously been associated with *dilp* regulation in the IPCs[151,190].

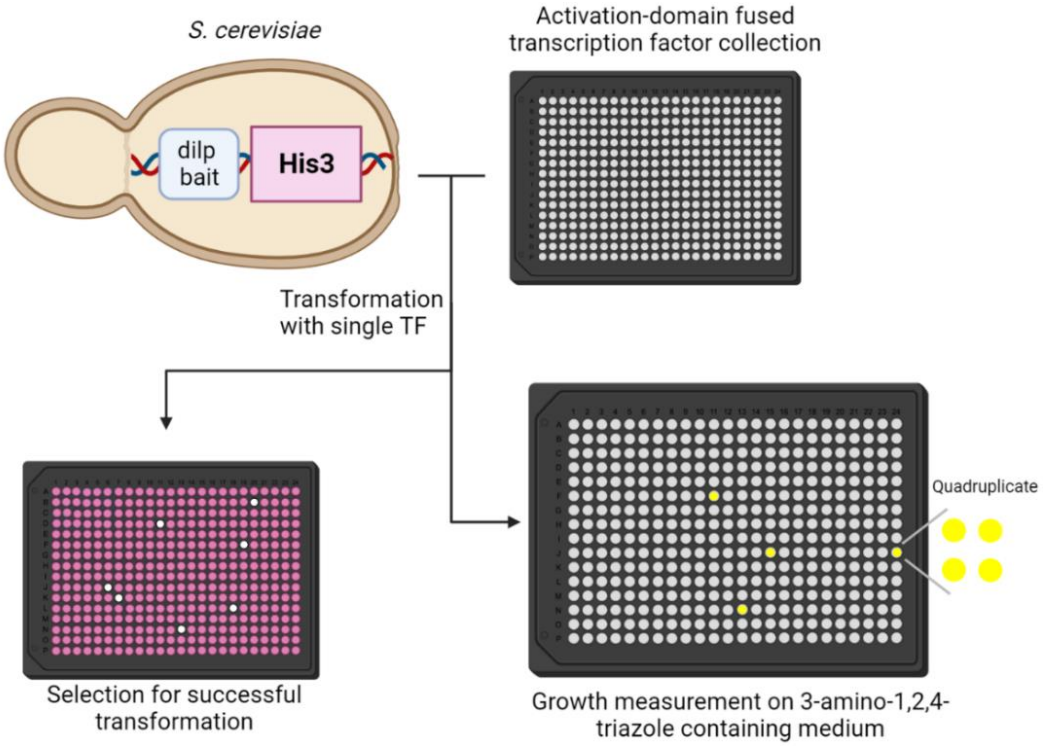


Figure 25. Workflow of Yeast-1-hybrid approach. A yeast strain containing regulatory regions for *dilp2*, *-3* or *-5* and the coding sequence of *His3* was distributed over a 384-well plate. Each well of this plate was then transformed with a different GAL4 activation domain-fused transcription factor clone. A total of 700 Transcription factors were screened. Transformed yeasts were checked by growing on a permissive plate containing SC –His –Trp medium and were separately spotted in quadruplicate on the selection plate containing 3-amino-1,2,4-triazole. DNA-transcription factors interactions were based on growth on this medium as expression of His3 allows for growth on Histidine deficient and 3-amino-1,2,4-triazole containing medium. *Figure made with BioRender.*

	TF	Positive hits *dilp2/5* screen (%)	Occurrence in all screens (%)		TF	Positive hits *dilp3* screen (%) A	B	C	Occurrence in all screens (%)
Dilp2	*E2f2*	50	1.1	*Dilp3*	*Max*		50	100	1.1
	Nf-YB	50	1.1		*CG14710*	75			1.1
	Hr78	75	4.3		*CG6683*	75			1.1
	en	75	6.4		*CG18764*	50	25		1.1
	drm	50	8.5		*HP1b*		50		1.1
	toy	50	11.7		*CG11902*		50		1.1
	CHES-1-like	50	44.7		*ara*			50	1.1
	wor	50	46.8		*CG7786*	50			2.1
	Rfx	100	87.2		*CkIIα-i1*			50	2.1
					Foxp			50	4.3
Dilp5	*vis*	50	1.1		*bap*	75		25	5.3
	nerfin-1	50	1.1		*sug*	75	75		6.4
	achi	75	6.4		*TfAP-2*	100		100	6.4
	en	50	6.4		*croc*	25		50	6.4
	CG4318	100	22.3		*CG32778*	75			9.6
	CG3526	50	22.3		*lmd*	50	25	50	24.5
	opa	75	28.7		*opa*	100		100	28.7
	Side	50	29.8		*CG11902*	100			29.8
	CHES-1-like	100	44.7		*sna*	50	75	50	29.8
	wor	50	46.8		*fkh*	50	25	25	33.0
	bsh	50	52.1		*sd*	50		50	35.1
	Rfx	100	87.2		*Hr38*	100	25	25	43.6
					CHES-1-like	100	50	100	44.7
					wor	50	50	50	46.8
					nub	100	50	25	47.9
					bsh	100	25	75	52.1
					Rfx	100	75	100	87.2

Table 1. Identified transcription factors in Yeast-1-hybrid screen. Gene names of all transcription factors that bind *dilp2*, *-3* or *-5* regulatory regions. Transcription factors are listed that allowed for successful growth on 3-amino-1,2,4-triazole medium in two independent integrations at least once where each integration was screened two times. 100% therefore means that it was detected 4 out 4 times. For *dilp3*, 'A' is the full regulatory element while 'B' and 'C' are overlapping smaller elements. Occurrence in all screens shows how often this transcription factor is detected in all other performed screens. A high percentage thus suggests a higher probability of being a false positive.

To narrow down the number of putative regulators, we first identified those transcription factors that are expressed in the IPCs by comparing this list with published transcriptomic data of both larval and adult IPCs[103,229]. A threshold of 5 transcripts per million (TPM) was used as a cut-off (Figure 26 A). Five transcription factors, *E2 transcription factor 2* (*E2f2*), *Nuclear factor Y-box B* (*Nf-YB*), *toy*, *CHES-1-like* and *Hormone receptor* 78 (*Hr78*) remained for further validation using RNAi-mediated knockdown. This showed that all but CHES-1-like knockdown resulted in a significant downregulation of *dilp2* mRNA levels (Figure 26 B). Follow-up experiments measuring Dilp2 on the protein level using immunohistochemistry showed that only *toy* knockdown significantly reduced Dilp2 accumulation in the IPCs (Figure 26 C -G). Given that Dilp2 is a major regulator of growth, we wondered if reduced dilp2 levels would result in smaller flies, but no significant reductions in adult weight were detected (Figure 26 H). We additionally measured *dilp3* and *-5* mRNA levels which revealed that knockdown of Hr78 and E2f2 results in the upregulation of *dilp3* while *dilp5* is mildly reduced (Figures S19 A - B). While *toy* knockdown slightly reduces *dilp5* levels in males only (Figure S19 B). It is, however, noteworthy that knockdown usually did not appear to result in sexually dimorphic effects and that downregulation of *dilp2* frequently co-occurred with a simultaneous reduction in *dilp5* and an upregulation of *dilp3*.

A similar approach for *dilp3* left us with 11 putative regulators (Figure 27 A) of which we evaluated nine for a role in *dilp3* regulation in the IPCs. Strikingly, only significantly increased *dilp3* levels were observed for several transcription factors: *Max*, *scalloped* (*sd*), *Forkhead box* (*FoxP*), *CG6683* and *CHES-1-like*), suggesting that these are negative regulators of *dilp3* expression (Figure 27 B). Two separate RNAi lines were used for *CHES-1-like*, *FoxP* and *CG6683* which had inconsistent effects on *dilp3* expression. Dilp3 protein intensity was additionally determined in *Max* knockdown, but this did not significantly alter Dilp3 levels in the IPCs (Figure 27 C). We further measured *dilp2* and *-5* levels which mostly showed reductions in both dilps upon knockdown of *FoxP*, *sd* and *bap* (Figure S20 A-B). *CG6683* knockdown had a male specific-effect resulting in the upregulation of *dilp2*. Knockdown of *fkh* and *Max* reduced *dilp5* levels in both sexes (Figure S20 B)

Lastly, from the identified *dilp5* regulators only 1 transcription factor, *CHES-1-like*, reached the expression level threshold in L3 or adult IPCs (Figure 28 A). We therefore opted to include two additional putative regulators, *bsh* and *achintya* (*achi*) that were also expressed in the adult IPCs albeit at lower levels (~1 TPM). RNAi-mediated knockdown showed that all but *achi*$^{RNAi\ \#2}$ and *CHES-1-like*$^{RNAi\ \#1}$ in females significantly reduced *dilp5* levels (Figure 28 B). Dilp5 protein intensity was, however, not altered in knockdown with *achi*$^{RNAi\ \#1}$ or *bsh*^{RNAi} (Figure 29 C - D). Knockdown with *achi*$^{RNAi\ \#2}$ slightly increased *dilp2* levels in both sexes and *CHES-1-like*$^{RNAi\ \#1}$ increased *dilp3* levels (Figure S21 A – B).

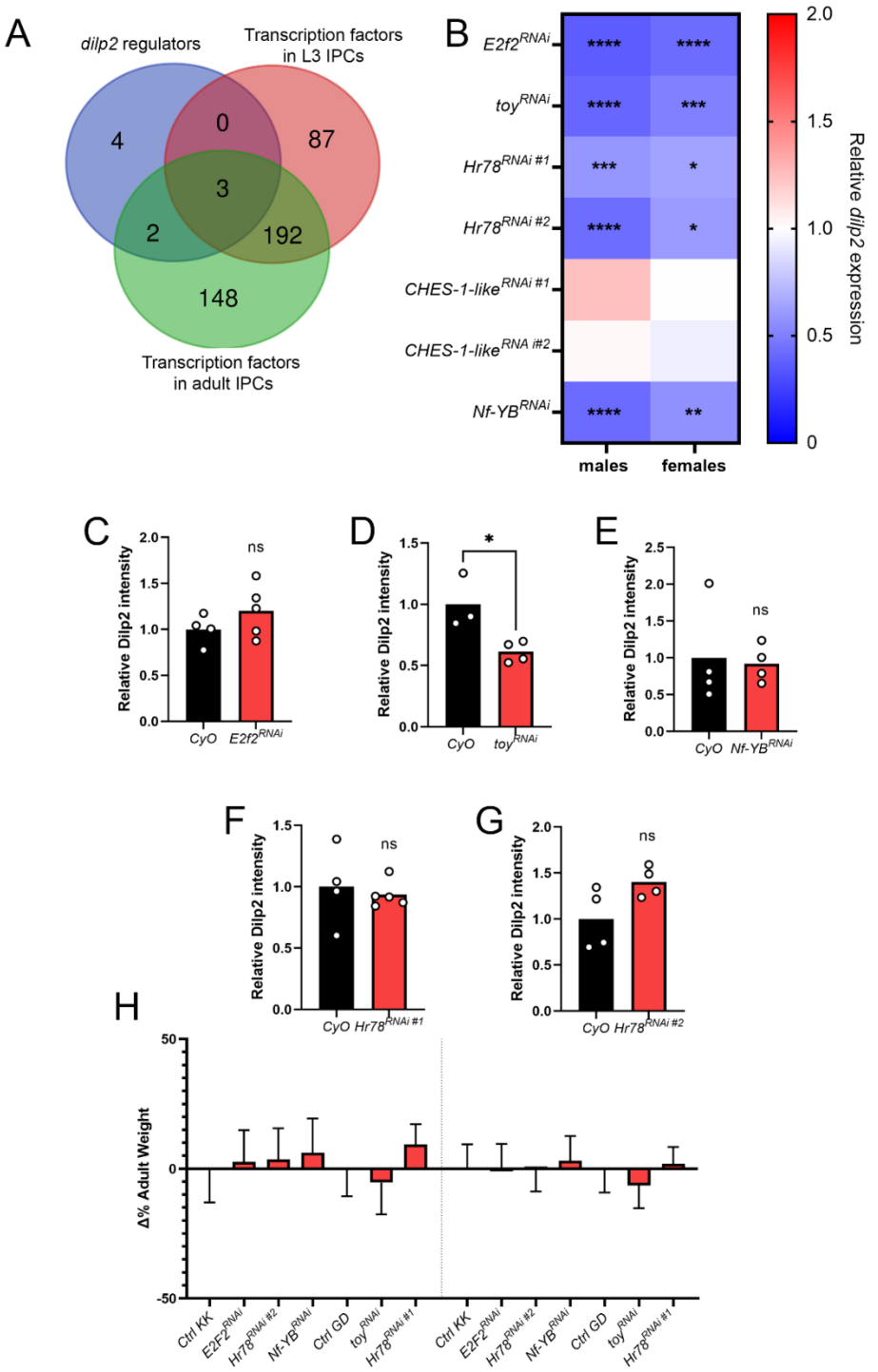


Figure 26. Validation of *dilp2* regulators reveals strong positive regulators of *dilp2* expression. (A) Overlap of 9 putative *dilp2* regulators with transcription factors expressed in larval and adult IPCs (TPM>5 as a cut-off). (B) RNAi-mediated knockdown of putative regulators shows that *E2f2*, *toy*, *Nf-YB* and *Hr78* positively regulate *dilp2* expression (* p < 0.05, ** p <0.01, *** p<0.001, **** p < 0.0001, One way ANOVA with post-hoc Dunnett's compared with KK or GD controls; n = 3-6 biological replicates per genotype per sex of > 8 heads). (C-G) Immunohistochemistry with specific antibodies against Dilp2 revealed that only knockdown of *toy* significantly reduced Dilp2 staining intensity. (H) No significant differences in adult weight were observed, one-way ANOVA with post-hoc Dunnett's test compared to one of two controls (KK ctrl or GD ctrl). Males left and females right.

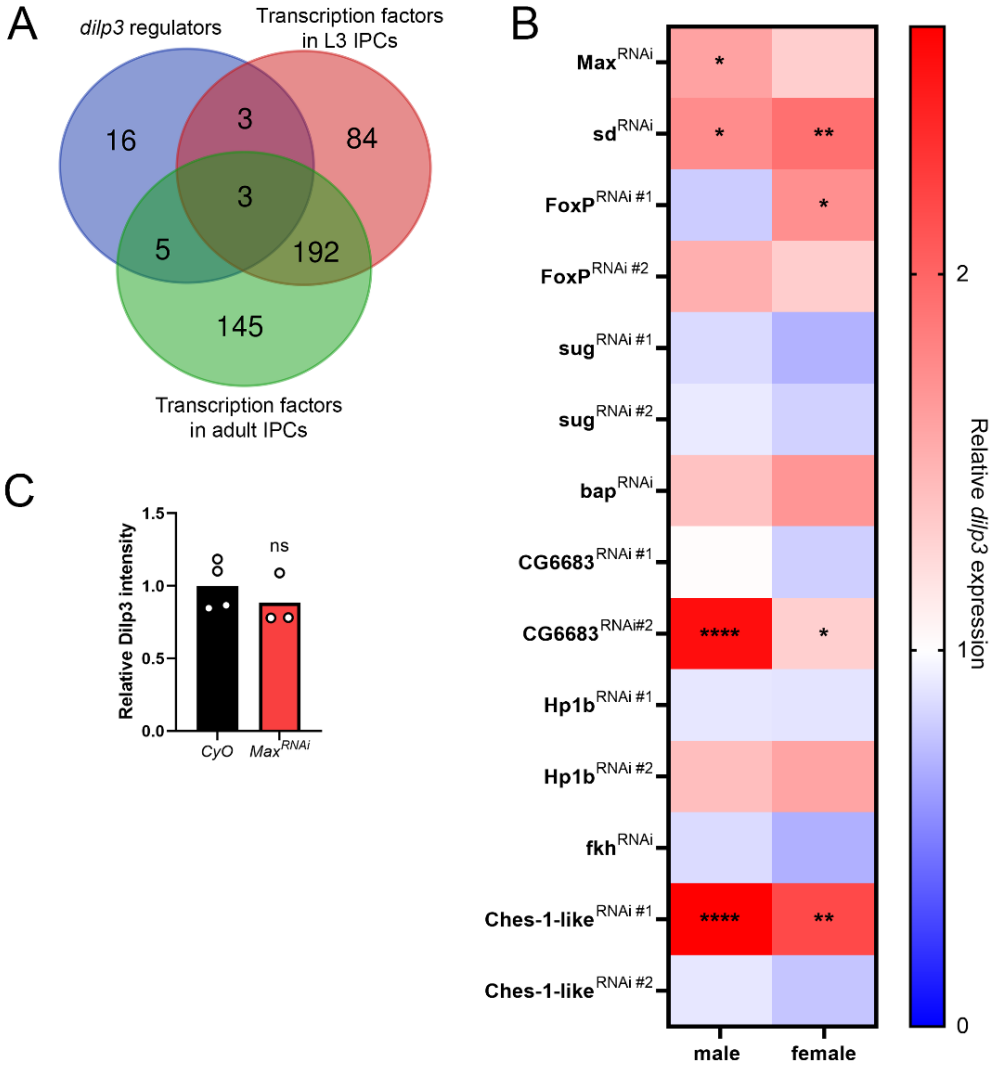


Figure 27. Validation of *dilp3* regulators reveals many negative regulators of *dilp3* expression. (A) Overlap of 29 putative *dilp2* regulators with larval and adult IPC expressed transcription factors (TPM>5). (B) RNAi-mediated knockdown of *Max*, *sd*, *FoxP*, *CG6683* and *CHES-1-like* significantly increased *dilp3* expression (* p < 0.05, ** p <0.01, *** p<0.001, **** p < 0.0001, One-way ANOVA with post-hoc Dunnett's compared with KK or GD controls; n = 3-6 biological replicates per genotype per sex of > 8 heads. (C) Immunohistochemistry with specific antibodies against Dilp3 revealed that only knockdown of *Max* did not significantly reduce Dilp3 staining intensity.

