Escola Superior de Saúde Egas Moniz

Mestrado em Biologia Molecular em Saúde

Characterizing the function of two *Drosophila* Leucine-rich repeat-containing G-protein coupled receptors

Andreia Palos Casimiro

Dissertação para obtenção do grau de Mestre em Biologia Molecular em Saúde

Setembro 2014

Escola Superior de Saúde Egas Moniz

Mestrado em Biologia Molecular em Saúde

Characterizing the function of two *Drosophila* Leucine-rich repeat-containing G-protein coupled receptors

Andreia Palos Casimiro

Dissertação orientada pelo Doutor Álisson Gontijo e pela Dra Fabiana Herédia.

Dissertação para obtenção do grau de Mestre em Biologia Molecular em Saúde

Setembro 2014

Agradecimentos

Aos meus orientadores, Álisson Gontijo e Fabiana Herédia, pela compreensão, confiança, paciência e tempo despendido e acima de tudo, por me ensinarem como fazer boa ciência. Agradeço também, por todo o incentivo e ajuda que me foram dando ao longo do trabalho, principalmente nos últimos meses, assim como por todas as oportunidades que me têm proporcionado neste último ano.

A todos os meus colegas do Integrative Biomedicine Laboratory, especialmente à Ângela por no início me ter transmitido os seus conhecimentos e me ter ajudado quando mais precisei e à Catarina pelo apoio e ajuda que me deu nos últimos tempos. Espero poder retribuir um dia.

Às minhas amigas, Sara, Ana e Vanessa, por muitas vezes terem aturado o meu "mau feitio" e a minha falta de tempo, mas nunca me terem deixado de apoiar e incentivar, estando sempre disponíveis para ouvir os meus desabafos.

À minha família: Mãe, Pai e Irmã, por me terem proporcionado seguir na área que mais gostava. Por sempre me terem apoiado e incentivado a concretizar os meus objetivos, sem nunca duvidarem de mim. Ao Paulo, por aguentar o stress, a ansiedade, por mesmo sem entender muito bem, sempre esteve disponível em me ouvir e apoiar. Com a certeza de que sem vocês não teria sido de todo possível, um muito obrigado!

Por último, me gustaría agradecer a nuestros colaboradores, la Prof. Dra. María Domínguez y el Dr. Andrés Garelli, la contribución que han hecho y siguen haciendo, con el desarrollo de esta historia. Sin ellos no hubiera sido posible, gracias!

Abstract

Relaxin is a hormone structurally similar to insulin, firstly described in 1926, as a substance with a significant influence on the reproductive system. While insulin activates a tyrosine kinase receptor and stimulates signaling pathway that includes phosphoinositide 3-kinase (PI3K) and serine/threonine kinase (AKT), relaxins bind to Leucine-rich repeat-containing G-protein-coupled receptors (LGRs) of the C1 subtype. In Drosophila, two *LGRs* of subtype C1 exhibit clear structural homology with relaxin receptors, described in mammals. Nevertheless, the ligands and the biological functions of these receptors remain unknown not only in *Drosophila,* but also in all invertebrates. Here our objective was to generate genetic tools to study the biological function of these two Type C1 LGRs in *Drosophila*. With this aim we generated mutant lines for both receptors through classical transposon remobilization techniques. The mutants obtained were characterized as strong loss-of-function deletions, yet no clear phenotype was observed for any of the receptors as regards viability and reproduction. We then tested the hypothesis that either Type C1 LGR could be part of the developmental delay pathway triggered by the *Drosophila* insulin-like peptide. This peptide is produced and secreted from abnormally growing imaginal discs and delays the onset of metamorphosis by inhibiting the biosynthesis of the major insect molting hormone ecdysone. Strikingly, we found that mutations in either Type C1 LGRs could suppress the insulin-like peptide-dependent delay in the onset of metamorphosis to different extents. These results provide the first glimpse into a biological function for invertebrate Type C1 LGR receptors and place them downstream or in parallel to *Drosophila* insulin-like peptide in this developmental timing control pathway. The resemblance between human and *Drosophila* core physiological and developmental pathways reinforce that the fly can be a powerful model system to study genes and pathways that are relevant for human development and disease.

Keywords: *Drosophila melanogaster,* Leucine-rich repeat-containing G-protein-coupled receptors, Drosophila insulin-like peptide, development.

Resumo

A relaxina é uma hormona, estruturalmente similar à insulina, descrita pela primeira vez em 1926, como uma substância com grande influência no sistema reprodutivo dos vertebrados. Apesar de semelhantes estruturalmente, a relaxina e a insulina atuam de forma distinta, ativando diferentes recetores. A insulina ativa um recetor tyrosine kinase e estimula a via de sinalização que inclui phosphoinositide 3-kinase (PI3K) e serine/threonine kinase (AKT), enquanto que as relaxinas ligam-se a *Leucine-rich repeat-containing G-protein-coupled receptors* (LGRs) do subtipo C1.

Em *Drosophila*, dois *LGRs* do subtipo C1, exibem uma clara homologia estrutural com os recetores de relaxina, descritos em mamíferos. Apesar disso, os ligandos e as funções biológicas destes recetores, são ainda desconhecidos em *Drosophila*. Uma vez que estes recetores são conservados ao longo da evolução, torna-se extremamente importante o seu estudo, com o intuito de perceber os reais efeitos que as interações ligando-recetor exercem no desenvolvimento dos organismos. Para tal, foram geradas linhas mutantes para ambos os recetores, através de um protocolo de excisão de inserções existentes nestes locus. Os mutantes obtidos foram caracterizados como deleções que, de acordo com o local e a extensão das mesmas, geraram proteínas truncadas com perda de função. Nenhum fenótipo relacionado com perda de fertilidade ou viabilidade foi observado, para nenhum dos mutantes.

Já é conhecido que um dano físico ou químico ocorrido nos discos imaginais de *Drosophila melanogaster* provoca um atraso na metamorfose, permitindo que ocorra um restabelecimento das condições normais para o bom desenvolvimento do animal. Estudos efetuados descobriram que existe um membro do grupo *insulin/IGF-I/relaxin family of peptides*, o *Drosophila insulin-like peptide 8*, responsável pela coordenação do atraso verificado no desenvolvimento em *Drosophila*, através da inibição da biossíntese da hormona *ecdysone*.

Com o intuito de perceber se algum dos dois recetores *LGRs* do subtipo C1 poderá atuar na via do *Drosophila insulin-like peptide 8*, o teste do tempo de desenvolvimento foi realizado. Este teste permitiu verificar que ambos os recetores suprimem o atraso na metamorfose produzido por esse péptido, embora com

diferentes intensidades. Estes achados demonstram, pela primeira vez, uma função biológica dos recetores *LGRs* do subtipo C1 de um invertebrado e os posicionam numa via específica de sinalização, a via do controle do tempo de desenvolvimento.

A semelhança que existe entre os processos biológicos importantes em distintos seres vivos, como é o caso do ser humano e da *Drosophila,* reforça que a mosca-da-fruta pode ser um poderoso modelo animal para o estudo de moléculas e vias conservadas com relevância, tanto para o desenvolvimento normal humano, quanto para as suas doenças.

Palavras – chave: *Drosophila melanogaster,* Leucine-rich repeat-containing G-protein-coupled receptors, *Drosophila* insulin-like peptide, desenvolvimento.