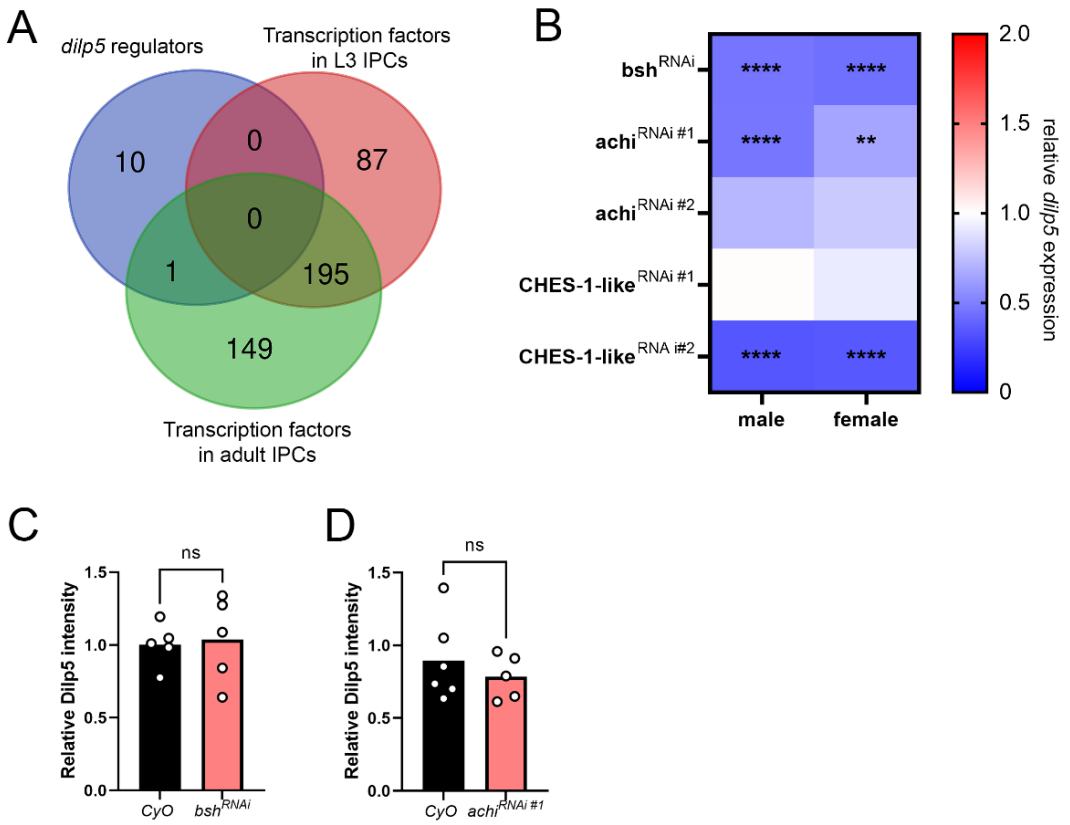


Figure 28. Validation of putative *dilp5* regulators identifies positive regulators of *dilp5* expression. (A) Overlap of 13 putative *dilp5* regulators with larval and adult IPC expressed transcription factors (TPM>5). Only *CHES-1-like* reached this threshold, two additional regulators were selected (*bsh* and *achi*) (B). Knockdown of *bsh* and *achi* resulted in significant downregulation of *dilp5* (* p < 0.05, ** p <0.01, *** p<0.001, **** p < 0.0001, One way ANOVA with post-hoc Dunnett's compared with KK or GD controls; n = 3-6 biological replicates per genotype per sex of > 8 heads). (C-D) Immunohistochemistry with specific antibodies against Dilp5 revealed that knockdown of *bsh* or *achi* did not significantly reduce Dilp5 staining intensity in the IPCs.

Toy is one of two PAX6 homologs, the other being Ey which has previously been shown to directly regulate *dilp5*[149,253]. The finding that Toy could directly regulate *dilp2* is therefore of considerable interest. We therefore further elaborated on these results by employing a dominant negative form of Toy (Toy^{DN}), which consists of Toy fused with the repressor domain of Engrailed. Overexpression of two independent Toy^{DN} lines in the IPCs resulted in significantly smaller flies, a phenotype previously shown to be associated with reduced *dilp2* expression (Figure 29 A-B)[89].

We further validated the expression of *toy* in adult IPCs using immunohistochemistry (Figure 29 C)[229]. We then evaluated *dilp2*, *-3* and *-5* expression upon Toy^{DN} overexpression and found that this strongly reduced *dilp2* levels (D), reduced *dilp5* levels by half (J) and had no effect on *dilp3* (Figure 29 D, J and G)). On the protein level, we found that Dilp2 and -5 were strongly reduced (Figure 29 E, F, K and L) and that Dilp3 was still present in the IPCs perinuclearly but

that staining intensity in median bundle which transports Dilp protein to the release sites is highly reduced (Figure 29 H - I). Morphological examination also revealed the presence of what we have previously termed ectopic pioneer neurites deriving from the median bundle[108] (arrows in Figure 29 F, I, L).

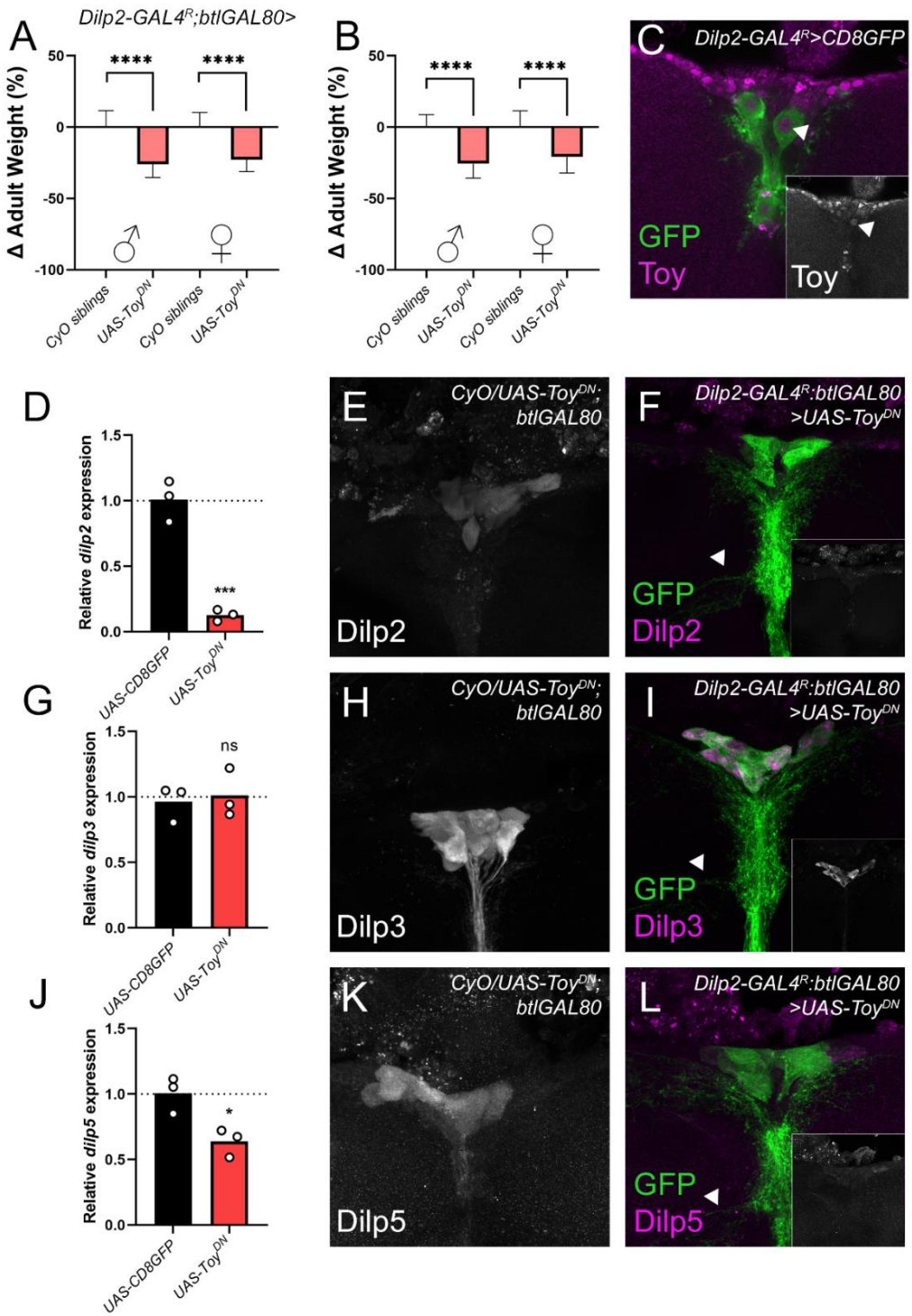


Figure 29. Toy regulates *dilp2* expression in the IPCs and a Toy^{DN} perturbs IPC differentiation. (A-B) Misexpression using two independent Toy fused with engrailed repressor domains (*UAS-Toy^{DN}*) results in significantly smaller flies with reduced body weight (unpaired t-test, **** p<0.000; n>20 for each condition, individual flies weighed). (C) Antibodies directed against the Toy paired domain readily detect it in adult IPC nuclei. Toy^{DN} highly reduces *dilp2* mRNA (D) and Dilp2 protein (E - F) levels. mRNA levels of *dilp3* and Dilp3 protein levels are not altered (G - I). Expression of *dilp5* is mildly reduced (J) and Dilp5 protein can no longer be detected in IPC somata (K – L). (* p < 0.05, *** p<0.001, ns: not significant, unpaired t-test; three replicates of >8 female heads). Representative images taken at identical confocal settings are shown for and insets (F, I, L) show Dilp2, -3 and -5 in grayscale. Brains from female flies were dissected and stained in the same tube and CD8GFP expression was used to distinguish experimental genotypes (*Dilp2-GAL4^{R}; btlGAL80 > UAS-Toy^{DN}*) from balanced sibling controls (*CyO/UAS-Toy^{DN}; btlGAL80*). Arrows point ectopic pioneer neurons (F, I, L).

Discussion

Regulation of *dilp* expression is both temporally and spatially dynamic but their transcriptional control has been understudied. In this study, we performed an exhaustive screen of the *dilp2*, *-3* and *–5* proximal promoters with 700 *Drosophila* transcription factors and performed RNAi-based validation on putative direct regulators. This is the first time such a comprehensive study has been done to identify direct regulators of *dilp* expression. The identified regulators were largely non-overlapping and those that were common were usually associated with a high hit rate in all screens, suggestive of a high false positive rate. These results thus suggest that dilp regulation is largely independent of each other (Table 1).

Of the transcription factors identified in the Yeast-1 Hybrid screen to regulate *dilp2* we provide evidence that E2f2, Nf-YB, Toy and Hr78 are direct regulators (Figure 26 B). E2f2 is a transcriptional repressor, together with Deadpan and is best known to suppress the expression of differentiation-related genes during cell cycle progression[254]. It is thus surprising that depleting E2f2, a repressor, reduces *dilp2* expression. Possibly, the observed effect is not direct but indirect via an additional role in IPC development. Nf-YB, with Nf-YA and Nf-YC, is involved in eye and thorax development in *Drosophila* and recent work has shown that its homolog is necessary for normal beta cell homeostasis[255–257]. Hr78 was identified as a positive regulator of *dilp2* and had additional effects on *dilp3* and *-5*. Hr78 plays an essential role in tracheal molting in third instar larvae and was long thought to be ecdysone-dependent but remains an orphan receptor as of yet[258].

Lastly, we found that knockdown of Toy reduces *dilp2* mRNA levels which is consistent with previous findings from our lab[229]. This is especially interesting given the role of its paralog, Ey, in *dilp5* regulation and IPC development, and the role of PAX6, the single mammalian homolog of *ey* and *toy*, in the direct transcriptional control of insulin expression and in beta cell development[231]. Direct regulation of *dilp5* by Ey occurs through its paired domain which differs in one amino acid from the paired domain of Toy[259]. This single substitution is sufficient to significantly alter which cis-regulatory elements are bound by Toy[253]. Ey knockdown also

slightly reduces *dilp2* and Toy knockdown seems to slightly reduce *dilp5* levels (Figure S19 B)[229]. Toy and Ey could possibly regulate *dilp2* and *-5* through the same cis-regulatory elements but with different affinities with Toy being more important for *dilp2* and Ey for *dilp5.* Alternatively, these effects can be secondary due to the role of Ey and Toy in IPC differentiation[260].

In addition to the direct regulation of *dilp5*, Ey is also involved in IPC morphological and functional differentiation with effects on IPC cell body clustering and median bundle morphology[149]. Toy^{DN} expression also results in morphological phenotypes in the IPCs, that are, however, distinct from those observed in Ey mutants. Specifically, the presence of long ectopic pioneer neurites emanating from the median bundle are not observed in Ey mutants. These ectopic pioneer neurites are also seen upon IPC-specific knockdown of PlexinA, a neuroglian receptor involved in pathfinding suggesting that Toy may be acting upstream of PlexinA[108,149,261]. In addition to the ectopic pioneer neurites, we observed that upon expression of Toy^{DN}, Dilp3 protein was not detected in de median bundle even though expression of *dilp3* was unaltered. Dilps are transported along the median bundle to the release sites. This phenotype suggests that Toy^{DN} affects the secretory pathway of the IPCs. Similar phenotypes have previously been observed upon overexpression of a dominant negative form of Rab1 in the IPCs which blocks Dilp2 transport to the axons[103]. Finally, contrary to the situation in Ey mutants where the clustering of IPC somata is perturbed, it is unaltered in upon Toy^{DN} expression[149]. We conclude that Toy and Ey both regulate expression of a *dilp* and IPC differentiation. However, they appear to do so in a largely non-overlapping and complementary way. Further experiments determining the precise transcriptional targets of Ey and Toy in the IPCs could identify the extent to which their downstream targets are shared. These results are an example of how paralogous genes are still partially redundant but have obtained specialized functions.

The identified *dilp3* regulators were less clear given the differences between two RNAi lines on *dilp3* expression of *CG6683*, *FoxP* and *CHES-1-like* (Figure 27 B). These differences could be due to insufficient knockdown in one of the RNAi lines or due to off-target effects[262]. Sug knockdown did not affect *dilp2*, *-3* and *-5* expression, consistent with previous results which showed that overexpression of Sug reduces their expression but that knockdown had no effect[151]. Knockdown of *sd*, a TF involved in Hippo signaling, significantly increased *dilp3* levels while significantly reducing *dilp2* and *-5*. For *dilp5*, we found that *bsh* knockdown significantly and selectively reduced *dilp5* while *CHES-1-like*, with two separate RNAi-lines gave inconsistent results on *dilp5* expression (Figure 28 B). The knockdown of *achi*, which is thus far only known for a role in spermatogenesis in *Drosophila*, yielded reductions in *dilp5*

levels with two RNAi lines but this did not reach statistical significance for *achi*$^{RNAi\ \#2}$. The mammalian homologs of Achi are best characterized for their role in tissue growth[263]

Given that previous work established that several transcription factors are shared between IPCs and mammalian beta cells, and that a second PAX6 homolog is functional in the IPCs[149,150,229], we wondered if the mammalian homologs of the identified regulators have a role in the beta cell. Compellingly, we found that several mammalian homologs play a role in pancreatic islet development and beta cell function (see Table 2). Our observation that three transcription factors of the forkhead family (*CHES-1-like*, *Foxp* and *fkh*) are required in the IPCs is intriguing, particularly in light of the fact that a key downstream effector of the insulin pathway, namely *foxo*, is also a forkhead domain transcription factor, and functional in the IPCs where it is involved in the regulation of *dilp3* and *-5* expression in an insulin signaling feedback loop[99,120]. Given that the forkhead domain transcription factors bind similar DNA-binding motifs, it is possible that these three identified forkhead transcription factors might have overlapping targets with *foxo* which could influence *dilp* expression [264]. Lastly, Bsx was shown to be required for Neuropeptide Y and agouti-related peptide expression in the murine hypothalamus[265]. This is interesting in light of the hypothesis that the pars intercerebralis is evolutionarily related to the hypothalamus. Our observation that *bsh* is required in the IPCs provides support for this hypothesis (Table 2).

Drosophila	Human	Relevant function of vertebrate homolog
Hr78	*NR2C1/2*	Expressed in islets and repressed by *JAZF1*, a T2D risk gene[266,267]
E2f2	*E2F2*	E2F1 and E2F2 double knockout mice have reduced islet size[268]
Nf-YB	*NF-YB*	NF-Y regulates beta cell mass and glucose homeostasis[257]
toy	*PAX6*	Regulates insulin transcription and endocrine cell differentiation[231]
sd	*TEAD1-4*	Regulates embryonic pancreatic development[269]
CG6683	-	
Max	MAX	MYC/MAX regulate beta cell regeneration[270]
fkh	FOXA1/2	FOXA2 cooperates with PDX1 to direct beta cell maturation[271,272]
FoxP	FOXP1-4	FOXP1/2 and -4 redundantly control alpha cell proliferation[273]
CHES-1-like	FOXN2/3	
Achi	TGIF1/2	TGIF2 is involved in early pancreatic specification in *Xenopus*[274]
bsh	BSX	BSX is involved hypothalamic neuropeptide expression[265]

Table 2. Summary of the relevant functions of the vertebrate homologs of identified *dilp* regulators.

We observed that knockdown of a transcription factor associated with one *dilp* often had additional effects on the two others which is perhaps not unexpected due to the complex epistatic behavior of the dilps. For example, *Δdilp2* mutants upregulate both *dilp3* and *-5* while *Δdilp3* mutants have slightly reduced *dilp2* and *-5* levels[89]. In the case of *Hr78* knockdown, a significant reduction in *dilp2* was observed and additionally increased *dilp3* while decreasing *dilp5* (Figure 26 B and Figure S20 A -B). This upregulation of *dilp3* could be due to compensatory effects caused by *dilp2* loss but the interaction with *dilp5* is less clear. Hr78 could possibly be a direct regulator that was missed by the yeast-1-hybrid screen if it functions as a heterodimer there or could act as a regulator of *dilp5* at a distal enhancer region. Alternatively, given that we performed knockdown experiments throughout development there is a possibility that knockdown causes a developmental defect in IPC differentiation which in turn more broadly affects *dilp* expression.

We also evaluated the effect of several transcription factors that resulted in a reduction of *dilp2*, *-3* or *-5* mRNA on the Dilp protein level. We found that only in the case of *toy*, there was a detectable difference in Dilp protein accumulation in the IPCs. Differences between mRNA level and protein levels are common and could be due to a number of reasons. We speculate that given that the production rate of highly expressed genes is mostly determined by translation efficiency and not by transcript abundance that this is the cause of the mRNA-protein differences[275,276].

This study focused on the identification of IPC-specific regulators of *dilp2*, *-3* and *5* and it has been shown that the proximal promoter regions of *dilp2*, *-3 -5* are sufficient to drive strong expression in the IPCs. However, parts of these regulatory regions additionally drive expression in different tissues[98,100,123]. The region between -763 and -165 of the *dilp3* promoter for example, drives expression solely in the muscles of the midgut, while its flanking regions are necessary for IPCs[98]. Putative regulators of *dilp3* that are not expressed in the IPCs or had no effect in knockdown experiments could thus be interesting candidates for transcription factors driving muscle-specific *dilp3* expression.

In summary, we identified several putative direct regulators of *dilp* expression and provide evidence for the direct regulation of *dilp2*, *-3* and *5* expression in the IPCs of several transcription factors. Many of their homologs are involved in pancreatic development and function. Lastly, we show that Toy, a second *PAX6* homolog, directly regulates *dilp2* and is involved in IPC development, distinct from *ey*.

Methods

Yeast-Hybrid screen

Insulin regulatory sequences (Supplemental table 1) were cloned upstream of the *HIS3* reporter to generate baits further tested in yeast one-hybrid. *HIS3* encodes an imidazoleglycerol-phosphate dehydratase, that catalyzes histidine synthesis, and the inhibitor 3-amino-1,2,4-triazole (3AT) competitively inhibits this activity. Thus, the higher the level of 3AT in the medium is, the higher *HIS3* expression needs to be to ensure yeast growth, thus testing the strength of transactivation of the *bait-HIS3* in response to multiple Transcription factors. Prior to the transcription factor/bait interaction test, a self-activation test was performed to assess whether natural *S. cerevisiae* transcription factors are sufficient to induce basal *HIS3* expression. This test was performed by measuring the growth of eight independent transformed yeasts for each bait on SC-His plates with varying concentrations of 3AT (0, 10, 20, 40, 60, 80mM). For each bait, a transformant yeast that can grow on SC-His medium but is unable to grow on medium supplemented with 3AT was selected.

The yeast one-hybrid assay was performed as previously[252]. Briefly, a pre-existing clone collection of 700 transcription factors fused to the Gal4 activation domain (prey) was transformed into each bait strain in a 384-well plate format, using LiAc-PEG, and plated on SC –His –Trp medium to select for successfully transformed yeast colonies. After 3d, yeasts were plated in quartets on plates containing increasing concentrations of 3AT (0, 10, 20 and 40 mM). Plates were incubated at 30°C for 3, 7, and 10d and imaged using a Bio-Rad gel doc system. Yeast not transformed with any transcription factor prey and yeasts transformed with the Gal4 activation domain alone served as negative control. Sets of quadruplicate colonies that showed growth above background levels were deduced to have a direct interaction between the transcription factor prey and the DNA bait.

Each dilp regulatory element was integrated two independent times into the yeast genome, the large *dilp3* regulatory element (1352 bases) was additionally subdivided into two smaller sequences and integrated as well. Each element integration was screened two times for a total of 4 screens per element. Every interaction that was detected at least once in both integrations was scored as a positive hit.

Fly Husbandry

All experiments were performed on Nutrifly (Genesee Scientific) at 25°C with a 12h/12h day night cycle unless otherwise specified. IPC-specific manipulations were performed with *Dilp2-GAL4^{R};btlGAL80*. Control lines used were GD control (VDRC # 60000) and KK control

(60100). RNAi lines were ordered from the Vienna *Drosophila* research center (VDRC) and are listed in Supplemental table 2.

RNA extraction, cDNA synthesis and qPCR

Flies were snap frozen on dry ice and 10 to 15 fly heads of 3-to-6-day old flies were used per replicate for RNA extraction. Total RNA was isolated using TRI™ reagent (Invitrogen) and reverse transcription was performed on 1µg RNA using SensiFAST cDNA Synthesis kit (Bioline). Primer sequences used are listed below. qPCR was performed on a QuantStudio™ 12K Flex Real-Time using the SYBR Green detection system. Transcript levels were normalized to *Rp49* transcript levels. Mean ΔCt values were statistically compared using a one-way ANOVA and compared with control genotypes using post-hoc Dunnett's test. Relative quantitation of transcript levels to control genotypes were calculated using the ΔΔCt method and plotted with GraphPad Prism software to visualize expression differences.

Immunohistochemistry

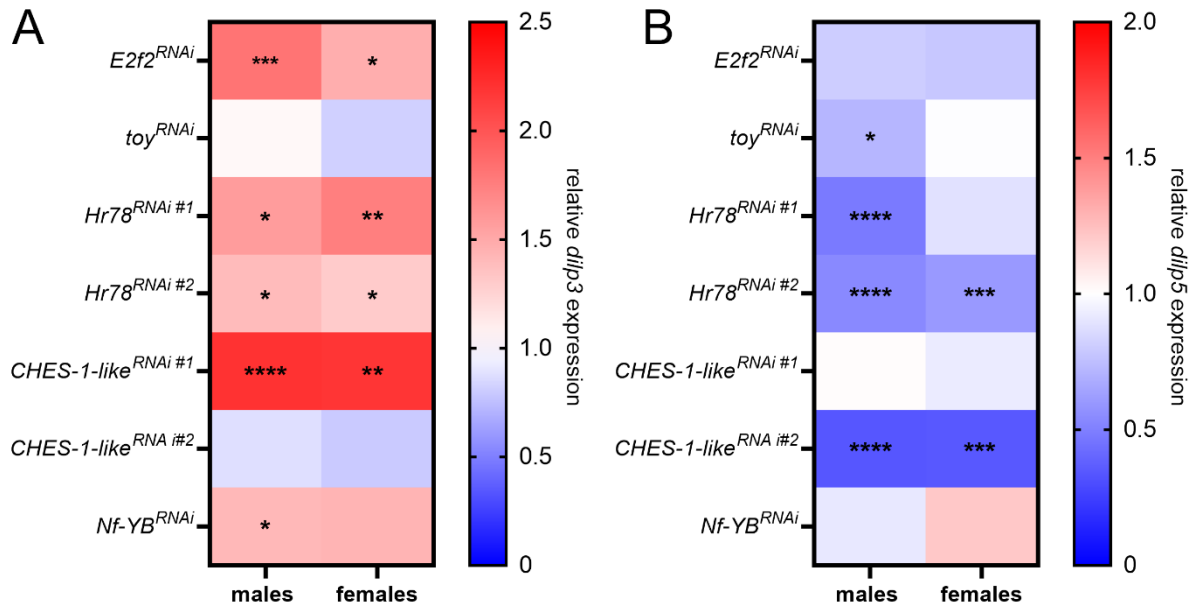
Larval and adult brains were dissected in cold 1X PBS, fixed for 20 minutes in 4% formaldehyde, and washed in PBS prior to 30'-60' blocking in PAXD (PBS containing 1% BSA, 0.3% Triton X-100, 0.3% deoxycholate). Brains were incubated overnight at 4°C in PAXD with primary antibodies, washed and incubated overnight at 4°C in PAXD with secondary antibodies. Samples are then washed and mounted in Vectashield Mounting Medium (Vector Laboratories Cat#H-1000). Images were taken with an Olympus FluoView FV1000 confocal microscope and processed using ImageJ64[206] and Photoshop software.

Antibodies used in this study include rat α-Dilp2 (1:100), rabbit α-Dilp3 (1:100), rabbit α-Dilp5 (1:400) and rat α-Toy (1:50), Secondary antibodies used were donkey α-rat Alexa-594 and donkey α-rabbit Alexa-594 (Jackson Immunoresearch). All secondary antibodies were used at 1:200 concentration. Rat antibodies directed against the Dilp2 partial peptide sequence TRQRQGIVERC, rabbit antibodies directed against the Dilp3 partial peptide sequence DEVLRYCAAKPRT and against the Dilp5 peptide sequence RRDFRGVVDSCCRKS were generated as a service by Thermo Fischer Scientific Inc and have been previously characterized[123].

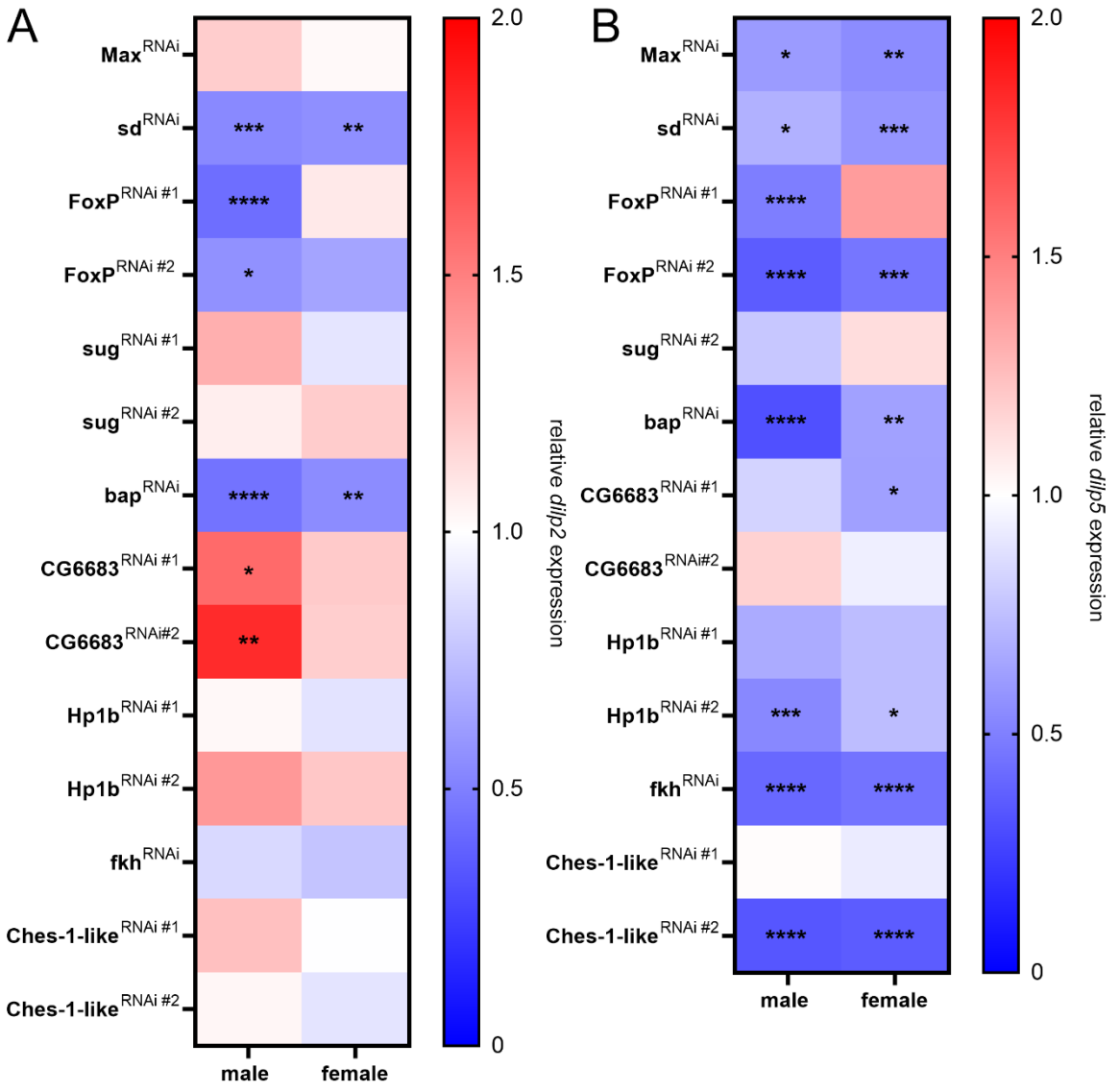
Adult fly weight

Adult weight was determined of chloroform anesthetized flies that were 3-6 days old and were raised at similar larval densities. Flies were weighed on a Mettler Toledo XS204 scale (Wet Weight; d = 0.1 mg). For Toy^{DN} weight measurements, balanced *CyO* siblings were used as an internal control.

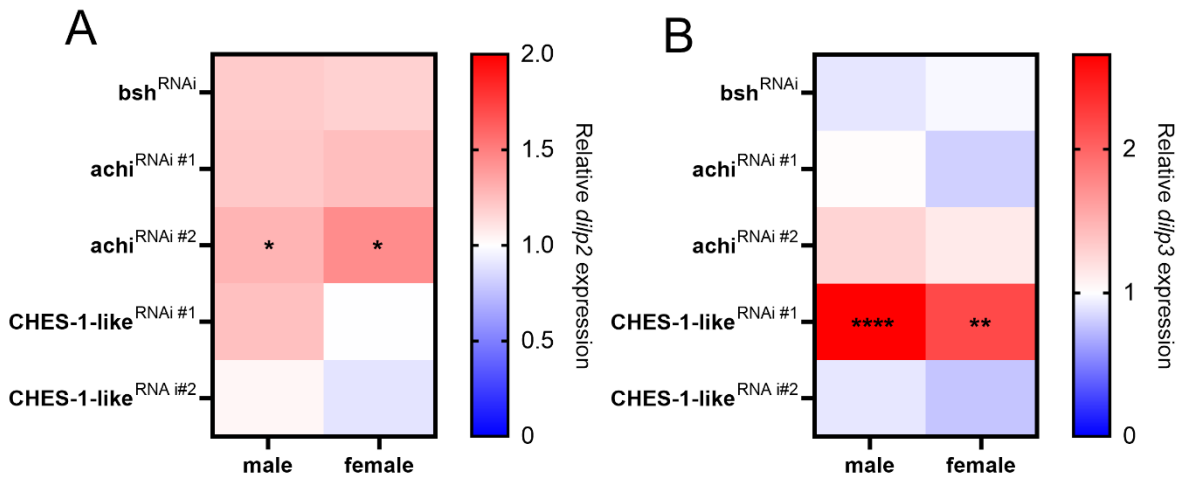
Supplementary information



Supplemental Figure 19. Knockdown of putative *dilp2* regulators results in upregulated *dilp3* and downregulated *dilp5*. Knockdown of *Hr78* consistently upregulated *dilp3* and downregulated *dilp5* (A - B). *E2f2* Knockdown upregulated *dilp3* in both sexes (A) (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, One way ANOVA with post-hoc Dunnett's tests).