Index

List of figures

amplification products of the precise and imprecise excision *lgr3[exp2]* and *lgr3[ex1],* respectively, which shows that the deletion occurred downstream (to the right) of the insertion site. (C) Electrophoretic band of PCR amplification product of *lgr3[ex1]*, from which we later obtained the gDNA sequence and identified a deletion of approximately 1.6 kb. (D) Scheme of the *lgr3[ex1]* deletion. The deletion removes most of the intron and completely removes exon seven and partially exon eight. The negative control used in the different experiments was dH_2O, showing no PCR amplifications, as expected. **69**

Figure 28. UCSC genome browser screenshot of the *D. melanogaster lgr3* locus with BLAT results (Your Sequence from Blat Search) obtained by sequencing the 1.6-kb PCR fragment obtained in Figure 29-C. This shows the breakpoints that characterize the 3827-bp deletion in *lgr3[ex1]* mutant. **69**

Figure 29. *RT-PCR amplification pattern of lgr3 mutations and controls using mRNA isolated from males.* (A) Electrophoretic pattern of PCR amplification products of *lgr3[exp2]* and *lgr3[ex1]*, with the set of primers *lgr3_around* (set of primers in blue). The *lgr3[ex1]* mutant gave a smear and a major product at around 200 bp, consistent with a non-spliced transcript reading directly through the deletion breakpoint. (B) RT negative control, showing no amplification products, ensuring that any gDNA contamination, if present, was under detection levels. (C) The housekeeping gene *rp49* was used as a positive control for the RT-PCR. The negative control used was dH_2O. (D) Schematic representation of the *lgr3* cDNA in the *lgr3[ex1]* mutant and in the *lgr3[exp2]* control line. Blue, normal splicing and exon reading frames. Red, aberrant splicing and frame shifted exons. We were unable to sequence a specific product from the *lgr3[ex1]* transcript. **70**

Figure 30. Transcript levels of the mutant *lgr3[ex1]* mutant and controls. An aberrant transcript can be detected in *lgr3[ex1]* which is present at higher levels than that of the other stocks. This could be due to the lack of the large intron deleted in this mutation. The correctly spliced *lgr3* transcript found in *lgr3[MB06848],*

scheme, the *lgr4* locus transcribes from right to the left. Exons are numbered. Colored arrows are show the approximate location of the primers used. (B) Electrophoretic pattern of PCR amplification products of the *lgr4[exp1]* precise excision and the *lgr4[ex1]* and *lgr4[ex3]* deletions. This shows that the *lgr4[ex1]* deletion occurred upstream of the insertion site while the *lgr4[ex3]* deletion occurred downstream. This also shows that the *lgr4[exp1]* is indeed a precise excision giving the expected product PCR size with the green primer pair (note that no 735-bp product is produced in the *MB03440* line because of the presence of the several-kb *MB03440* insertion in between the primer sites). The positive control, used to check the deletions, was the original line used in the excision protocol. As expected, the negative control used, dH_2O, showed no PCR amplification. (C) Gene scheme depicting the deletion in *lgr4[ex1],* showing that exon 2 was partially deleted. (D) Gene scheme depicting the deletion in *lgr4[ex3],* showing that the deletion affects introns two, three and four and the exons three and four. **75**

Figure 35. UCSC genome browser screenshot of part of the *D. melanogaster lgr4* locus containing BLAT results (Your Sequence from Blat Search) obtained by sequencing the PCR product with a set of primers upstream and downstream of the *lgr4[MB03440]* position (blue arrowhead). Instead of the predicted 255-bp product, we obtained a product of 96 bp, confirming the *ex1* mutation. Notice that the *lgr4* locus transcribes from the right to the left in this scheme. Hence, the exon to the right is exon 2, and the exon to the left is exon 3. Some of the sequences adjacent to the deletion missalign in exon 2 rather than aligning in the intron. Actually the deletion starts immediately to the right of the *MB03440* element. **76**

Figure 36. Protein sequence of WT Lgr4 (Reference NP_001096966.2) and the predicted Lgr4[ex1] protein. According to the sequence obtained, the deletion causes the formation of a truncated 137-aa protein containing the very N-terminus of Lgr4 up to the LDLa domain. In red are the aa that form due to readthrough from exon 2 into the downstream intron. This sequence was confirmed by sequencing the RT-PCR product of this site. The major domains of the Lgr4 protein are the signal peptide, LDLa domain, LRR N-terminal domain (LRRNT), LRRs, the hinge region

(Hinge) and the 7TM domains. This cartoon is not to scale. It depicts the approximate region of the Lgr4[ex1] protein truncation. There could be other protein isoforms if splicing still occurs downstream of the breakpoint, but these are also predicted to encode premature stop codons and hence truncate the protein. **76**

Figure 37. RT-PCR amplification pattern of *lgr4* mutations and controls using mRNA isolated from males. (A) Electrophoretic pattern of RT-PCR amplification products in the depicted genotypes with the set of primers for *lgr4 gene* (set of primers forward in red and reverse in purple). (B) RT negative control, showing no amplification, as expected, ensuring that no gDNA contamination occurred. (C) Positive control, the housekeeping gene *RP49* expression. The negative control used was dH_2O, showing no PCR amplification, as expected. (D) Scheme of the *lgr4* transcripts in the mutants and in the control *exp1* line. Notice that in this scheme the *lgr4* locus is transcribed from the right to the left. **78**

Figure 38. qRT-PCR analyses of transcript levels of the *lgr4[ex1]* deletion mutant in comparison with the control lines, *lgr4[exp1] and w[1118],* and the original insertion line *lgr4[MB03440]*. Primer pairs in exonic sequences flanking the *ex1* deletion were used and normalized to the housekeeping gene *rp49*. A transcript was detected in *lgr4[ex1],* which was nevertheless aberrant and had a distinct dissociation curve (B). One can also observe that the *lgr4[MB03440]* insertion induces a small reduction in *lgr4* levels, as previously detected (see Figure 34), which is restored to normal in *lgr4[exp1]*. Only males cDNA were analyzed. Shown are the averages of three repeats, except for *lgr4[exp1]*. Error bars are standard deviations of the means. (B) The dissociation curves of selected qRT-PCR products shown in (A) for the three depicted genotypes. These results confirm that *lgr4[ex1]* produced an aberrant transcript, different from the *w[1118]* and *lgr4[exp1]* controls. **79**

Figure 39. *lgr4[ex3]* cDNA sequence. UCSC genome browser screenshot of part of the *D. melanogaster lgr4* locus containing BLAT results (Your Sequence from Blat Search) obtained by sequencing the RT-PCR product obtained in Figure 39-A for *lgr4[ex3]* with a set of primers in exon2 and in exon7. The *lgr4[MB03440]* position

is depicted with a blue arrowhead. Notice that the *lgr4* locus transcribes from the right to the left in this scheme. The *lgr4[ex3]* transcript splices perfectly from exon two to exon seven, consistent with the prediction of the deletion of all introns two, three and four, and all exons three and four. Exons 5 and 6 were also probably deleted. **80**