Supplemental Figure 20. Additional effects on *dilp2* and *-5* by putative *dilp3* regulators. Knockdown of putative *dilp3* regulators often has pleiotropic effects on *dilp2* (A) and *dilp5* expression but generally results in downregulation of *dilp2* and *-5*. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, One way ANOVA with post-hoc Dunnett's tests).



Supplemental Figure 21. Slight upregulation of *dilp2 and -3* expression upon knockdown of putative *dilp5 regulators.* Knockdown of putative *dilp5* regulators tends to result in upregulated *dilp2* and *-3.* (* $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$, One way ANOVA with post-hoc Dunnett's tests).

Regulatory sequence	Start (Chr 3L)	End (Chr 3L)	Length (Nucleotides)
Dilp2	9791999	9792866	868
Dilp3 (A)	9795147	9796498	1352
Dilp3 (B)	9795877	9796498	622
Dilp3 (C)	9795147	9796008	862
Dilp5	9816968	9817681	714

Supplemental table 1. Summary information pertaining to the regulatory sequences used in the yeast-1-hybrid screen. Nucleotide positions relate to those of *dm3* genome assembly. Nucleotide sequences are additionally provided below

Name	VDRC stock number
E2f2^{RNAi}	45743
toy^{RNAi}	15919
Hr78$^{RNAi\ \#1}$	109435
Hr78$^{RNAi\ \#2}$	48978
CHES-1-like$^{RNAi\ \#1}$	105641
CHES-1-like$^{RNA\ i\#2}$	15742
Nf-YB^{RNAi}	103655
Max^{RNAi}	15351
sd^{RNAi}	101497
FoxP$^{RNAi\ \#1}$	50200
FoxP$^{RNAi\ \#2}$	15732
sug$^{RNAi\ \#1}$	21719
sug$^{RNAi\ \#2}$	109584
bap^{RNAi}	101354
CG6683$^{RNAi\ \#1}$	35109
CG6683$^{RNAi\ \#2}$	102695
Hp1b$^{RNAi\ \#1}$	26097
Hp1b$^{RNAi\ \#2}$	109735
fkh^{RNAi}	49961
achi$^{RNAi\ \#1}$	49635
achi$^{RNA\ \#2}$	49637
bsh^{RNAi}	15716

Supplemental table 2. RNAi-lines used in this study.

Full regulatory sequence of *dilp2*

GACACTTGGCCAACACACACACATTCATTACCCAGCATGCATACACATACATACATACATACGAACACTTGAA
CATGGCATTCCCACATACATACTATATGTACACCCGTTCTCCTAACAGCATTCGTAATCCAATTCGTAATCAC
TTATTTATGCAGCACCAACTTTGACTAATGAATAAAAGTAGACAGTGCAAAAAGAAGAATTTGTGTGAGCGGC
AAAGACAAGACAGTCAAAGGAAAAGTACCAAGGAGGCGTAATTCGTTGCCAGGATGCTGCTGGCTCCAATC
AGCTGGCTCGGGTTCTTGTTGGCCTACTCCGGCCTGCCGGGCGCCCAGGCGCGGATTGGTAAGTATAGGG
TAGAAATGGTAGAGGCAAACGGGCTGTAAAGGCGAGGTAACTTTAAGACCAAGGACTTATTTGGTATTCATA
AATGCAGGGCTTAGTACGGCCTGGCTGTTTATACATAAATTGTTTAATTGGCTGACAGCAAAACGCCCACAA
AGCCCTTCTTCTCCCAAACACAGCCACCCAAAAACCACCCACAATTTTCCAATAGCCATTTGCATAAGTCAGC
TTTGACATTGAATGAGTTTAGCCAGGCGTTAGAGAGAAGAGCCAACTCCGGAATGCACTCATCGCATCCTAA
CACAATTTGACAACCTGACTTTTGAGGATAACATGTGTGTGGATCCAGACCCCGACCCCGAGAGACCGACC
CCGACCCAGACCCAGTTTAAGTTTCTTACGAGGCAGTATATAAGCCAGTTACTTTGGGCCAAACATCATCAG

TTAGCTTTCGACCCGGCTCGACCCAACTTAATCCATTTGATCGTAAAGCAACCTAAGCAGTAAACCCATAACC
ATG

Full Regulatory sequence of *dilp3* (A is underlined and B is highlighted in grey)

AAACAAAAGAAGAGTCCATTGCTTTTTTCTACATTCTACGCCCTGCAGCATTCCAGCTGTGAGGCATGGGGA
ATCCCCTTGTTATTCAAACCACCCGAAGCCACCCAAACCATCGAGCCACCCACAAGCAGCTGCCATTCAGCA
CCTCGAGTGCGGTGCCCTTGTTTTCCGAGAACAATAATGAAAAATATGAATTTTTAATTAGATGACGTTCTGA
TTTTAATAAGCAAAACAAAAGGTGGAGACAAAACGAACTCGGTAATACACTCGGATTCGAATTTACAGCTTCC
TTTTTATCCATAATTTTTGTTATTATCGAAGGAGCGATATCAAAACTAGAAAACAACTTCCAATCAGTAGCGGG
ATTTTCCGAAGATAACACTCTATTCAACCGAAGGGTTTTGAAATGATAATAATTCCGTTCTTACAGGTAAAAAT
CTATACTAATACCTGTTTTTTTGCGGACGGAAAAAAGGCTCAGTTGGCTTATCATTGGCAAAAGGGACTTGG
GGAAACCATAAAGTATCGAAGGTACTGAGTCAAGATAATGAGATAACAGAAGGCGACTTTATTGTTTTCCACT
CAAAAGCAATTGAATAAGTTGGCACTCGTTTTTAATTGAATGGGAATGAAATAAGCTCTAAAAGTGTTGTTAAA
ACGTAATGGCTTTTGTGTTAATTTAAAGAATTTAAGTAGTTTTGAAAGTATCATTATTCTTTAGGTAATTTTTAT
TACATTCCAAATTTAATAAATGACTAATTCGAAAAAGTGTTTATTTAATCAATGAATATATTTCAAGTAAGTTTA
CTTTTAGTAGCTTGCCAAATGTGAGTTTAAATATGTATGCATAGAACTATATAGTTAAACTGCTAAACTTTACA
GTTAAACTTTCTGAACCCACCAAAATGGATGAACATCCTCGTCTGCCGAAGGGAACTCGATGCACGTCATTT
TGTTTTTCAACAATCCAGATCCGTGCGCTACTTCTTGGGCGAGAAAGTAAACAAACGCCAGCTGATATGCGT
CAGACCCCCCGGGCTCATCATCATCTCACCATTTCAGACATCCCATGCCAGCCCGAATCCTCACGAGAAACT
AGACCAGACCAGGGCGAACTACATATGTGGATGATGCTAACTGACACTACGGCTGACTCATGCTGACAGTG
CTCAGACGCTGGATACAGCCCGCAGACATCCAACTCGTATCCTATCCGATTCTGCCCCATATATATAACCCT
CAGTCGATGGCTGGGAGGCAAACAGTTGAGGCCGTGCCACTTGGCAGACACATACTACACACTCCCCGGG
GGATTCACGCATCCATACTTAAACACCACTTCATCACTCATG

Full Regulatory sequence of *dilp5*

TTCAGGGGCTTCGTTATTTGATCGTCGACTTCAATCCTAAAAATAACGTAAAAATACAATACAATTAGTAATGT
ATAATGCGTTAGGTAAAATTTCAAATTTCGCGGCATGTACTGCTTGTGTGTTTGATGCACTGCTTGGACTGCT
TGCAAGCATCTGGCGCCAAATTTAAATTTATTTTTAAACATGCTTTCCTCGCAGGCATGTATTCGATTGTCATG
CATTCTTTTGGTATAAGTTCTAGAATTATACATGTGAACAGGCAAAAATGTAATATACGACTTTACTTTCTTAG
ATATACAATAAGTACATTACTTAAATATTTCATGAAATACAAACTTATGATCGTATGCACTATTTGCACGATAAG
CCAATTAAACATATTGAAAATGAAAGCCCTTCTAACTAAAGACCATATATATCGTTATACATTTTGATTTCAGC
TATTATATATAGATTCCGCCACCCCCTTGGTGCATCCGATTAAGCGAAAGTTACGCTAATCCATGGGCATAAA
TTAATGATAGCACCCTACGATGTCCTGACATTGAAAATTAGCAAGTACGCTCCTGCCCACGCTGACCCCAAA
AGACTAATCCAATCTGACAGGTGAAAGTGCGCCGGCGGATATAAAAATGGCTGACAGCGACGGAAACTGCA
CCACATCCTATCCAGCAGCAGTTCCAGCAAGGCAATGATG

6 General discussion

Significant efforts over the years have contributed to a deeper understanding of the control of IIS in *Drosophila*. This has identified an impressive number of non-autonomous signals, many of which are nutritionally regulated. These signals are integrated by the IPCs to produce and secrete appropriate Dilp levels to achieve metabolic homeostasis[109,178]. This tight control of IIS by nutritional cues is essential in the larval stage to ensure proper final adult size[233]. In the adult stage, IIS levels are additionally involved in controlling fecundity and longevity[92]. In contrast, the transcription factors that instruct IPC cell fate, function and control *dilp2*, *-3* and *-5* expression remain less well understood. Here, I will discuss the results presented in this thesis which contributes to a deeper understanding of the transcriptional control of IPC *dilp* expression, IPC differentiation and what this can tell us about redundancy between dilps and the evolutionary conservation of "insulin-producing cells".

Connecting autonomous to non-autonomous signals.

Metabolic homeostasis fundamentally depends on inter-organ communication which allows cells to relay information to other cell types which in turn can respond appropriately. This is dependent on the secretion of proteins into the circulatory system where they can act on receptors on target tissues or alternatively via direct innervation and neurotransmitter release from neurons[277]. In the context of *Drosophila* IIS, a key source of non-autonomous signals is the fat body, which acts as the primary nutrient-sensing organ in the fly[135]. These signals typically act to increase *dilp* expression in nutrient-rich conditions while others are secreted in conditions of dietary restriction to decrease *dilp* levels. Interestingly, their effects on *dilp2*, *-3* and *-5* are usually not uniform. The fat body secretes Eiger in response to dietary restriction which inhibits *dilp2* and *-5* expression but leaves *dilp3* unaltered[142]. Other signals such as neuronally derived octopamine specifically inhibit *dilp3* expression[147]. Clearly, these results indicate that the three IPC *dilps* are independently transcriptionally regulated. Three different approaches were used in this thesis to identify novel IPC-autonomous regulators: (i) using variation in *dilp2*, *-3* and *-5* expression across the DGRP (Chapter 3), (ii) characterizing the homolog of a diabetes GWAS candidate (Chapter 4) and (iii) performing a Yeast-1-hybrid screen for the *dilp2*, *-3* and *-5* proximal promotors (Chapter 5).

The results from chapter 3 revealed that most members of the retinal determination gene network (RDGN), for example, specifically controlled *dilp2* and *-5* expression while Hth/Exd were essential only for *dilp2* and *-3* (Figure 13 and 14). Additionally, the results from the Yeast-1-Hybrid screen showed that the direct regulation of *dilp2*, *-3* and *-5* had very little overlap as identified transcription factors were usually regulating one *dilp* only (Table 1). The co-

regulation of *dilp2* and *-5*, which is often noted in the literature therefore does not seem to depend on shared direct transcriptional regulators.

Several transcription factors for which we identified a role in *dilp* regulation in the IPCs are already expressed early during IPC development[97,229]. This is reminiscent of what is seen in the beta cells where PAX6 and PDX1 play a role in beta cell differentiation and pancreatic development but are additionally directly involved in *INS* expression[22]. A similar feedforward mechanism has previously been described for peptidergic neurons in the *Drosophila* ventral nerve cord[278–280]. Early transcription factors in these cells specify cellular identity by promoting the expression of other transcription factors. This combinatorial code activates the transcriptional program for the specific neuronal subtype and drives expression of the neuropeptide gene[281].

However, effects on *dilp* expression observed in experiments performed throughout development, could also reflect a developmental issue with IPC differentiation. This is in and of itself an interesting result but does not necessarily reflect a direct role in the regulation of *dilp* expression. Indeed, when we specifically expressed Pan^{DN} in the adult IPCs this did not have the dramatic effects on *dilp* expression seen upon its expression throughout development (Figure 22). This shows that while Pan plays a key developmental role in the IPCs, its role in *dilp* regulation post-differentiation is only minor. In light of this, we propose that the RDGN members with an important impact on *dilp2* and *-5* expression are not directly involved in *dilp2* and *-5* regulation but are rather involved in IPC development and/or act upstream of direct regulators. Determining whether the RDGN members regulate transcription of the direct *dilp*-regulators is a challenge for the future. In addition, the direct regulators which we identified in chapter 5 using the yeast-1-hybrid screen and validated using an RNAi-based approach, should be further validated using dominant-negative versions of the transcription factors. It is our experience that RNAi of transcription factors typically yields weaker or no phenotypes than the use of dominant-negative versions of the transcription factors which can lead to false negative results. Moreover, using methods to limit expression to a specific stage will allow us to determine if some of these transcription factors are stage/context specific.

Further research should try to establish links between non-autonomous signals and individual or groups of transcription factors. One already known example of such a connection is the nutritional protein-dependent increase in *dilp5* expression during the larval stage which depends on a functional interaction between Ey and Dac[99]. Hth and Exd, which are involved in *dilp2* and *-3* expression, are however, seemingly not nutritionally regulated and are, based on the yeast-1-hybrid data, not involved in the direct transcriptional control of *dilp2* and *-3* (Table 1). However, due to the heterodimeric association of Hth and Exd, the absence of

detection in the yeast-1-hybrid could reflect their interdependent nature regarding DNA binding. Regardless, they could function to allow for baseline *dilp2* and *-3* expression or to interact with other transcription factors which in turn promote *dilp2* and *-3* expression. We can speculate about the possibility of a link between Tachykinin and Hth/Exd, for example, as knockdown of its receptor specifically increases *dilp2* and *-3*[282]. This specific effect on *dilp* expression could be facilitated by repression of Hth/Exd either via nuclear exclusion or indirectly via reduced expression of *hth/exd.* Such interactions are purely speculative but are an example of how future research could try to connect the autonomous with the non-autonomous.

Gene redundancy and diversification: the case of IPC *dilp* expression

Why do the IPCs express *dilp2*, -3 *and* -5 (and *dilp1* in certain conditions)? It has been postulated that IIS performs such a variety of functions that one explanation for the presence of multiple IIS ligands in most multicellular organisms is that this allows for functional diversification of the ligands. The last common ancestor of *Drosophila* and humans likely contained a single insulin gene that has separately expanded due to lineage specific gene duplications[86,87]. In *Drosophila*, the expansion of the *dilp* family is evolutionarily old, more than 40 million years, and is thus stable[89]. Fully redundant genes, i.e. genes that perform the same functions in the same conditions equally well, are because of a lack of selective pressure not stable[283]. The overlapping but spatially and temporally distinct expression profiles could provide one explanation for the evolutionary stability of *dilp2*, *-3* and *-5*. Additionally, IPC *dilp* expression is likely not fully redundant as seen in the phenotypic consequences of loss of *dilp2*, which increases circulating trehalose, which is something not seen in *Δdilp3* or *Δdilp5* flies[89]. Although, given that *dilp2* is the most abundant IPC *dilp*, it is possible that its removal cannot be fully compensated by *dilp3* and *-5*. One intriguing related result is the finding that Toy directly regulates *dilp2* while previous results show that Ey directly regulates *dilp5* (Figure 17, 26 and 29)[149]. There is only a single *ey/toy* homolog in mammals, PAX6, which is also involved in the direct transcriptional control of *INS*[231]. In general, there are multiple homologs of a single *Drosophila* gene due to two genome duplications in the early vertebrate lineage[284]. Regarding the relationship between *toy/ey* and *dilp2* and *-5*, this seems to have been reversed. We hypothesize that the duplication of *toy* and *ey* early in insect evolution and the later emergence of *dilp2* and *-5* in the *Drosophila* branch led to a specialization of *toy* in the regulation of *dilp2* and of Ey with *dilp5*[89,253]. This also shows that a connection between PAX6 and insulin is evolutionarily conserved.

On the origin of the beta cell

The first occurrence of an islet-like organ occurred in vertebrates and can be detected in lampreys and hagfish where it consists mostly of cells that produce insulin and a few SS^{+} cells[285]. It is only in species of the gnathostome clade that real islets containing cells that also produce glucagon, PP and ghrelin emerge. Developmentally, the pancreas derives from the endodermal digestive gut tube. Nonetheless, it is possible that the ancestral origin of insulin expression is neuronal[23]. A few lines of evidence support this notion. First, the beta cell is neuron-like. Insulin is packaged into neurotransmitter-like vesicles and the general gene expression profile of beta cells is neuron-like[286,287]. Furthermore, the REST complex, which represses a neuronal phenotype is not expressed in the beta cell[158]. NEUROD and NGN3 are some of the downstream targets that are repressed by REST and are transcription factors that the beta cells share with hypothalamic neuroendocrine cells[288,289]. Furthermore, there are several lines of evidence showing that insulin is expressed in some neurons, specifically in the hypothalamus[290]. Moreover, in many insects other than *Drosophila*, like aphids and the silk moth, insulin expression is detected in cells located in the pars intercerebralis of the brain[291,292]. Lastly, in species of the genus *Hydra*, belonging to the Cnidaria, the sister group of Bilateria (that includes vertebrates and invertebrates), there is expression of insulin-like peptides in neurons[293]. The emergence of the beta cell in islets could thus be a derived feature in vertebrates with the ancestral origin being neuronal. The beta cell could therefore have co-opted the already existing regulatory networks active in this ancestral insulin-producing neuron into the endoderm.

In this context, we have found that several transcription factors that are active in the beta cells are also active in the neurosecretory IPCs. First, we have shown that Hth/Exd are implicated in the (indirect) regulation of *dilp2* and *-3*, while their homologs MEIS/PBX are involved in promoting PAX6 expression[188]. In *Drosophila*, this regulatory relationship does not exist with Ey (figure S3) but might exist for its paralog Toy. Additionally, we found that the *TCF7L2* homolog *pangolin* is essential for IPC differentiation and that it acts upstream of several key transcription factors like Ey and Dimm (Chapter 4). This seems to reflect an evolutionary conserved function, given the strong effects of *TCF7L2* deletion on the expression of several key beta cell transcription factors like PDX1 and NEUROD1[241]. Furthermore, a significant number of the transcription factors identified with the Yeast-1-hybrid screen are involved in pancreas/beta cell development (chapter 5). All these results combined provide support for a common regulatory program associated with insulin regulation among Nephrozoa. In other model organisms where there is genetic evidence available, like zebrafish, Pax6b was found to be involved in beta cell development as well[294]. Interestingly, the PDX1 transcription factor which is key in pancreatic development and the transcriptional control of *INS* together with

NEUROD1 has been identified in other invertebrates like mollusks where it also regulates expression of an insulin[26,295,296]. However, PDX1 has been lost in all Ecdysozoa, showing that there are divergences between *Drosophila* IPCs and mammalian beta cells[297]. Moreover, the exact regulatory relationships have diverged as illustrated by e.g., the absence of an interaction between Hth and Ey. However, an evolutionary ancient transcriptional program that controls insulin expression dating back to the split between vertebrates and invertebrates seems to exist.

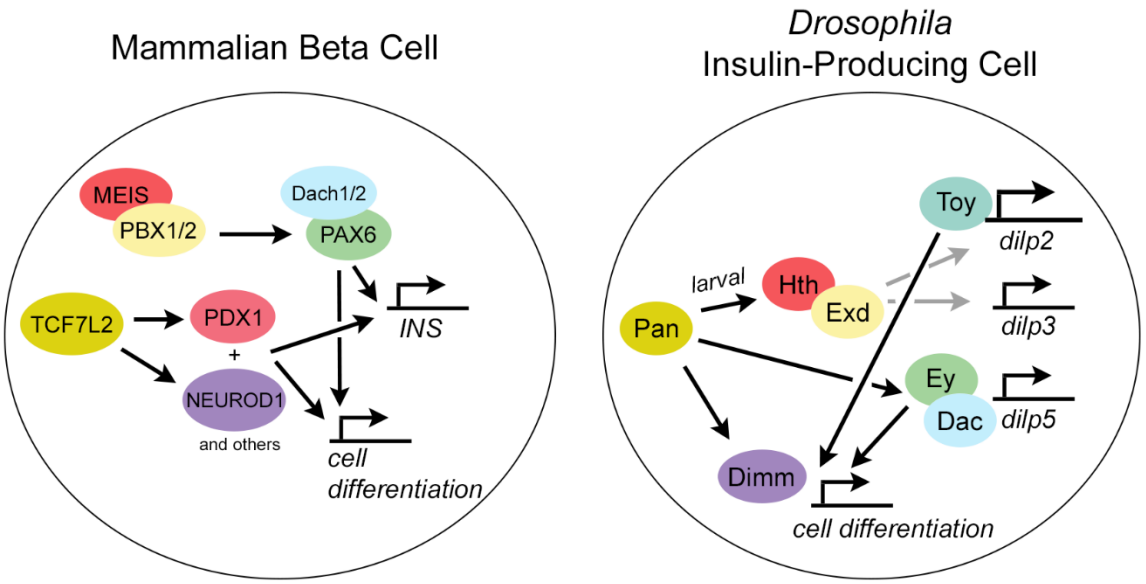


Figure 30. Model comparing the transcriptional networks at play in the mammalian beta cell with *Drosophila* IPCs. The transcription factors that orchestrate beta cell development and *INS* regulation show significant overlap with those active in the IPCs while their specific regulatory interactions might have diverged. Several transcription factors are omitted for the sake of simplicity.

Conclusions

To conclude, we have identified several novel autonomous regulators of *dilp* expression in the IPCs, many of which are conserved. This work can form the basis to further elucidate the interaction between non-autonomous and autonomous regulation of the IPCs. Furthermore, the identified transcription factors provide further evidence for an ancient regulatory transcriptional network controlling insulin production.

Future perspectives

We identified several transcription factors that are involved in IPC development and *dilp* regulation. Further research can elaborate on these results in several ways.

First, since we have mostly performed tissue-specific knockdowns active throughout development, we have missed the temporal aspects of *dilp* regulation and IPC development. Charting the expression of the identified transcription factors throughout IPC development and through various nutritional challenges will give us a less static picture of the IPCs. Additionally, temporal restrictions via a temperature sensitive GAL80 or using inducible GAL4 systems will allow us to distinguish the role of a transcription factor in *dilp* expression from its role on IPC development.

Secondly, in this work we have mostly considered the IPCs as a homogenous cell population but work from other labs, mostly based on immunohistochemical experiments, has shown that there are distinct subpopulations. Clearly, there will be many transcription factors that are important in the entire IPC population. Yet, there must be specific transcriptional networks that instruct these subpopulations, but they remain unidentified. Here, we think that deep and thorough single cell sequencing of the IPCs will be able to chart these subpopulations and allow us to take targeted approaches to understand their specific functions. Furthermore, given the large effects of the depletion of even a single transcription factor on the IPCs we believe that RNA-sequencing of the manipulated IPCs, single cell or bulk, can provide a more holistic view of the role of the transcription factor in the IPCs. For example, to what extent are the transcriptional programs of Ey and Toy in the IPCs shared or what is the exact dysregulation caused by Pan^{DN} in the IPCs? These questions will be more readily answerable using higher throughput approaches.

Lastly, we have provided some evidence for genetic interactions between transcription factors themselves in the IPCs. Building the regulatory network using bio-informatic approaches followed by *in vivo* validation will be informative from an evolutionary perspective.

References

1. Hiller-Sturmhöfel, S. & Bartke, A. The endocrine system: an overview. *Alcohol Health Res. World* **22**, 153–164 (1998).
2. Rahman, M. S. *et al.* Role of Insulin in Health and Disease: An Update. *Int. J. Mol. Sci.* **22**, 6403 (2021).
3. Brogiolo, W. *et al.* An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Curr. Biol.* **11**, 213–221 (2001).
4. Efstratiadis, A. Genetics of mouse growth. *Int. J. Dev. Biol.* **42**, 955–976 (1998).
5. Kimura, K. D., Tissenbaum, H. A., Liu, Y. & Ruvkun, G. daf-2, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science* **277**, 942–946 (1997).
6. Walker, J. T., Saunders, D. C., Brissova, M. & Powers, A. C. The Human Islet: Mini-Organ With Mega-Impact. *Endocr. Rev.* **42**, 605–657 (2021).
7. Bosco, D. *et al.* Unique arrangement of alpha- and beta-cells in human islets of Langerhans. *Diabetes* **59**, 1202–1210 (2010).
8. Cabrera, O. *et al.* The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proc. Natl. Acad. Sci.* **103**, 2334–2339 (2006).
9. Nurk, S. *et al.* The complete sequence of a human genome. *Science* **376**, 44–53 (2022).
10. Ezkurdia, I. *et al.* Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum. Mol. Genet.* **23**, 5866–5878 (2014).
11. CRICK, F. Central Dogma of Molecular Biology. *Nature* **227**, 561–563 (1970).
12. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
13. Lenhard, B., Sandelin, A. & Carninci, P. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat. Rev. Genet.* **13**, 233–245 (2012).
14. Levine, M. Transcriptional enhancers in animal development and evolution. *Curr. Biol.* **20**, R754-63 (2010).
15. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
16. Andersson, R. & Sandelin, A. Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet.* **21**, 71–87 (2020).
17. Noguchi, H. *et al.* Mutations in the C1 element of the insulin promoter lead to diabetic phenotypes in homozygous mice. *Commun. Biol.* **3**, 309 (2020).
18. Fu, Z., Gilbert, E. R. & Liu, D. Regulation of insulin synthesis and secretion and pancreatic Beta-cell dysfunction in diabetes. *Curr. Diabetes Rev.* **9**, 25–53 (2013).
19. Walker, M. D., Edlund, T., Boulet, A. M. & Rutter, W. J. Cell-specific expression controlled by the 5'-flanking region of insulin and chymotrypsin genes. *Nature* **306**, 557–561 (1983).
20. Qiu, Y., Guo, M., Huang, S. & Stein, R. Insulin Gene Transcription Is Mediated by Interactions between the p300 Coactivator and PDX-1, BETA2, and E47. *Mol. Cell. Biol.* **22**, 412–420 (2002).

21. Melloul, D., Marshak, S. & Cerasi, E. Regulation of insulin gene transcription. *Diabetologia* **45**, 309–326 (2002).

22. Hay, C. W. & Docherty, K. Comparative Analysis of Insulin Gene Promoters : Implications for Diabetes Research. *Diabetes* **55**, 3201–3213 (2006).

23. Jennings, R. E. *et al.* Development of the human pancreas from foregut to endocrine commitment. *Diabetes* **62**, 3514–3522 (2013).

24. Kim, S. K., Hebrok, M. & Melton, D. A. Notochord to endoderm signaling is required for pancreas development. *Development* **124**, 4243–4252 (1997).

25. Bort, R., Martinez-Barbera, J. P., Beddington, R. S. P. & Zaret, K. S. Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. *Development* **131**, 797–806 (2004).

26. Stoffers, D. A., Zinkin, N. T., Stanojevic, V., Clarke, W. L. & Habener, J. F. Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. *Nat. Genet.* **15**, 106–110 (1997).

27. Pan, F. C. & Brissova, M. Pancreas development in humans. *Curr. Opin. Endocrinol. Diabetes. Obes.* **21**, 77–82 (2014).

28. Jennings, R. E., Berry, A. A., Strutt, J. P., Gerrard, D. T. & Hanley, N. A. Human pancreas development. *Development* **142**, 3126–3137 (2015).

29. Lyttle, B. M. *et al.* Transcription factor expression in the developing human fetal endocrine pancreas. *Diabetologia* **51**, 1169–1180 (2008).

30. Byrnes, L. E. *et al.* Lineage dynamics of murine pancreatic development at single-cell resolution. *Nat. Commun.* **9**, 3922 (2018).

31. Salinno, C. *et al.* β-Cell Maturation and Identity in Health and Disease. *Int. J. Mol. Sci.* **20**, 5417 (2019).

32. Sarkar, S. A. *et al.* Global gene expression profiling and histochemical analysis of the developing human fetal pancreas. *Diabetologia* **51**, 285–297 (2008).

33. Otonkoski, T., Andersson, S., Knip, M. & Simell, O. Maturation of insulin response to glucose during human fetal and neonatal development. Studies with perifusion of pancreatic isletlike cell clusters. *Diabetes* **37**, 286–291 (1988).

34. Barsby, T. & Otonkoski, T. Maturation of beta cells: lessons from in vivo and in vitro models. *Diabetologia* **65**, 917–930 (2022).

35. Hang, Y. *et al.* The MafA transcription factor becomes essential to islet β-cells soon after birth. *Diabetes* **63**, 1994–2005 (2014).

36. Zhong, F. & Jiang, Y. Endogenous Pancreatic β Cell Regeneration: A Potential Strategy for the Recovery of β Cell Deficiency in Diabetes. *Front. Endocrinol.* **10**, 101 (2019).

37. Vasiljević, J., Torkko, J. M., Knoch, K.-P. & Solimena, M. The making of insulin in health and disease. *Diabetologia* **63**, 1981–1989 (2020).

38. Ramzy, A., Asadi, A. & Kieffer, T. J. Revisiting Proinsulin Processing: Evidence That Human β-Cells Process Proinsulin With Prohormone Convertase (PC) 1/3 but Not PC2. *Diabetes* **69**, 1451–1462 (2020).

39. Li, Y. V. Zinc and insulin in pancreatic beta-cells. *Endocrine* **45**, 178–189 (2014).

40. McCulloch, L. J. *et al.* GLUT2 (SLC2A2) is not the principal glucose transporter in human pancreatic beta cells: implications for understanding genetic association signals

at this locus. *Mol. Genet. Metab.* **104**, 648–653 (2011).

41. Campbell, J. E. & Newgard, C. B. Mechanisms controlling pancreatic islet cell function in insulin secretion. *Nat. Rev. Mol. Cell Biol.* **22**, 142–158 (2021).

42. Villalpando, S., Cazevieille, C., Fernandez, A., Lamb, N. J. & Hani, E.-H. Type II PKAs are anchored to mature insulin secretory granules in INS-1 β-cells and required for cAMP-dependent potentiation of exocytosis. *Biochimie* **125**, 32–41 (2016).

43. Kalwat, M. A. & Cobb, M. H. Mechanisms of the amplifying pathway of insulin secretion in the β cell. *Pharmacol. Ther.* **179**, 17–30 (2017).

44. Lu, C., Lam, H. N. & Menon, R. K. New Members of the Insulin Family: Regulators of Metabolism, Growth and Now … Reproduction. *Pediatr. Res.* **57**, 70–73 (2005).

45. Hede, M. S. *et al.* E-peptides control bioavailability of IGF-1. *PLoS One* **7**, e51152 (2012).

46. Roith, D. Le. The Insulin-Like Growth Factor System. *Exp. Diabesity Res.* **4**, 459056 (2003).

47. Livingstone, C. & Borai, A. Insulin-like growth factor-II: its role in metabolic and endocrine disease. *Clin. Endocrinol. (Oxf).* **80**, 773–781 (2014).

48. Baxter, R. C., Martin, J. L. & Beniac, V. A. High molecular weight insulin-like growth factor binding protein complex. Purification and properties of the acid-labile subunit from human serum. *J. Biol. Chem.* **264**, 11843–11848 (1989).

49. Firth, S. M. & Baxter, R. C. Cellular Actions of the Insulin-Like Growth Factor Binding Proteins. *Endocr. Rev.* **23**, 824–854 (2002).

50. Kawamura, K. *et al.* Paracrine regulation of mammalian oocyte maturation and male germ cell survival. *Proc. Natl. Acad. Sci.* **101**, 7323–7328 (2004).

51. Morgan, D. O. *et al.* Insulin-like growth factor II receptor as a multifunctional binding protein. *Nature* **329**, 301–307 (1987).

52. Seino, S. & Bell, G. I. Alternative splicing of human insulin receptor messenger RNA. *Biochem. Biophys. Res. Commun.* **159**, 312–316 (1989).

53. Belfiore, A., Frasca, F., Pandini, G., Sciacca, L. & Vigneri, R. Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocr. Rev.* **30**, 586–623 (2009).