Figure 40. Protein sequence of WT Lgr4 (Reference NP_001096966.2) and the predicted Lgr4[ex3] protein. According to the sequence obtained, the deletion causes the formation of a truncated ~163-aa protein (153 aa of Lgr4 sequence and 10 aa of framshifted sequence from exon 7) containing the very N-terminus of Lgr4 up to the LDLa domain. In red are the aa that form due to out-of-frame splicing from exon 2 into the exon 7. The major domains of the Lgr4 protein are the signal peptide, LDLa domain, LRR N-terminal domain (LRRNT), LRRs, the hinge region (Hinge) and the 7TM domains. This cartoon is not to scale. It depicts the approximate region of the Lgr4[ex3] protein truncation. There could be other protein isoforms if splicing still occurs downstream of exon7, but most of these are also predicted to encode premature stop codons and hence truncate the protein. **80**

Figure 41. *lgr3* mutation rescues the *dilp8*-dependent delay in the onset of pupariation. (A) Pupariation timing curves for different genotypes. Shown are averages +/- SEMs of 8-16 experiments. (B) Bar graph showing the median pupariation time of the depicted genotypes. ANOVA: $p<0.001$. Groups sharing the same letters are not statistically significantly different at alpha=5% according to the Tukey's HSD post-hoc test. Shown are averages +/- SEMs of 8-16 experiments. Bar colors depict the dosage of *lgr3* (blue = two wild-type alleles, red = two loss-of-function alleles; hashed bars = heterozygotes). Green arrow depicts full suppression. Yellow arrow depicts partial suppression. **82**

Figure 42. Preliminary evidence that *lgr3[ex1]* is a null mutation and that *dilp8* activity is sensitive to *lgr3* dosage. (A) Pupariation timing curves for different genotypes. Shown are averages +/- SEMs of 3 experiments except for *tub-dilp8/+;lgr3[ex1]/Df(3)BSC321* and *tub-dilp8/+;lgr3[ex1] lgr3[ex1]*, which had two

List of tables

List of abbreviations

7TM - seven transmembrane domain

20E - 20-hydroxyecdysone

aa - aminoacid

AKT - serine/threonine-specific protein kinase

AMP - adenosine monophosphate

AP2 - clathrin adaptor

BLAST- basic local alignment search tool

bp - base pair(s)

Ca^{2+} - calcium ion

cDNA - complementary DNA

CRR - cysteine-rich region

Cys - cysteines

DAG - diacylglycerol

Df - deficiency chromossome

dH2O - distilled water

Dilp - *D. melanogaster* insulin-like peptides

DNA - deoxyribonucleic acid

EDTA - ethylenediaminetatracetate

FSHR - vertebrate gonadotropin receptor

Gα/Gβ/Gγ - heterotrimeric G proteins subunits

gDNA - genomic deoxyribonucleic acid

GDP - guanosine diphosphate

GFP - green fluorescent protein

GPCR – G-protein coupled receptors

GRK - G-protein coupled receptors kinases

GTP - guanosine triphosphate

H2O - water

HCl - hydrochloric acid

HR - hinge region

IGF - insulin-like growth factor

IGFR - insulin-like growth factor receptor

IIS - insulin/IGF Signaling

Ilp - insulin-like peptide

InR - insulin receptor

Insl - insulin-like peptide

JH - juvenile hormone

Kb - kilobase(s)

LDLa - low density lipoprotein receptor domain class A

LGR - leucine-rich repeat containing G protein-coupled receptor

LH/CGR - luteinizing hormone receptor

LRR - leucine-rich repeat

MAPK - mitogen-activated protein kinase

mRNA - messenger RNA

MYA - million years ago

NaCl - sodium chloride

NCBI - National Center for Biotechnology Information

O.N. - overnight

PCR - polymerase chain reaction

PI3K - phosphoinositide 3-kinase

PTTH - prothoracicotropic hormone

qPCR - quantitative polymerase chain reaction

RLN - relaxin

RNA - ribonucleic acid

RNAi - RNA interference

rpm - rotations per minute

RT - room temperature

RTK - tyrosine kinase receptor

RT-PCR - reverse transcription polymerase chain reaction

qRT-PCR - quantitative reverse transcription polymerase chain reaction

rp49 - housekeeping gene, also knowed as *rpl32*

RXFP - human relaxin receptor

SD - standard deviation

SNP - single nucleotide polymorphism

TAE - tris-acetate-ethylenediaminetatracetate

TE - transposable element

Tm - melting temperature

tRNA - transfer ribonucleic acid

TSHR - thyroid hormone receptor

Tub - tubulin promoter

UAS - upstream activating sequence

Chapter 1. Introduction

1.1G-protein Coupled receptors

G protein-coupled receptors (GPCRs) are an ancient and large family that share structural characteristics of having at least an extracellular Amino (N)-terminus, an intracellular Carboxy (C)-terminus, the characteristic seven transmembrane (7TM)-spanning helical domains, and three interhelical loops on each side of the membrane (Figure 1). GPCRs transduce a diverse array of extracellular signals, such as light, calcium, neurotransmitters, odorants, peptides, proteins, among others. Hence, GPCRs are involved in most physiological processes, from smell to vision and from taste to neurological, cardiovascular, endocrine and reproductive responses. Because of this, GPCRs are an extremely important therapeutic and pharmacological target, as well as a subject of considerable research interest (Broeck, 1996; Bargmann, 1998; Broeck, 2001; Tyndall *et al.*, 2005; Overington *et al.*, 2006; Bridges *et al.*, 2008).

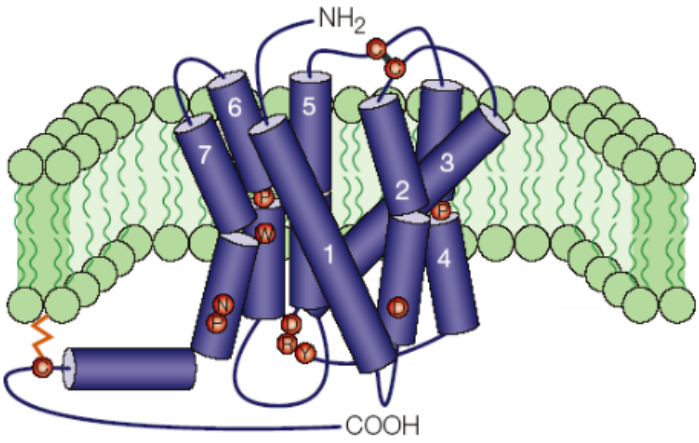


Figure 1. Characteristic structure of a GPCR in the cellular membrane. The seven transmembrane (7TM) domains are numbered. The Amino-terminus of the protein is extracellular, while the Carboxy-terminus is intracellular. The structure depicted is that of a typical rhodopsin-like GPCR. Image from Ellis (2004).

In a classical perspective, GPCR signaling begins when an external signaling molecule, such as an agonist ligand molecule, binds to the extracellular loops of the GPCR 7TM domains and induces a conformational change in the GPCR that stimulates its interaction with the intracellular guanine nucleotide–binding regulatory proteins (G proteins) (Sealfon, 2005; Figure 2).