54. Coppola, I., Brouwers, B., Meulemans, S., Ramos-Molina, B. & Creemers, J. W. M. Differential Effects of Furin Deficiency on Insulin Receptor Processing and Glucose Control in Liver and Pancreatic β Cells of Mice. *Int. J. Mol. Sci.* **22**, 6344 (2021).

55. Hernández-Sánchez, C., Mansilla, A., de Pablo, F. & Zardoya, R. Evolution of the Insulin Receptor Family and Receptor Isoform Expression in Vertebrates. *Mol. Biol. Evol.* **25**, 1043–1053 (2008).

56. Belfiore, A. *et al.* Insulin Receptor Isoforms in Physiology and Disease: An Updated View. *Endocr. Rev.* **38**, 379–431 (2017).

57. Nagao, H. *et al.* Distinct signaling by insulin and IGF-1 receptors and their extra- and intracellular domains. *Proc. Natl. Acad. Sci.* **118**, e2019474118 (2021).

58. Boucher, J., Kleinridders, A. & Kahn, C. R. Insulin receptor signaling in normal and insulin-resistant states. *Cold Spring Harb. Perspect. Biol.* **6**, a009191 (2014).

59. Khalid, M., Alkaabi, J., Khan, M. A. B. & Adem, A. Insulin Signal Transduction

Perturbations in Insulin Resistance. *Int. J. Mol. Sci.* **22**, 8590 (2021).

60. Haeusler, R. A., McGraw, T. E. & Accili, D. Biochemical and cellular properties of insulin receptor signalling. *Nat. Rev. Mol. Cell Biol.* **19**, 31–44 (2018).

61. Taniguchi, C. M., Emanuelli, B. & Kahn, C. R. Critical nodes in signalling pathways: insights into insulin action. *Nat. Rev. Mol. Cell Biol.* **7**, 85–96 (2006).

62. Katsarou, A. *et al.* Type 1 diabetes mellitus. *Nat. Rev. Dis. Prim.* **3**, 17016 (2017).

63. Pociot, F. & Lernmark, Å. Genetic risk factors for type 1 diabetes. *Lancet* **387**, 2331–2339 (2016).

64. DeFronzo, R. A. *et al.* Type 2 diabetes mellitus. *Nat. Rev. Dis. Prim.* **1**, 15019 (2015).

65. González-Muniesa, P. *et al.* Obesity. *Nat. Rev. Dis. Prim.* **3**, 17034 (2017).

66. Rohm, T. V, Meier, D. T., Olefsky, J. M. & Donath, M. Y. Inflammation in obesity, diabetes, and related disorders. *Immunity* **55**, 31–55 (2022).

67. Laakso, M. & Fernandes Silva, L. Genetics of Type 2 Diabetes: Past, Present, and Future. *Nutrients* **14**, 3201 (2022).

68. Barroso, I. *et al.* Candidate gene association study in type 2 diabetes indicates a role for genes involved in beta-cell function as well as insulin action. *PLoS Biol.* **1**, E20 (2003).

69. McIntyre, H. D. *et al.* Gestational diabetes mellitus. *Nat. Rev. Dis. Prim.* **5**, 47 (2019).

70. Buchanan, T. A. Pancreatic B-cell defects in gestational diabetes: implications for the pathogenesis and prevention of type 2 diabetes. *J. Clin. Endocrinol. Metab.* **86**, 989–993 (2001).

71. Huopio, H. *et al.* Association of risk variants for type 2 diabetes and hyperglycemia with gestational diabetes. *Eur. J. Endocrinol.* **169**, 291–297 (2013).

72. Lega, I. C. & Lipscombe, L. L. Review: Diabetes, Obesity, and Cancer—Pathophysiology and Clinical Implications. *Endocr. Rev.* **41**, 33–52 (2020).

73. Ryu, T. Y., Park, J. & Scherer, P. E. Hyperglycemia as a risk factor for cancer progression. *Diabetes Metab. J.* **38**, 330–336 (2014).

74. Vaughn, A. E. & Deshmukh, M. Glucose metabolism inhibits apoptosis in neurons and cancer cells by redox inactivation of cytochrome c. *Nat. Cell Biol.* **10**, 1477–1483 (2008).

75. Vella, V., Milluzzo, A., Scalisi, N. M., Vigneri, P. & Sciacca, L. Insulin Receptor Isoforms in Cancer. *Int. J. Mol. Sci.* **19**, 3615 (2018).

76. Scalia, P., Giordano, A. & Williams, S. J. The IGF-II-Insulin Receptor Isoform-A Autocrine Signal in Cancer: Actionable Perspectives. *Cancers (Basel).* **12**, 366 (2020).

77. Garofalo, R. S. Genetic analysis of insulin signaling in *Drosophila*. *Trends Endocrinol. Metab.* **13**, 156–162 (2002).

78. Fernandez, R., Tabarini, D., Azpiazu, N., Frasch, M. & Schlessinger, J. The *Drosophila* insulin receptor homolog: a gene essential for embryonic development encodes two receptor isoforms with different signaling potential. *EMBO J.* **14**, 3373–3384 (1995).

79. Böhni, R. *et al.* Autonomous control of cell and organ size by CHICO, a *Drosophila* homolog of vertebrate IRS1-4. *Cell* **97**, 865–875 (1999).

80. Slack, C. *et al.* The Ras-Erk-ETS-Signaling Pathway Is a Drug Target for Longevity. *Cell* **162**, 72–83 (2015).

81. Garelli, A., Gontijo, A. M., Miguela, V., Caparros, E. & Dominguez, M. Imaginal Discs Secrete Insulin-Like Peptide 8 to Mediate Plasticity of Growth and Maturation. *Science* **336**, 579–582 (2012).

82. Colombani, J. *et al. Drosophila* Lgr3 Couples Organ Growth with Maturation and Ensures Developmental Stability. *Curr. Biol.* **25**, 2723–2729 (2015).

83. Vallejo, D. M., Juarez-Carreno, S., Bolivar, J., Morante, J. & Dominguez, M. A brain circuit that synchronizes growth and maturation revealed through Dilp8 binding to Lgr3. *Science* **350**, aac6767 (2015).

84. Imambocus, B. N. *et al.* A neuropeptidergic circuit gates selective escape behavior of *Drosophila* larvae. *Curr. Biol.* **32**, 149-163.e8 (2021).

85. Sajid, W. *et al.* Structural and biological properties of the *Drosophila* insulin-like peptide 5 show evolutionary conservation. *J. Biol. Chem.* **286**, 661–673 (2011).

86. Jin Chan, S. & Steiner, D. F. Insulin Through the Ages: Phylogeny of a Growth Promoting and Metabolic Regulatory Hormone. *Am. Zool.* **40**, 213–222 (2000).

87. Irwin, D. M. Evolution of the Insulin Gene: Changes in Gene Number, Sequence, and Processing. *Front. Endocrinol.* **12**, 649255 (2021).

88. Jiaying Chan, L., Akhter Hossain, M., S. Samuel, C., Separovic, F. & D. Wade, J. The Relaxin Peptide Family – Structure, Function and Clinical Applications. *Protein & Peptide Letters* vol. 18 220–229 (2011).

89. Grönke, S., Clarke, D.-F., Broughton, S., Andrews, T. D. & Partridge, L. Molecular Evolution and Functional Characterization of *Drosophila* Insulin-Like Peptides. *PLOS Genet.* **6**, e1000857 (2010).

90. Rulifson, E. J., Kim, S. K. & Nusse, R. Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. *Science* **296**, 1118–1120 (2002).

91. Tatar, M. *et al.* A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science* **292**, 107–110 (2001).

92. Broughton, S. J. *et al.* Longer lifespan, altered metabolism, and stress resistance in *Drosophila* from ablation of cells making insulin-like ligands. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 3105–3110 (2005).

93. LaFever, L. & Drummond-Barbosa, D. Direct control of germline stem cell division and cyst growth by neural insulin in *Drosophila*. *Science* **309**, 1071–1073 (2005).

94. Brown, E. B., Shah, K. D., Faville, R., Kottler, B. & Keene, A. C. *Drosophila* insulin-like peptide 2 mediates dietary regulation of sleep intensity. *PLoS Genet.* **16**, e1008270 (2020).

95. Sudhakar, S. R. *et al.* Insulin signalling elicits hunger-induced feeding in *Drosophila*. *Dev. Biol.* **459**, 87–99 (2020).

96. de Velasco, B. *et al.* Specification and development of the pars intercerebralis and pars lateralis, neuroendocrine command centers in the *Drosophila* brain. *Dev. Biol.* **302**, 309–323 (2007).

97. Wang, S., Tulina, N., Carlin, D. L. & Rulifson, E. J. The origin of islet-like cells in *Drosophila* identifies parallels to the vertebrate endocrine axis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19873–19878 (2007).

98. Ikeya, T., Galic, M., Belawat, P., Nairz, K. & Hafen, E. Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in *Drosophila*. *Curr. Biol.* **12**, 1293–1300 (2002).

99. Okamoto, N. & Nishimura, T. Signaling from Glia and Cholinergic Neurons Controls Nutrient-Dependent Production of an Insulin-like Peptide for *Drosophila* Body Growth. *Dev. Cell* **35**, 295–310 (2015).

100. Liu, Y., Liao, S., Veenstra, J. A. & Nassel, D. R. *Drosophila* insulin-like peptide 1 (DILP1) is transiently expressed during non-feeding stages and reproductive dormancy. *Sci. Rep.* **6**, 26620 (2016).

101. Hartenstein, V. The neuroendocrine system of invertebrates: a developmental and evolutionary perspective. *J. Endocrinol.* **190**, 555–570 (2006).

102. Hwang, H. J. & Rulifson, E. Serial specification of diverse neuroblast identities from a neurogenic placode by Notch and Egfr signaling. *Development* **138**, 2883–2893 (2011).

103. Cao, J. *et al.* Insight into insulin secretion from transcriptome and genetic analysis of insulin-producing cells of *Drosophila*. *Genetics* **197**, 175–192 (2014).

104. Kim, S. K. & Rulifson, E. J. Conserved mechanisms of glucose sensing and regulation by *Drosophila* corpora cardiaca cells. *Nature* **431**, 316–320 (2004).

105. Lee, G. & Park, J. H. Hemolymph sugar homeostasis and starvation-induced hyperactivity affected by genetic manipulations of the adipokinetic hormone-encoding gene in *Drosophila* melanogaster. *Genetics* **167**, 311–323 (2004).

106. Siegmund, T. & Korge, G. Innervation of the ring gland of *Drosophila* melanogaster. *J. Comp. Neurol.* **431**, 481–491 (2001).

107. Hughson, B. N. The Glucagon-Like Adipokinetic Hormone in *Drosophila* melanogaster – Biosynthesis and Secretion. *Front. Physiol.* **12**, 710652 (2021).

108. Clements, J., Buhler, K., Winant, M., Vulsteke, V. & Callaerts, P. Glial and Neuronal Neuroglian, Semaphorin-1a and Plexin A Regulate Morphological and Functional Differentiation of *Drosophila* Insulin-Producing Cells. *Front. Endocrinol.* **12**, 600251 (2021).

109. Nässel, D. R. & Zandawala, M. Hormonal axes in *Drosophila*: regulation of hormone release and multiplicity of actions. *Cell Tissue Res.* **382**, 233–266 (2020).

110. Lin, F. *et al.* Total Solid-Phase Synthesis of Biologically Active *Drosophila* Insulin-Like Peptide 2 (DILP2). *Aust. J. Chem.* **70**, 208–212 (2017).

111. Hadžić, T. *et al.* Genome-wide features of neuroendocrine regulation in *Drosophila* by the basic helix-loop-helix transcription factor DIMMED. *Nucleic Acids Res.* **43**, 2199–2215 (2015).

112. Nässel, D. R. Insulin-producing cells and their regulation in physiology and behavior of *Drosophila*. *Can. J. Zool.* **90**, 476–488 (2012).

113. Hewes, R. S., Park, D., Gauthier, S. A., Schaefer, A. M. & Taghert, P. H. The bHLH protein Dimmed controls neuroendocrine cell differentiation in *Drosophila*. *Development* **130**, 1771–1781 (2003).

114. Park, D., Veenstra, J. A., Park, J. H. & Taghert, P. H. Mapping Peptidergic Cells in *Drosophila*: Where DIMM Fits In. *PLoS One* **3**, e1896 (2008).

115. Hamanaka, Y. *et al.* Transcriptional Orchestration of the Regulated Secretory Pathway in Neurons by the bHLH protein DIMM. *Curr. Biol.* **20**, 9–18 (2010).

116. Slaidina, M., Delanoue, R., Gronke, S., Partridge, L. & Leopold, P. A *Drosophila* insulin-like peptide promotes growth during nonfeeding states. *Dev. Cell* **17**, 874–884 (2009).

117. Okamoto, N. *et al.* A Fat Body-Derived IGF-like Peptide Regulates Postfeeding Growth in *Drosophila*. *Dev. Cell* **17**, 885–891 (2009).

118. Liao, S. *et al.* Regulatory Roles of *Drosophila* Insulin-Like Peptide 1 (DILP1) in Metabolism Differ in Pupal and Adult Stages. *Front. Endocrinol.* **11**, 180 (2020).

119. Post, S. *et al. Drosophila* insulin-like peptide dilp1 increases lifespan and glucagon-like Akh expression epistatic to dilp2. *Aging Cell* **18**, e12863 (2019).

120. Broughton, S. *et al.* Reduction of DILP2 in *Drosophila* triages a metabolic phenotype from lifespan revealing redundancy and compensation among DILPs. *PLoS One* **3**, e3721–e3721 (2008).

121. Géminard, C., Rulifson, E. J. & Léopold, P. Remote Control of Insulin Secretion by Fat Cells in *Drosophila*. *Cell Metab.* **10**, 199–207 (2009).

122. Chell, J. M. & Brand, A. H. Nutrition-responsive glia control exit of neural stem cells from quiescence. *Cell* **143**, 1161–1173 (2010).

123. Buhler, K. *et al.* Growth control through regulation of insulin signalling by nutrition-activated steroid hormone in *Drosophila*. *Development* **145**, dev165654 (2018).

124. O'Brien, L. E., Soliman, S. S., Li, X. & Bilder, D. Altered modes of stem cell division drive adaptive intestinal growth. *Cell* **147**, 603–614 (2011).

125. Söderberg, J. A. E., Birse, R. T. & Nässel, D. R. Insulin Production and Signaling in Renal Tubules of *Drosophila* Is under Control of Tachykinin-Related Peptide and Regulates Stress Resistance. *PLoS One* **6**, e19866 (2011).

126. Kim, J. & Neufeld, T. P. Dietary sugar promotes systemic TOR activation in *Drosophila* through AKH-dependent selective secretion of Dilp3. *Nat. Commun.* **6**, 6846 (2015).

127. Post, S. *et al. Drosophila* Insulin-Like Peptides DILP2 and DILP5 Differentially Stimulate Cell Signaling and Glycogen Phosphorylase to Regulate Longevity. *Front. Endocrinol.* **9**, 245 (2018).

128. Söderberg, J. A. E., Carlsson, M. A. & Nässel, D. R. Insulin-Producing Cells in the *Drosophila* Brain also Express Satiety-Inducing Cholecystokinin-Like Peptide, Drosulfakinin. *Front. Endocrinol.* **3**, 109 (2012).

129. Ohhara, Y., Kobayashi, S., Yamakawa-Kobayashi, K. & Yamanaka, N. Adult-specific insulin-producing neurons in *Drosophila* melanogaster. *J. Comp. Neurol.* **526**, 1351–1367 (2018).

130. Park, S. *et al.* A Genetic Strategy to Measure Circulating *Drosophila* Insulin Reveals Genes Regulating Insulin Production and Secretion. *PLoS Genet.* **10**, e1004555 (2014).

131. Kréneisz, O., Chen, X., Fridell, Y.-W. C. & Mulkey, D. K. Glucose increases activity and Ca2+ in insulin-producing cells of adult *Drosophila*. *Neuroreport* **21**, 1116–1120 (2010).

132. Manière, G., Ziegler, A. B., Geillon, F., Featherstone, D. E. & Grosjean, Y. Direct Sensing of Nutrients via a LAT1-like Transporter in *Drosophila* Insulin-Producing Cells. *Cell Rep.* **17**, 137–148 (2016).

133. Ziegler, A. B., Manière, G. & Grosjean, Y. JhI-21 plays a role in *Drosophila* insulin-like peptide release from larval IPCs via leucine transport. *Sci. Rep.* **8**, 1908 (2018).

134. Barry, W. E. & Thummel, C. S. The *Drosophila* HNF4 nuclear receptor promotes

glucose-stimulated insulin secretion and mitochondrial function in adults. *Elife* **5**, e11183 (2016).

135. Colombani, J. *et al.* A nutrient sensor mechanism controls *Drosophila* growth. *Cell* **114**, 739–749 (2003).

136. Delanoue, R. *et al. Drosophila* insulin release is triggered by adipose Stunted ligand to brain Methuselah receptor. *Science* **353**, 1553–1556 (2016).