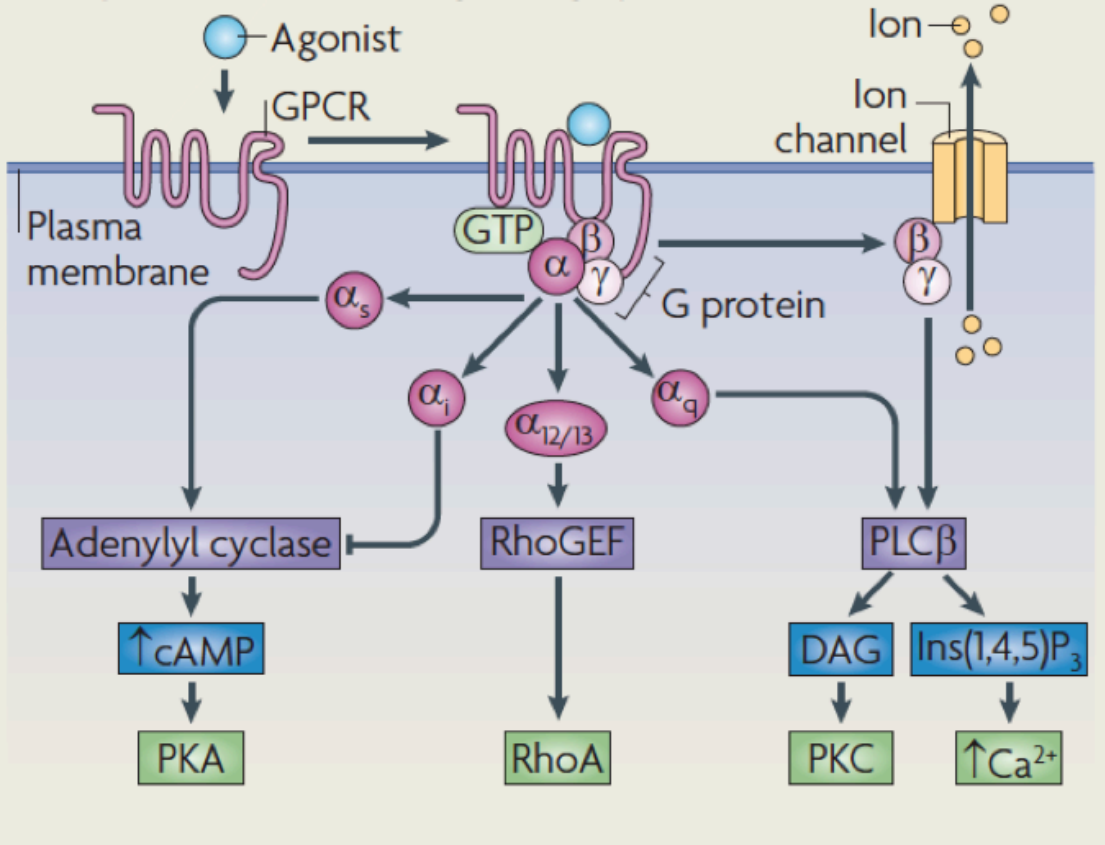


Figure 2. GPCR signaling in a classical perspective, from Ritter and Hall (2009).

G proteins are a subfamily of GTPase proteins that can bind the nucleotides guanosine triphosphate (GTP) and guanosine diphosphate (GDP) (Purves *et al.,* 2001; Oldham & Hamm *et al.*, 2008; Ritter & Hall, 2009). GTPases bind and hydrolyze GTP through their highly conserved *G* domain. Typically, when a G protein is bound to GTP it is "active", and when it is bound to GDP, it is "inactive". G proteins can be either monomeric (e.g., the Ras superfamily of small GTPases) or heterotrimeric (Purves *et al.,* 2001). Heterotrimeric G proteins consist of alpha (Gα), beta (Gβ) and gamma (Gγ) subunits, with the *G* domain localized in the Gα subunit, the largest one (Figure 2). The Gα and Gβ subunits associate with the cellular membrane by lipid anchors.

Upon stimulation, the GPCR catalyzes the exchange of GDP to GTP in the Gα subunit, leading to dissociation of the Gα from the Gβγ subunit and the GPCR (Figure 2). Both activated Gα (Gα-GTP) and the dissociated Gβγ can activate signaling cascades. The specific cellular response will depend on the specific subtypes of Gα [*e.g.*, $G\alpha_S$ (stimulatory), $G\alpha_i$ (inhibitory), $G\alpha_{12/13}$ or $G\alpha_q$], just to cite

a few subtypes (Oldham & Hamm *et al.*, 2008; Ritter & Hall, 2009) that couples to the specific GPCR in question. The different Gα subtypes and the dissociated Gβγ can bind to different effectors such as adenylyl cyclase, RhoGEF, phospholipase Cβ, and exercise different downstream regulation by producing second messengers (cyclic AMP, diacylglycerol – DAG, inositol-1,4,5-triphosphate) that will themselves regulate further downstream effectors like protein kinase A and C, ion channel and phospholipase Cβ (Figure 2) (Oldham & Hamm *et al.*, 2008; Ritter & Hall, 2009).

If the stimulating ligand is still present, the activated GPCR will bind to and activate the next heterotrimeric G protein, and signaling will continue as long as ligand and inactive G proteins are available. In the absence of stimuli, the GPCR signal terminates when the Gα subunit eventually hydrolyzes the attached GTP to GDP (Oldham & Hamm *et al.*, 2008; Ritter & Hall, 2009).

1.1.1 Role of GPCR in diseases

Many diseases are directly linked with mutations in GPCR receptors. Known diseases include, fertility disorders, congenital night blindness, familial gestational hyperthyroidism, nephrogenic diabetes insipidus, and carcinomas (Schöneberg *et al.,* 2005; Thompson *et al*., 2008). Some of the mechanisms by which the mutations lead to disease comprise changes either in the GPCR gene expression levels (*e.g.,* transcription regulatory mutations), in the abundance of the receptor in the plasma membrane (*e.g.,* receptor recycling), in receptor signaling (*e.g.,* G protein coupling), in receptor desensitization, or alteration of ligand binding by loss of ligand specificity or in agonist-independent receptor signaling (Sealfon, 2005; Schöneberg *et al.,* 2005; Thompson *et al.*, 2008).

Like GPCRs, mutations in G-proteins that lead to deficient or excessive G protein signaling can also lead to different types of diseases, such as cancer, cardiovascular, neurological, or endocrine disorders. Additionally, many drugs used to treat asthma, hypertention or depression are targeted to disrupt G protein signaling (Ibegbu *et al.,* 2012). Loss-of-function mutations in different subtypes of Gα subunits can lead to different disorders, such as Pseudohypoparathyroidism (mutations $G\alpha_s$) and Night Blindness (mutations in $G\alpha_t$). On the other hand, gain-of-function mutations in Gα subunits, that reduce their capacity of hydrolyzing GTP and hence

terminating GPCR signaling, have been linked to adenoma formation. Gβ subunits mutations have been linked to hypertension (Ibegbu *et al.,* 2012).

1.1.2 GPCR evolution

GPCRs are present in the majority of eukaryotic organisms, testifying for their extraordinary evolutionary success (Fredriksson & Schioth, 2005; Perez, 2005; Figure 3). GPCRs can be generally classified into different subfamilies the Adhesion, Secretin, Frizzled, Glutamate, and Rhodopsin families (Fredriksson, 2003). These families likely arose and expanded in bilateria about 750-430 million years ago (MYA), before the nematodes split from the chordate lineage (Fredriksson & Schioth, 2005; Perez, 2005).

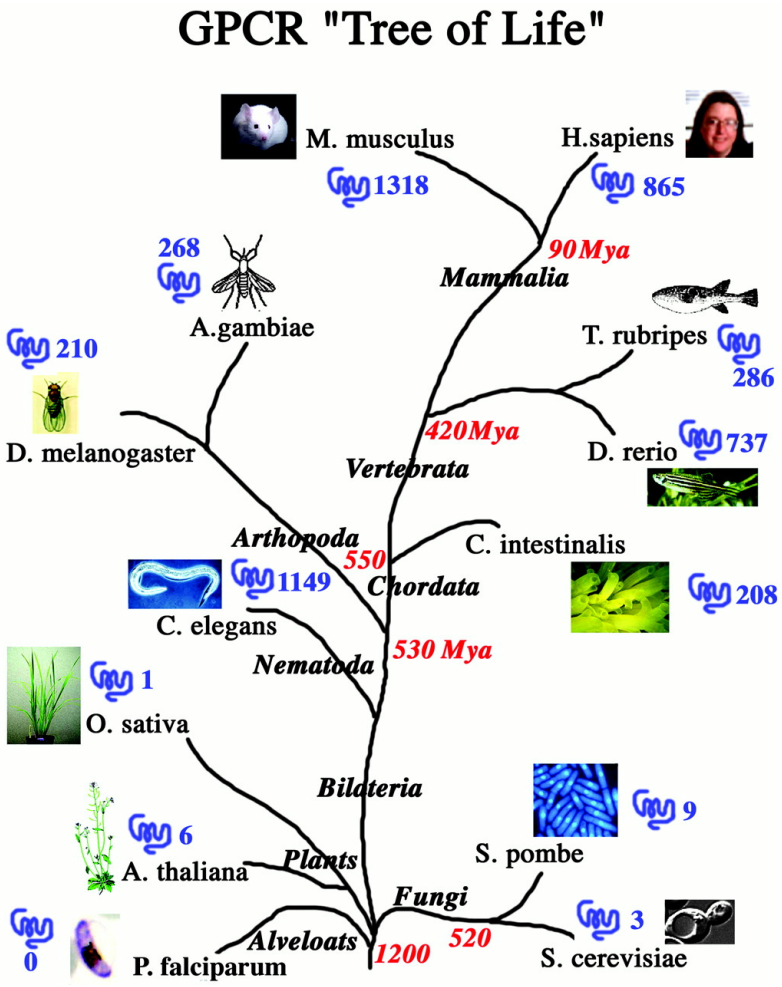


Figure 3. GPCR phylogenetic "Tree of life", suggested by Fredriksson and Schioth (2005). The red numbers represent the time in million of years (MYA) since the split. GPCRs and their number in the different classes are represented in blue. Image from Perez (2005).

The most well studied and understood GPCR family is probably the Rhodopsin GPCR family (also known as Rhodopsin-like receptor family), which were also the first to be discovered (Brody & Cravchik, 2000). This family includes Opsin receptors, receptors for biogenic Amines, Purine receptors, Peptide receptors and Orphan receptors that can be activated by diverse ligands or stimuli, such as neurotransmitters, neuropeptides, peptide hormones, light, nucleotides, prostaglandins, leukotrienes, chemotactic peptides and chemokines. These receptors vary considerably in structure, but are characterized by their conserved 7TM domain structure (Figure 1). Besides the well studied function as photoreceptors, other members of the Rhodopsin-like GPCR family, such as the Leucine-rich repeat-containing G-protein coupled receptors (LGRs) (see below) also play an important role in development (Brody & Cravchik, 2000).

The LGR subgroup of Rhodopsin-like GCPRs are structurally characterized by a large N-terminal extracellular domain flanked by a cysteine-rich region (Figure 4), in addition to the typical 7TM of the Rhodopsin-like receptor family. LGRs can be classified in three subtypes: according to the numbers of Leucine-rich repeat motifs (LRRs) in their ectodomains and the sequence of the hinge domain in between the LRRs and the 7TMs. Type A and C LGRs contain between 7-9 LRRs, while Type B LGRs have 13-18 LRRs (Figure 4; Hiel *et al*., 2011).

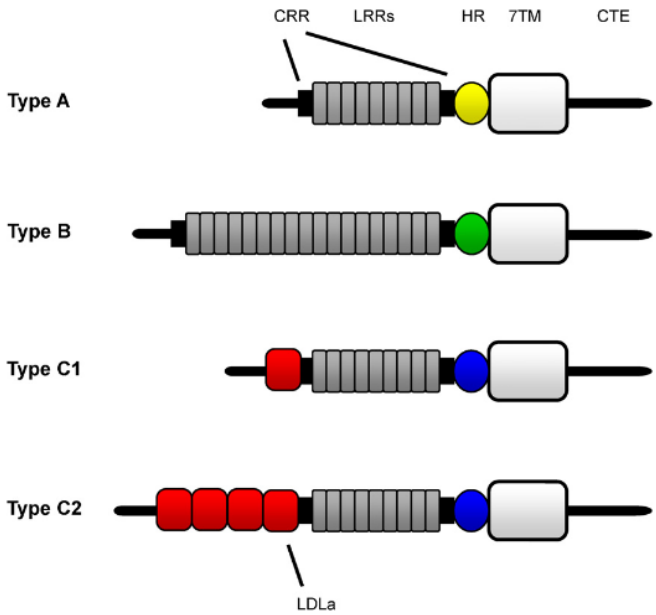


Figure 4. Representation of the major LGR types. Image from Van Hiel *et al.,* (2011). 7TM – seven transmembrane domain; CRR – cysteine-rich region; CTE – C terminal endodomain; HR – hinge region; LDLa – low density lipoprotein receptor domain class A; LRR – leucine-rich repeat.

Type A LGRs are further characterized by having the longest hinge region, with a conserved sequence containing 6 Cysteines that are expected to form disulfide bonds and stabilize the receptor structure. Examples of Type A LGRs are the Follicle Stimulating Hormone Receptor FSHR receptor (vertebrate gonadotropin receptors), the LH/CGR receptor (luteinizing hormone receptor) and the thyroid hormone receptor (TSHR). Apart from having more LRRs than Type A, Type B LGRs have a shorter hinge region with only 4 Cysteines. An example of a Type B LGR is the Drosophila Bursicon receptor, rickets, responsible for the critical wing expansion and cuticle hardening after eclosion of the fly (Baker & Truman, 2002).

Type C LGRs have the shortest hinge region among the LGRs and have the consensus sequence LxxLxHIxFDRFxYCxYAPHV at its beginning and CxPxxDGISSxEDLLSNxVLRV at its end (Hiel *et al*., 2011). They are further characterized by having at least one Low-density lipoprotein domain class A (LDLa) domain located N-terminally to the LRRs (Figure 4). Type C LGRs are divided in Type C1 and Type C2, depending on the number of Cysteines in the hinge region and the number of LDLa repeats in the extracellular domain. Type C2 have many LDLa repeats and four Cysteines, while type C1 has only one LDLa domain and two Cysteines in the hinge region. Examples of Type C1 LGRs are the human relaxin receptors RXFP1 and RXFP2, also known as LGR7 and LGR8, respectively. In the fruit fly genome, two LGRs type C1 genes have already been identified, Lgr3 and Lgr4 (as well known as CG31096 and CG34411 respectively) (Hiel *et al.,* 2011). The biological functions of these two putative *Drosophila* Relaxin-like Type C1 LGRs are unknown and are the focus of this study.