137. Koyama, T. & Mirth, C. K. Growth-Blocking Peptides As Nutrition-Sensitive Signals for Insulin Secretion and Body Size Regulation. *PLOS Biol.* **14**, e1002392 (2016).

138. Meschi, E., Léopold, P. & Delanoue, R. An EGF-Responsive Neural Circuit Couples Insulin Secretion with Nutrition in *Drosophila*. *Dev. Cell* **48**, 76-86.e5 (2019).

139. Sun, J. *et al. Drosophila* FIT is a protein-specific satiety hormone essential for feeding control. *Nat. Commun.* **8**, 14161 (2017).

140. Rajan, A. & Perrimon, N. *Drosophila* cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion. *Cell* **151**, 123–137 (2012).

141. Brent, A. E. & Rajan, A. Insulin and Leptin/Upd2 Exert Opposing Influences on Synapse Number in Fat-Sensing Neurons. *Cell Metab.* **32**, 786-800.e7 (2020).

142. Agrawal, N. *et al.* The *Drosophila* TNF Eiger Is an Adipokine that Acts on Insulin-Producing Cells to Mediate Nutrient Response. *Cell Metab.* **23**, 675–684 (2016).

143. Alfa, R. W. *et al.* Suppression of Insulin Production and Secretion by a Decretin Hormone. *Cell Metab.* **21**, 323–334 (2015).

144. Oh, Y. *et al.* A glucose-sensing neuron pair regulates insulin and glucagon in *Drosophila*. *Nature* **574**, 559–564 (2019).

145. Kapan, N., Lushchak, O. V, Luo, J. & Nässel, D. R. Identified peptidergic neurons in the *Drosophila* brain regulate insulin-producing cells, stress responses and metabolism by coexpressed short neuropeptide F and corazonin. *Cell. Mol. Life Sci.* **69**, 4051–4066 (2012).

146. Luo, J., Becnel, J., Nichols, C. D. & Nässel, D. R. Insulin-producing cells in the brain of adult *Drosophila* are regulated by the serotonin 5-HT1A receptor. *Cell. Mol. Life Sci.* **69**, 471–484 (2012).

147. Luo, J., Lushchak, O. V, Goergen, P., Williams, M. J. & Nässel, D. R. *Drosophila* insulin-producing cells are differentially modulated by serotonin and octopamine receptors and affect social behavior. *PLoS One* **9**, e99732 (2014).

148. Liessem, S. *et al.* Behavioral state-dependent modulation of insulin-producing cells in *Drosophila*. *Curr. Biol.* **33**, 449-463.e5 (2023).

149. Clements, J., Hens, K., Francis, C., Schellens, A. & Callaerts, P. Conserved role for the *Drosophila* Pax6 homolog Eyeless in differentiation and function of insulin-producing neurons. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16183–16188 (2008).

150. Okamoto, N., Nishimori, Y. & Nishimura, T. Conserved role for the Dachshund protein with *Drosophila* Pax6 homolog Eyeless in insulin expression. *Proc. Natl. Acad. Sci.* **109**, 2406–2411 (2012).

151. Varghese, J., Lim, S. F. & Cohen, S. M. *Drosophila* miR-14 regulates insulin production and metabolism through its target, sugarbabe. *Genes Dev.* **24**, 2748–2753 (2010).

152. Docherty, J. E. B., Manno, J. E., McDermott, J. E. & DiAngelo, J. R. Mio acts in the

Drosophila brain to control nutrient storage and feeding. *Gene* **568**, 190–195 (2015).

153. Alfa, R. W. & Kim, S. K. Using *Drosophila* to discover mechanisms underlying type 2 diabetes. *Dis. Model. Mech.* **9**, 365–376 (2016).

154. Honegger, B. *et al.* Imp-L2, a putative homolog of vertebrate IGF-binding protein 7, counteracts insulin signaling in *Drosophila* and is essential for starvation resistance. *J. Biol.* **7**, 10 (2008).

155. Arquier, N. *et al. Drosophila* ALS regulates growth and metabolism through functional interaction with insulin-like peptides. *Cell Metab.* **7**, 333–338 (2008).

156. Ghosh, S. *et al.* A local insulin reservoir in *Drosophila* alpha cell homologs ensures developmental progression under nutrient shortage. *Curr. Biol.* **32**, 1788-1797.e5 (2022).

157. Okamoto, N. *et al.* A secreted decoy of InR antagonizes insulin/IGF signaling to restrict body growth in *Drosophila*. *Genes Dev.* **27**, 87–97 (2013).

158. Atouf, F., Czernichow, P. & Scharfmann, R. Expression of neuronal traits in pancreatic beta cells: implication of neuron-restrictive silencing factor/repressor element silencing transcription factor, a neuron-restrictive silencer. *J. Biol. Chem.* **272**, 1929–1934 (1997).

159. Sherwood, N. M., Adams, B. A. & Tello, J. A. Endocrinology of protochordates. *Can. J. Zool.* **83**, 225–255 (2005).

160. Smit, A. B. *et al.* Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. *Nature* **331**, 535–538 (1988).

161. Nakamura, M. *et al.* Autocrine insulin pathway signaling regulates actin dynamics in cell wound repair. *PLOS Genet.* **16**, e1009186 (2020).

162. de Tredern, E. *et al.* Glial glucose fuels the neuronal pentose phosphate pathway for long-term memory. *Cell Rep.* **36**, 109620 (2021).

163. Jünger, M. A. *et al.* The *Drosophila* forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling. *J. Biol.* **2**, 20 (2003).

164. Sousa-Nunes, R., Yee, L. L. & Gould, A. P. Fat cells reactivate quiescent neuroblasts via TOR and glial insulin relays in *Drosophila*. *Nature* **471**, 508–512 (2011).

165. Miguel-Aliaga, I., Thor, S. & Gould, A. P. Postmitotic Specification of *Drosophila* Insulinergic Neurons from Pioneer Neurons. *PLOS Biol.* **6**, e58 (2008).

166. Yang, C., Belawat, P., Hafen, E., Jan, L. Y. & Jan, Y.-N. *Drosophila* Egg-Laying Site Selection as a System to Study Simple Decision-Making Processes. *Science* **319**, 1679–1683 (2008).

167. Cognigni, P., Bailey, A. P. & Miguel-Aliaga, I. Enteric neurons and systemic signals couple nutritional and reproductive status with intestinal homeostasis. *Cell Metab.* **13**, 92–104 (2011).

168. Castellanos, M. C., Tang, J. C. Y. & Allan, D. W. Female-biased dimorphism underlies a female-specific role for post-embryonic Ilp7 neurons in *Drosophila* fertility. *Development* **140**, 3915–3926 (2013).

169. Veenstra, J. A., Rombauts, S. & Grbić, M. In silico cloning of genes encoding neuropeptides, neurohormones and their putative G-protein coupled receptors in a spider mite. *Insect Biochem. Mol. Biol.* **42**, 277–295 (2012).

170. Colombani, J., Andersen, D. S. & Léopold, P. Secreted Peptide Dilp8 Coordinates

Drosophila Tissue Growth with Developmental Timing. *Science* **336**, 582–585 (2012).

171. Liao, S. & Nässel, D. R. *Drosophila* Insulin-Like Peptide 8 (DILP8) in Ovarian Follicle Cells Regulates Ovulation and Metabolism. *Front. Endocrinol.* **11**, 461 (2020).

172. Giannakou, M. E. & Partridge, L. Role of insulin-like signalling in *Drosophila* lifespan. *Trends Biochem. Sci.* **32**, 180–188 (2007).

173. Itskov, P. & Ribeiro, C. The Dilemmas of the Gourmet Fly: The Molecular and Neuronal Mechanisms of Feeding and Nutrient Decision Making in *Drosophila*. *Front. Neurosci.* **7**, 12 (2013).

174. Vitali, V., Horn, F. & Catania, F. Insulin-like signaling within and beyond metazoans. *Biol. Chem.* **399**, 851–857 (2018).

175. Garelli, A. *et al.* Dilp8 requires the neuronal relaxin receptor Lgr3 to couple growth to developmental timing. *Nat. Commun.* **6**, 8732 (2015).

176. Veenstra, J. A., Agricola, H.-J. & Sellami, A. Regulatory peptides in fruit fly midgut. *Cell Tissue Res.* **334**, 499–516 (2008).

177. Kannan, K. & Fridell, Y.-W. C. Functional implications of *Drosophila* insulin-like peptides in metabolism, aging, and dietary restriction. *Front. Physiol.* **4**, 288 (2013).

178. Nässel, D. R., Kubrak, O. I., Liu, Y., Luo, J. & Lushchak, O. V. Factors that regulate insulin producing cells and their output in *Drosophila*. *Front. Physiol.* **4**, 252 (2013).

179. Mackay, T. F. C. *et al.* The *Drosophila* melanogaster Genetic Reference Panel. *Nature* **482**, 173–178 (2012).

180. Shorter, J. *et al.* Genetic architecture of natural variation in *Drosophila* melanogaster aggressive behavior. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3555-63 (2015).

181. Zwarts, L. *et al.* The genetic basis of natural variation in mushroom body size in *Drosophila* melanogaster. *Nat. Commun.* **6**, 10115 (2015).

182. Baker, B. M., Carbone, M. A., Huang, W., Anholt, R. R. H. & Mackay, T. F. C. Genetic basis of variation in cocaine and methamphetamine consumption in outbred populations of *Drosophila* melanogaster. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

183. Ivanov, D. K. *et al.* Longevity GWAS Using the *Drosophila* Genetic Reference Panel. *J. Gerontol. A. Biol. Sci. Med. Sci.* **70**, 1470–1478 (2015).

184. Huang, W. *et al.* Genetic basis of transcriptome diversity in *Drosophila* melanogaster. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E6010-9 (2015).

185. Kumar, J. P. Retinal determination the beginning of eye development. *Curr. Top. Dev. Biol.* **93**, 1–28 (2010).

186. Clements, J. *et al.* Mutational analysis of the eyeless gene and phenotypic rescue reveal that an intact Eyeless protein is necessary for normal eye and brain development in *Drosophila*. *Dev. Biol.* **334**, 503–512 (2009).

187. Abu-Shaar, M., Ryoo, H. D. & Mann, R. S. Control of the nuclear localization of Extradenticle by competing nuclear import and export signals. *Genes Dev.* **13**, 935–945 (1999).

188. Zhang, X. *et al.* Pax6 is regulated by Meis and Pbx homeoproteins during pancreatic development. *Dev. Biol.* **300**, 748–757 (2006).

189. Owa, T. *et al.* Meis1 Coordinates Cerebellar Granule Cell Development by Regulating Pax6 Transcription, BMP Signaling and Atoh1 Degradation. *J. Neurosci.* **38**, 1277–1294

(2018).

190. Park, S. *et al.* A Genetic Strategy to Measure Circulating *Drosophila* Insulin Reveals Genes Regulating Insulin Production and Secretion. *PLoS Genet.* **10**, e1004555 (2014).

191. Mirth, C., Truman, J. W. & Riddiford, L. M. The Role of the Prothoracic Gland in Determining Critical Weight for Metamorphosis in *Drosophila* melanogaster. *Curr. Biol.* **15**, 1796–1807 (2005).

192. Shingleton, A. W., Das, J., Vinicius, L. & Stern, D. L. The temporal requirements for insulin signaling during development in *Drosophila*. *PLoS Biol.* **3**, e289 (2005).

193. Rehman, N. & Varghese, J. Larval nutrition influences adult fat stores and starvation resistance in *Drosophila*. *PLoS One* **16**, e0247175 (2021).

194. Bessa, J., Gebelein, B., Pichaud, F., Casares, F. & Mann, R. S. Combinatorial control of *Drosophila* eye development by eyeless, homothorax, and teashirt. *Genes Dev.* **16**, 2415–2427 (2002).

195. Lopes, C. S. & Casares, F. hth maintains the pool of eye progenitors and its downregulation by Dpp and Hh couples retinal fate acquisition with cell cycle exit. *Dev. Biol.* **339**, 78–88 (2010).

196. Post, S. & Tatar, M. Nutritional Geometric Profiles of Insulin/IGF Expression in *Drosophila* melanogaster. *PLoS One* **11**, e0155628 (2016).

197. Peiris, H. *et al.* Discovering human diabetes-risk gene function with genetics and physiological assays. *Nat. Commun.* **9**, 3855 (2018).

198. Bürgy-Roukala, E., Miellet, S., Mishra, A. K. & Sprecher, S. G. Early Eye Development: Specification and Determination BT - Molecular Genetics of Axial Patterning, Growth and Disease in the *Drosophila* Eye. in (eds. Singh, A. & Kango-Singh, M.) 1–36 (Springer New York, 2013). doi:10.1007/978-1-4614-8232-1_1.

199. Adachi, Y. *et al.* Conserved cis-regulatory modules mediate complex neural expression patterns of the eyeless gene in the *Drosophila* brain. *Mech. Dev.* **120**, 1113–1126 (2003).

200. Huang, W. *et al.* Natural variation in genome architecture among 205 *Drosophila* melanogaster Genetic Reference Panel lines. *Genome Res.* **24**, 1193–1208 (2014).

201. Nagoshi, E. *et al.* Dissecting differential gene expression within the circadian neuronal circuit of *Drosophila*. *Nat. Neurosci.* **13**, 60–68 (2010).

202. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

203. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

204. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2015).

205. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

206. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

207. Van Handel, E. Rapid determination of total lipids in mosquitoes. *J. Am. Mosq. Control Assoc.* **1**, 302–304 (1985).

208. Torquato, L. S., Mattos, D., Matta, B. P. & Bitner-Mathé, B. C. Cellular basis of morphological variation and temperature-related plasticity in *Drosophila* melanogaster strains with divergent wing shapes. *Genetica* **142**, 495–505 (2014).

209. Castro-Mondragon, J. A. *et al.* JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **50**, D165–D173 (2022).

210. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).

211. Sun, H. *et al.* IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res. Clin. Pract.* **183**, 109119 (2022).

212. Galicia-Garcia, U. *et al.* Pathophysiology of type 2 diabetes mellitus. *Int. J. Mol. Sci.* **21**, 1–34 (2020).

213. Sanches, J. M., Zhao, L. N., Salehi, A., Wollheim, C. B. & Kaldis, P. Pathophysiology of type 2 diabetes and the impact of altered metabolic interorgan crosstalk. *FEBS J.* **290**, 620–648 (2023).

214. McCarthy, M. I. & Zeggini, E. Genome-wide association studies in type 2 diabetes. *Curr. Diab. Rep.* **9**, 164–171 (2009).

215. Del Bosque-Plata, L., Martínez-Martínez, E., Espinoza-Camacho, M. Á. & Gragnoli, C. The Role of TCF7L2 in Type 2 Diabetes. *Diabetes* **70**, 1220–1228 (2021).

216. Grant, S. F. A. *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323 (2006).

217. Sladek, R. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881–885 (2007).

218. Nusse, R. & Varmus, H. E. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* **31**, 99–109 (1982).

219. Nusse, R. & Varmus, H. E. Wnt genes. *Cell* **69**, 1073–1087 (1992).

220. Cadigan, K. M. & Waterman, M. L. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harb. Perspect. Biol.* **4**, a007906 (2012).

221. Nusse, R. WNT targets. Repression and activation. *Trends Genet.* **15**, 1–3 (1999).

222. Chen, X. *et al.* The Diabetes Gene and Wnt Pathway Effector TCF7L2 Regulates Adipocyte Development and Function. *Diabetes* **67**, 554–568 (2018).

223. Boj, S. F. *et al.* Diabetes Risk Gene and Wnt Effector Tcf7l2/TCF4 Controls Hepatic Response to Perinatal and Adult Metabolic Demand. *Cell* **151**, 1595–1607 (2012).

224. Schäfer, S. A. *et al.* Impaired glucagon-like peptide-1-induced insulin secretion in carriers of transcription factor 7-like 2 (TCF7L2) gene polymorphisms. *Diabetologia* **50**, 2443–2450 (2007).

225. Takamoto, I. *et al.* TCF7L2 in mouse pancreatic beta cells plays a crucial role in glucose homeostasis by regulating beta cell mass. *Diabetologia* **57**, 542–553 (2014).

226. Facchinello, N. *et al.* Tcf7l2 plays pleiotropic roles in the control of glucose homeostasis,

pancreas morphology, vascularization and regeneration. *Sci. Rep.* **7**, 9605 (2017).

227. Musselman, L. P. *et al.* A high-sugar diet produces obesity and insulin resistance in wild-type *Drosophila*. *Dis. Model. Mech.* **4**, 842–849 (2011).