1.2 Relaxin like-peptides

Relaxin was first discovered in 1926, while Dr. Frederick Lee Hisaw observed the relaxation of the interpubic ligament in virgin guinea pig females, following the injection of serum from pregnant guinea pigs or rabbits (Hisaw, 1926). In 1930, the Hisaw group named the active substance “relaxin”, after purifying it from pig corpora lutea and postulating it to be a peptide hormone. The determination of the relaxin peptide structure and cloning of the rat and pig relaxin cDNAs, revealed a close structural similarity to insulin and their appurtenance to the larger insulin-like peptide (Ilp) superfamily (Hudson *et al.,* 1981; Haley *et al*., 1982).

All Ilps are structurally related and are thought to originate from a common Insulin/IGF-like ancestral peptide (Blundell *et al*., 1980; Wilkinson & Bathgate, 2006; Bathgate *et al*., 2013). Ilps are now subdivided into insulins, insulin-like growth factors (IGFs), and relaxins on the basis of gene and protein structure, processing, and receptor binding preferences (Blundell *et al.,* 1980; Wilkinson & Bathgate, 2006; Figure 5). Insulin binds to a receptor tyrosine kinase (RTK), the insulin receptor (InR), and stimulates a signaling pathway that includes phosphoinositide 3-kinase (PI3K) and serine/threonine kinase (AKT). IGFs preferentially bind related RTKs, named IGF receptors (IGFRs), which share an evolutionary history with the InR, and activate the MAPK pathway. In contrast, relaxins, as will be further described below, bind to GPCRs and regulate diverse downstream signaling pathways (Hsu *et al*., 2002; Hiel *et al.,* 2011; Bathgate *et al.,* 2013). Functionally, while the insulin/IGF pathway regulates growth, metabolism and ageing (Blundell *et al.,* 1980; Wilkinson & Bathgate, 2007; Gronke *et al.,* 2010), members of the vertebrate relaxin family, have antifibrotic, neuropeptide, vasodilator and cardiac stimulatory functions, affecting many processes, including - but not limited to -, female and male reproduction, such as menstruation, sperm motility, and pregnancy (reviewed in Bathgate *et al*., 2013). Therefore, it is important to stress, that a peptide's appurtenance to the Ilp superfamily per se by no means implies that it will act via insulin/IGF Signaling (IIS).

Humans have 10 Ilps (1 insulin, 2 IGFs and 7 relaxins). The single insulin gene and all the relaxin genes are thought to encode preprohormones (Figure 6), while both IGFs are encoded as prohormones. All have an N-terminal signal peptide which is processed out of the mature form followed by two chains, B and A, which have two and four conserved Cysteines, respectively, and form characteristic inter and intra-chain disulfide bonds in the mature peptides. In insulin and relaxins, the sequence in between chains B and A, called "C peptide", is further proteolitically processed out of the prohormone and secreted together with the mature peptide in an equimolar ratio. In this way, both mature insulin and relaxin is formed by discontinuous chains B and A, that are nevertheless held tightly together by the Cys-Cys disulfide bonds (Figure 6).

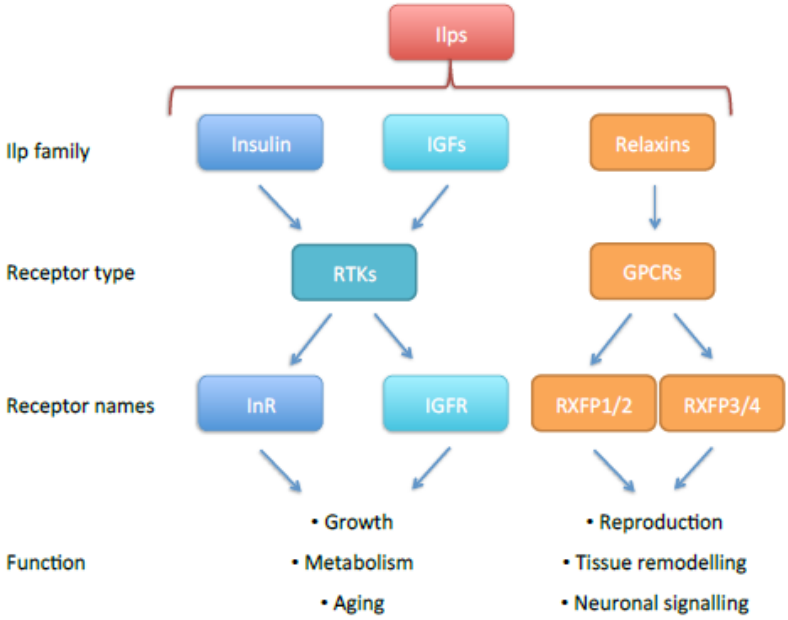


Figure 5. General scheme of the subdivision of Ilps into insulins, insulin-like growth factors (IGFs), and relaxins based on receptor binding preference and function.

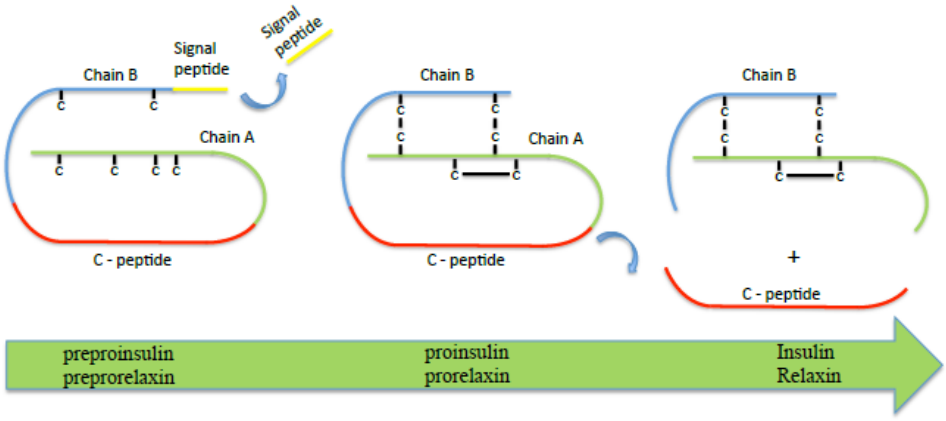


Figure 6. Preprohormone processing steps for Ilps. Insulin and relaxins are encoded as preprohormones and are processed twice as described in the text. IGFs do not undergo C peptide processing. "C" stands for the amino acid Cysteine. The three disulfide bonds made by the six stereotype Cysteines of the Ilp superfamily are shown.

The seven members of the human relaxin family peptides are: Relaxin-1 (RLN1), Relaxin-2 (RLN2), Relaxin-3 (RLN3) and INSL3 (*aka* Leydig cell insulin-like peptide and relaxin-like factor), INSL4 (*aka* placenta insulin-like peptide, INSL5 and INSL6. The two first relaxin genes to be described in humans were RLN1 and RLN2 (Hudson *et al.,* 1983; Hudson *et al.,* 1984). RLN1 and RLN2 arose from a duplication that occurred in the last common ancestor of catarrhini primates, but only

apes retained the RLN1 duplication (Arroyo *et al.,* 2014). Hence, despite the historical name, RLN2, aka as H2 relaxin in humans, is the ancestral and major form of circulating relaxin found in the great apes and is the functional equivalent to the first relaxin identified by F.L. Hisaw, which is produced by the corpus luteum and by placenta during pregnancy (Hisaw, 1926).

The evolution of vertebrate relaxin-like peptides has been focus of several studies (Maere *et al*., 2005; Makino *et al.,* 2009), but there is limited consensus on how and when these peptides diverged from ancenstral Insulin/IGF-like peptides. Probably the most comprehensive analyses of the evolution of the relaxin-like peptides was carried out by Sergey Yegorov and Sara Good in 2012, who applied paleogenomic models to elucidate the role of whole genome duplications on the story of these peptides and their GPCR receptors. The most likely scenario is that a single relaxin-like peptide, probably most similar to the neuropeptide RLN3, existed in the ancestral vertebrate (Bathgate & Wilkinson , 2007; Yegorov & Good, 2012).

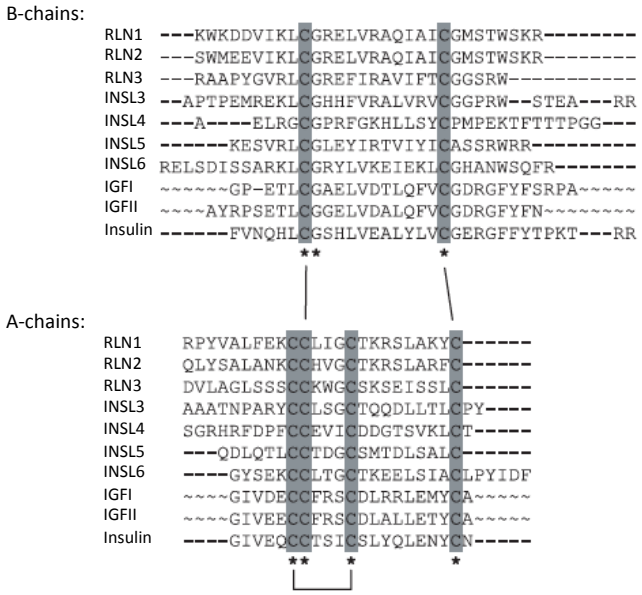


Figure 7. Comparison of amino-acid structures of the 10 human Ilps, adapted and modified from Bathgate *et al.* (2013). Note the so called "relaxin signature" [RxxxRxxI/V] in some of the B chain of the mature relaxin-like peptides.

Some, but not all, human relaxin-like peptides, most notably RLN1, RLN2 and RLN3, share the so called "relaxin signature" [RxxxRxxI/V] in the B chain of the mature peptide (Figure 7). These amino acids, as will be described further below,

are all important for their binding to the relaxin receptor. Despite this similarity between RLN1, RLN2 and RLN3, the human relaxin-like peptides actually have a different phylogenetic history. While RLN1 and RLN2 are intimately related, they seem more related to INLS3, 4 and 6 than to relaxin-3 and INLS5, which branch earlier in the relaxin-like peptide tree (Figure 8; Wilkinson *et al.,* 2005).

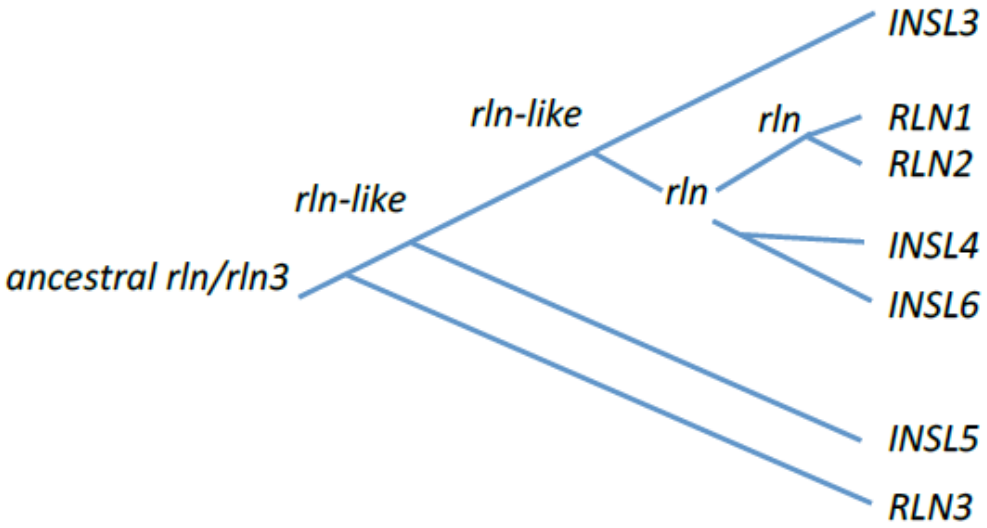


Figure 8. Evolutionary relationship between rln/insl genes in vertebrates based on Yegorov *et al.,* (2014). While the endpoints are the seven human relaxin-like genes, not all vertebrates have these all of these genes, yet they have rln-like genes in their genome. INSL4 and INSL6 is found in placental mammals and RLN1 and RLN2 in catarrhini primates. The last common ancestor to all vertebrates is thought to have had one rln/rln3-like gene and one copy of an RXFP1/2-like receptor and one copy of an RXFP3/4 receptor (Bathgate & Wilkinson , 2007; Yegorov & Good, 2012; Yegorov *et al.*, 2014).

While different relaxin-like peptides have known specialized functions, the functions of other peptides, such as RLN1, INSL5, and INSL6 remain unknown, albeit some evidences point to neuroreproductive roles, at least for INSL5 and INSL6 (Conklin *et al.*, 1999; Lok *et al*., 2000; Kong, 2010). The major circulating relaxin in human, RLN2, has many known functions both related and unrelated to reproduction. For instance, RLN2 has a strong antifibrotic role which is associated with many of its reproductive functions, such as that of promoting cervical softening, or ripening, in pregnant women. Amongst other functions, it is also known to promote nipple and mammary gland development in other mammals. In males, it is produced in the prostate and promotes sperm motility (Unemori *et al.,* 1990; Bell *et al.,* 1993; Unemori *et al.,* 1993; Sherwood, 1994; Unemori *et al., 1*996; Eriksen *et al.,* 2001; Garber *et al.,* 2001; Williams *et al.,* 2001). In contrast to RLN2, RLN3 seems to act as a neuropeptide with a role in the central nervous system (Bathgate &

Wilkinson, 2007). INSL3 has a role in male reproductive system being involved in the mediation of testicular descent (Nef & Parada, 1999; Zimmermann *et al.,* 1999). INSL4 plays a role during pregnancy specifically in placenta (Chassin *et al.,* 1995; Koman *et al*., 1996).

While relaxin was discovered in the late 20s, as described above, the link between relaxins and GPCRs was only definitively established many decades later, in 2002 (Hsu *et al*., 2002; Ivell, 2002). Due to the structural similarity between relaxins and insulin/IGF-like peptides, the relaxins had been expected to also act via RTKs (Ivell, 2002; Bathgate & Wilkinson, 2007). The discovery in the late 90s, that mice lacking the relaxin family peptide INSL3 showed a similar abnormal testis descent phenotype, as mice with a disrupted *GREAT* gene, which encodes a GPCR, provided the first strong hint towards GPCRs as putative relaxin receptors (Nef & Parada, 1999; Zimmermann *et al.,* 1999; Overbeek *et al.,* 2001). The fact that relaxin stimulated cAMP in some tissues (Sherwood, 1994), but not all (Ivell, 2002), also added strength to this link. Hence, the discovery, reported in 2002, that the relaxin peptide (RLN2) was able to activate the RXFP1 and RXFP2 GPCRs (also known as LGR7 and LGR8, respectively) provided the definitive evidence that relaxins were not RTK ligands. The ligand and receptor relationship between INSL3 and *GREAT* was later confirmed (Kumagai *et al.,* 2002; Bogatcheva *et al*., 2003), and as the mouse *GREAT* gene encodes the homologue of RXFP2 that is expressed in a similar distribution as INSL3 rather than RLN2, it is INSL3 that is regarded as the endogenous ligand for RXFP2 despite the fact that it can also be activated by RLN2.