228. Birse, R. T. *et al.* High-fat-diet-induced obesity and heart dysfunction are regulated by the TOR pathway in *Drosophila*. *Cell Metab.* **12**, 533–544 (2010).

229. Winant, M. *et al.* Genome-wide analysis identifies Homothorax and Extradenticle as regulators of insulin in *Drosophila* Insulin-Producing cells. *PLoS Genet.* **18**, e1010380 (2022).

230. Gosmain, Y. *et al.* Pax6 is crucial for β-cell function, insulin biosynthesis, and glucose-induced insulin secretion. *Mol. Endocrinol.* **26**, 696–709 (2012).

231. Sander, M. *et al.* Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. *Genes Dev.* **11**, 1662–1673 (1997).

232. Li, H. *et al.* Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science* **375**, eabk2432 (2022).

233. Mirth, C. K. & Shingleton, A. W. Integrating body and organ size in *Drosophila*: Recent advances and outstanding problems. *Front. Endocrinol.* **3**, 49 (2012).

234. Overton, P. M., Chia, W. & Buescher, M. The *Drosophila* HMG-domain proteins SoxNeuro and Dichaete direct trichome formation via the activation of shavenbaby and the restriction of Wingless pathway activity. *Development* **134**, 2807–2813 (2007).

235. Chao, A. T., Jones, W. M. & Bejsovec, A. The HMG-box transcription factor SoxNeuro acts with Tcf to control Wg/Wnt signaling activity. *Development* **134**, 989–997 (2007).

236. Cavallo, R. A. *et al. Drosophila* Tcf and Groucho interact to repress Wingless signalling activity. *Nature* **395**, 604–608 (1998).

237. Simionato, E. *et al.* Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* **7**, 33 (2007).

238. Direnzo, D. *et al.* Induced Mist1 expression promotes remodeling of mouse pancreatic acinar cells. *Gastroenterology* **143**, 469–480 (2012).

239. Pin, C. L., Rukstalis, J. M., Johnson, C. & Konieczny, S. F. The bHLH transcription factor Mist1 is required to maintain exocrine pancreas cell organization and acinar cell identity. *J. Cell Biol.* **155**, 519–530 (2001).

240. Luo, X., Shin, D. M., Wang, X., Konieczny, S. F. & Muallem, S. Aberrant localization of intracellular organelles, Ca2+ signaling, and exocytosis in Mist1 null mice. *J. Biol. Chem.* **280**, 12668–12675 (2005).

241. Zhou, Y. *et al.* TCF7L2 is a master regulator of insulin production and processing. *Hum. Mol. Genet.* **23**, 6419–6431 (2014).

242. Lee, J., Kim, K., Yu, S. W. & Kim, E. K. Wnt3a upregulates brain-derived insulin by increasing NeuroD1 via Wnt/β-catenin signaling in the hypothalamus. *Mol. Brain* **9**, 24 (2016).

243. Barber, A. F., Erion, R., Holmes, T. C. & Sehgal, A. Circadian and feeding cues integrate to drive rhythms of physiology in *Drosophila* insulin-producing cells. *Genes Dev.* **30**, 2596–2606 (2016).

244. Bylund, M., Andersson, E., Novitch, B. G. & Muhr, J. Vertebrate neurogenesis is

counteracted by Sox1-3 activity. *Nat. Neurosci.* **6**, 1162–1168 (2003).

245. Ferrero, E., Fischer, B. & Russell, S. SoxNeuro orchestrates central nervous system specification and differentiation in *Drosophila* and is only partially redundant with Dichaete. *Genome Biol.* **15**, R74 (2014).

246. Seymour, P. A. *et al.* SOX9 is required for maintenance of the pancreatic progenitor cell pool. *Proc. Natl. Acad. Sci.* **104**, 1865–1870 (2007).

247. Jenett, A. *et al.* A GAL4-Driver Line Resource for *Drosophila* Neurobiology. *Cell Rep.* **2**, 991–1001 (2012).

248. Veenstra, J. A. Arthropod IGF, relaxin and gonadulin, putative orthologs of *Drosophila* insulin-like peptides 6, 7 and 8, likely originated from an ancient gene triplication. *PeerJ* **8**, e9534 (2020).

249. Koyama, T., Texada, M. J., Halberg, K. A. & Rewitz, K. Metabolism and growth adaptation to environmental conditions in *Drosophila*. *Cell. Mol. Life Sci.* **77**, 4523–4551 (2020).

250. Reece-Hoyes, J. S. & Marian Walhout, A. J. Yeast one-hybrid assays: a historical and technical perspective. *Methods* **57**, 441–447 (2012).

251. Hens, K., Feuz, J.-D. & Deplancke, B. A High-throughput Gateway-Compatible Yeast One-Hybrid Screen to Detect Protein–DNA Interactions BT - Gene Regulatory Networks: Methods and Protocols. in *Methods in Molecular Biology* (eds. Deplancke, B. & Gheldof, N.) vol. 786 335–355 (Humana Press, 2012).

252. Hens, K. *et al.* Automated protein-DNA interaction screening of *Drosophila* regulatory elements. *Nat. Methods* **8**, 1065–1070 (2011).

253. Czerny, T. *et al.* twin of eyeless, a Second Pax-6 Gene of *Drosophila*, Acts Upstream of eyeless in the Control of Eye Development. *Mol. Cell* **3**, 297–307 (1999).

254. Dimova, D. K., Stevaux, O., Frolov, M. V & Dyson, N. J. Cell cycle-dependent and cell cycle-independent control of transcription by the *Drosophila* E2F/RB pathway. *Genes Dev.* **17**, 2308–2320 (2003).

255. Yoshioka, Y., Suyari, O. & Yamaguchi, M. Transcription factor NF-Y is involved in regulation of the JNK pathway during *Drosophila* thorax development. *Genes to Cells* **13**, 117–130 (2008).

256. Yoshioka, Y., Ly, L. L. & Yamaguchi, M. Transcription factor NF-Y is involved in differentiation of R7 photoreceptor cell in *Drosophila*. *Biol. Open* **1**, 19–29 (2012).

257. Liu, Y. *et al.* Nuclear Factor-Y in Mouse Pancreatic β-Cells Plays a Crucial Role in Glucose Homeostasis by Regulating β-Cell Mass and Insulin Secretion. *Diabetes* **70**, 1703–1716 (2021).

258. Astle, J., Kozlova, T. & Thummel, C. S. Essential roles for the Dhr78 orphan nuclear receptor during molting of the *Drosophila* tracheal system. *Insect Biochem. Mol. Biol.* **33**, 1201–1209 (2003).

259. Punzo, C. *et al.* Functional divergence between eyeless and twin of eyeless in *Drosophila* melanogaster. *Development* **131**, 3943–3953 (2004).

260. Gera, T., Jonas, F., More, R. & Barkai, N. Evolution of binding preferences among whole-genome duplicated transcription factors. *Elife* **11**, e73225 (2022).

261. Winberg, M. L. *et al.* Plexin A is a neuronal semaphorin receptor that controls axon guidance. *Cell* **95**, 903–916 (1998).

262. Heigwer, F., Port, F. & Boutros, M. RNA Interference (RNAi) Screening in *Drosophila*. *Genetics* **208**, 853–874 (2018).

263. Wang, Y., Wang, L. & Wang, Z. Transgenic analyses of TGIF family proteins in *Drosophila* imply their role in cell growth. *J. Genet. Genomics* **35**, 457–465 (2008).

264. Golson, M. L. & Kaestner, K. H. Fox transcription factors: from development to disease. *Development* **143**, 4558–4570 (2016).

265. Sakkou, M. *et al.* A Role for Brain-Specific Homeobox Factor Bsx in the Control of Hyperphagia and Locomotory Behavior. *Cell Metab.* **5**, 450–463 (2007).

266. Fogarty, M. P., Panhuis, T. M., Vadlamudi, S., Buchkovich, M. L. & Mohlke, K. L. Allele-Specific Transcriptional Activity at Type 2 Diabetes–Associated Single Nucleotide Polymorphisms in Regions of Pancreatic Islet Open Chromatin at the JAZF1 Locus. *Diabetes* **62**, 1756–1762 (2013).

267. Eizirik, D. L. *et al.* The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-inflammatory cytokines. *PLoS Genet.* **8**, e1002552 (2012).

268. Iglesias, A. *et al.* Diabetes and exocrine pancreatic insufficiency in E2F1/E2F2 double-mutant mice. *J. Clin. Invest.* **113**, 1398–1407 (2004).

269. Cebola, I. *et al.* TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nat. Cell Biol.* **17**, 615–626 (2015).

270. Rosselot, C. *et al.* The many lives of Myc in the pancreatic β-cell. *J. Biol. Chem.* **296**, 100122 (2021).

271. Sund, N. J. *et al.* Tissue-specific deletion of Foxa2 in pancreatic beta cells results in hyperinsulinemic hypoglycemia. *Genes Dev.* **15**, 1706–1715 (2001).

272. Bastidas-Ponce, A. *et al.* Foxa2 and Pdx1 cooperatively regulate postnatal maturation of pancreatic β-cells. *Mol. Metab.* **6**, 524–534 (2017).

273. Spaeth, J. M. *et al.* The FOXP1, FOXP2 and FOXP4 transcription factors are required for islet alpha cell proliferation and function in mice. *Diabetologia* **58**, 1836–1844 (2015).

274. Spagnoli, F. M. & Brivanlou, A. H. The Gata5 target, TGIF2, defines the pancreatic region by modulating BMP signals within the endoderm. *Development* **135**, 451–461 (2008).

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

276. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).

277. Castillo-Armengol, J., Fajas, L. & Lopez-Mejia, I. C. Inter-organ communication: a gatekeeper for metabolic health. *EMBO Rep.* **20**, e47903 (2019).

278. Baumgardt, M., Miguel-Aliaga, I., Karlsson, D., Ekman, H. & Thor, S. Specification of Neuronal Identities by Feedforward Combinatorial Coding. *PLOS Biol.* **5**, e37 (2007).

279. Benito-Sipos, J. *et al.* A genetic cascade involving klumpfuss, nab and castor specifies the abdominal leucokinergic neurons in the *Drosophila* CNS. *Development* **137**, 3327–3336 (2010).

280. Baumgardt, M., Karlsson, D., Terriente, J., Díaz-Benjumea, F. J. & Thor, S. Neuronal

Subtype Specification within a Lineage by Opposing Temporal Feed-Forward Loops. *Cell* **139**, 969–982 (2009).

281. Stratmann, J. & Thor, S. Neuronal cell fate specification by the molecular convergence of different spatio-temporal cues on a common initiator terminal selector gene. *PLoS Genet.* **13**, e1006729 (2017).

282. Birse, R. T., Söderberg, J. A. E., Luo, J., Winther, Å. M. E. & Nässel, D. R. Regulation of insulin-producing cells in the adult *Drosophila* brain via the tachykinin peptide receptor DTKR. *J. Exp. Biol.* **214**, 4201–4208 (2011).

283. Nowak, M. A., Boerlijst, M. C., Cooke, J. & Smith, J. M. Evolution of genetic redundancy. *Nature* **388**, 167–171 (1997).

284. Sacerdot, C., Louis, A., Bon, C., Berthelot, C. & Roest Crollius, H. Chromosome evolution at the origin of the ancestral vertebrate genome. *Genome Biol.* **19**, 166 (2018).

285. Heller, R. S. The comparative anatomy of islets. *Adv. Exp. Med. Biol.* **654**, 21–37 (2010).

286. van Arensbergen, J. *et al.* Derepression of Polycomb targets during pancreatic organogenesis allows insulin-producing beta-cells to adopt a neural gene activity program. *Genome Res.* **20**, 722–732 (2010).

287. Arntfield, M. E. & van der Kooy, D. β-Cell evolution: How the pancreas borrowed from the brain: The shared toolbox of genes expressed by neural and pancreatic endocrine cells may reflect their evolutionary relationship. *Bioessays* **33**, 582–587 (2011).

288. Pelling, M. *et al.* Differential requirements for neurogenin 3 in the development of POMC and NPY neurons in the hypothalamus. *Dev. Biol.* **349**, 406–416 (2011).

289. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497–1502 (2007).

290. Dakic, T. *et al.* The Expression of Insulin in the Central Nervous System: What Have We Learned So Far? *Int. J. Mol. Sci.* **24**, 6586 (2023).

291. Huygens, C. *et al.* Evolutionary diversification of insulin-related peptides (IRPs) in aphids and spatiotemporal distribution in Acyrthosiphon pisum. *Insect Biochem. Mol. Biol.* **141**, 103670 (2022).

292. Akira, M. *et al.* A monoclonal antibody against a synthetic fragment of bombyxin (4K-prothoracicotropic hormone) from the silkmoth, Bombyx mori: characterization and immunohistochemistry. *Mol. Cell. Endocrinol.* **51**, 227–235 (1987).

293. Xiang, X. Towards understanding the ancestral role of insulin-like-peptides (ILPs) in Hydra. (2016).

294. Verbruggen, V. *et al.* The Pax6b Homeodomain Is Dispensable for Pancreatic Endocrine Cell Differentiation in Zebrafish. *J. Biol. Chem.* **285**, 13863–13873 (2010).

295. Xu, F. *et al.* Evidence from oyster suggests an ancient role for Pdx in regulating insulin gene expression in animals. *Nat. Commun.* **12**, 3117 (2021).

296. Yee, N. S., Yusuff, S. & Pack, M. Zebrafish pdx1 morphant displays defects in pancreas development and digestive organ chirality, and potentially identifies a multipotent pancreas progenitor cell. *Genesis* **30**, 137–140 (2001).

297. Ferrier, D. E. & Holland, P. W. Sipunculan ParaHox genes. *Evol. Dev.* **3**, 263–270 (2001).

Scientific acknowledgments and personal contribution

Chapter 3

Mattias Winant performed all experiments related to the data generated in chapter 3, except for those in Figure 11A-B, Figure 13 A-C, Figure 13 E-E', Figure 13 G-G', Figure 14 A-B, Figure 14 E-E', Figure 15E, Figure 16D, Figure 17 A-K, Figure 17 M-S, Figure 17 X-Z, Figure S1, Figure S3 and Figure S9.

Korneel Hens generated the adult IPC transcriptome. Kurt Buhler performed IHC related to figures 13, 14, 16, 17, S3 and S9. Jason Clements performed IHC related to Figure 17 and qPCR related to Figure S1. Sofie De Groef scraped DGRP data for Figure 11 A-B, generated qPCR data related to Figure 17 and TAG measurements of figure 15 E. Veerle Vulsteke assisted with qPCR.

Mattias Winant and Patrick Callaerts wrote the manuscript based on drafts from Kurt Buhler.

Chapter 4

Mattias Winant performed all experiments related to Chapter 4.

Mattias Winant and Patrick Callaerts wrote the manuscript.

Chapter 5

Korneel Hens performed and analyzed Yeast-1-Hybrid screen of the *dilp2*, *-3* and *-5* proximal promoter regions. Tom Wilms performed RNA extraction and qPCR experiments. Mattias Winant analyzed qPCR data, performed all IHC experiments and weight measurements.

Mattias Winant and Patrick Callaerts wrote the manuscript.

Conflict of interest statement

This work has been created in the absence of any conflicts of interest.

Publications

Winant, M.[#], Buhler K.[#], Clements J[#], De Groef, S., Hens, K., Vulsteke V. & Callaerts, P. Genome-wide analysis identifies Homothorax and Extradenticle as regulators of insulin in *Drosophila* Insulin-Producing cells. *PLoS Genet.* **18**, e1010380 (2022).

Clements, J., Buhler, K., **Winant, M**., Vulsteke, V. & Callaerts, P. Glial and Neuronal Neuroglian, Semaphorin-1a and Plexin A Regulate Morphological and Functional Differentiation of *Drosophila* Insulin-Producing Cells. *Front. Endocrinol*. **12**, 600251 (2021).

Buhler, K.[#], Clements, J.[#], **Winant, M**., Bolckmans, L., Vulsteke, V. & Callaerts, P. Growth control through regulation of insulin signalling by nutrition-activated steroid hormone in *Drosophila*. *Development* **145**, dev165654 (2018).

# Authors contributed equally.

Conferences

2018 Lowlands Fly Meeting, Utrecht, The Netherlands. How are growth, maturation and nutrition integrated? Oral Presentation

2021 Lowlands Fly Meeting, Leuven, Belgium. Genes and gene networks regulating *Drosophila* Insulin-Producing Cell development and function. Oral Presentation

2021 Genetics and Genomics Symposium, Leuven, Belgium. Genes and gene networks regulating *Drosophila* Insulin-Producing Cell development and function. Oral Presentation

2022 Neurofly Meeting, Saint Malo, France. GWAS identifies Homothorax and Extradenticle as regulators of insulin in *Drosophila* Insulin-Producing Cells. Poster Presentation