In 2003, another two orphan GPCRs, RXFP3 and RXFP4, also known as GPCR135 and GPCR142, respectively, were identified as receptors for RLN3 and INSL5, respectively (Liu *et al.,* 2003a; Liu *et al.,* 2003b). These two GPCRs have a distinct structure from the RXFP1 and RXFP2, in fact they do not have a large N-terminal extracellular domain and ligand binding is thought to occur via the 7TM, as occurs in type 1 small peptide receptors (Bathgate *et al*., 2013; Figure 9). Despite these structural differences, RLN3 and INSL5 can cross-activate RXFP3 and RXFP4, respectively (Bathgate *et al.,* 2013). Regarding the downstream signaling, RXFP3 couples to Gi/o and inhibits adenylyl cyclase (Liu *et al.*, 2003a; Westhuizen

et al., 2007). The receptors for INSL4 and INSL6 are currently unknown, and these peptides do not interact with any of the RXFP receptors (Bathgate *et al.*, 2013).

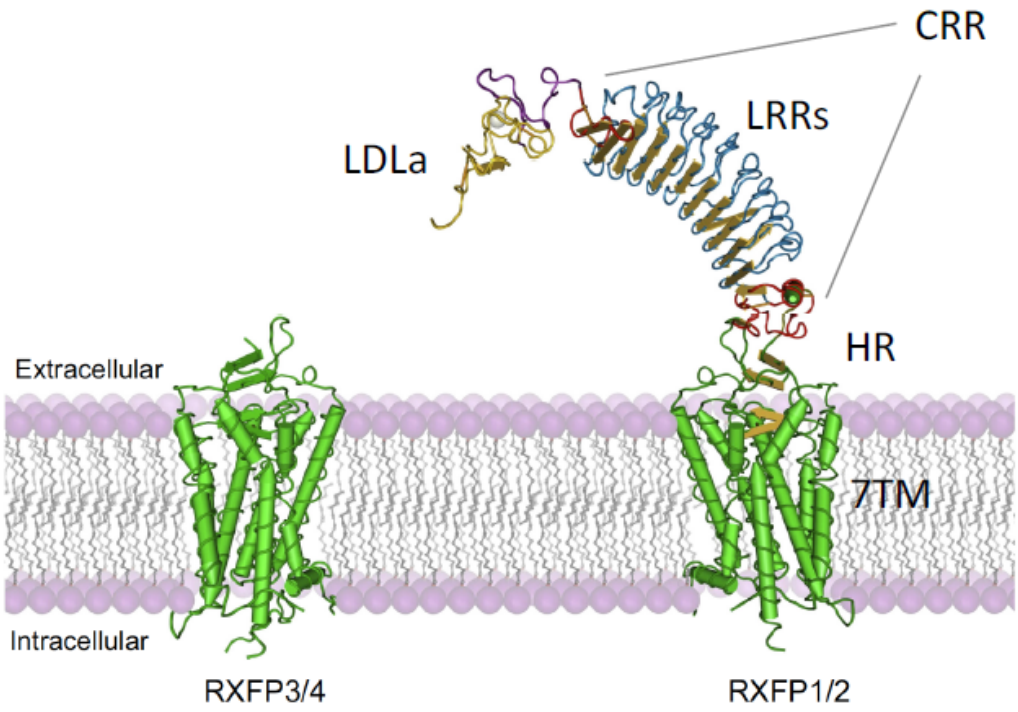


Figure 9. The two groups of vertebrate relaxin-like GPCR receptors. Figure adapted from Bathgate *et al.* (2013). RXFP1/2 are Type C1 LGRs. 7TM – seven transmembrane domain; CRR – cysteine-rich region; HR – hinge region; LDLa – low density lipoprotein receptor domain class A; LRR – leucine-rich repeat. Most of the C-terminal endodomain is not visible in this figure. The LRR is the ligand (relaxin-peptide) binding domain for RXFP1/2 receptors. We note that in both RXFP1 and RXFP2 there is an extra transmembrane domain localized N-terminally to the depicted structure (immediately upstream of the LDLa), which is nevertheless not depicted in this scheme. Whether or not this transmembrane domain serves as a processed signal peptide or is retained in the mature receptor form is not clear.

As mentioned above, two genes encoding LGR type C1 GPCRs have already been identified in the fruit fly genome, Lgr3 and Lgr4 (as well known as CG31096 and CG34411, respectively) (Hiel *et al.,* 2011). Lgr3 and Lgr4 share the exact same domain architecture with RXFP1 and RXFP2 (Figure 9), and while Lgr3 shows 46% sequence similarity to human RXFP2 and 45% to human RXFP1, Lgr4 shows 40% sequence similarity to both human receptors RXFP1 and RXFP2 (Hiel *et al.,* 2011).

Lgr4 is encoded by 18 exons like human LGR7 and LGR8, while Lgr3 is encoded by 11 exons (Hiel *et al*., 2014). The function of the *Drosophila* relaxin-like receptors Lgr3 and Lgr4 is unknown and strikingly, there have been no loss- or gain-of-function studies on these genes to the best of our knowledge. To gain insight into the function of these receptors is the principal aim of this work.

1.2.1 The relaxin paradox

As seen above, the presence of the RXFP1 and RXFP2 homologues Lgr3 and Lgr4 in *Drosophila* attests to the fact that this receptor family is an ancient one, probably found in all metazoans (Hiel *et al.,* 2011 ; Yegorov & Good, 2012). However, the same is not true for their ligands. While the Lgr3 and Lgr4 are orphan *Drosophila* GPCRs (*i.e.,* they have no known ligand), the ligands of the RXFP1 and RXFP2 receptors are relaxin-like peptides of the insulin-like peptide superfamily. As ligands and receptors are known to co-evolve, one would expect the Lgr3 and Lgr4, as well as other invertebrate homologues of the relaxin-like receptors, to be activated by ILPs. However, one must also consider the fact that the relaxins are, per se, an example that ligands and receptors do not always co-evolve: *i.e.,* relaxins are members of the ILP superfamily and they do not activate RTKs as the insulin/IGF peptides do. This "paradox" was strengthened by initial phylogenetic analysis that suggested that the RXFP3 receptor evolved early during vertebrate evolution, at the same time as RLN3 emerged (Wilkinson *et al.*, 2005; Bathgate & Wilkinson, 2007). The co-evolution of the ligand receptor pair RLN3 and RXFP3, suggested that RXFP3 was the ancestral form of the relaxin-receptor. As there are no clear homologues of RXFP3 and RXFP4 in invertebrates, this was considered as further evidence that the relaxins evolved in vertebrates (Wilkinson *et al.*, 2005; Bathgate & Wilkinson, 2007). Hence, the idea arose that invertebrates do not have relaxins, and therefore that the invertebrate relaxin-like receptors RXFP1/2-type are receptors for other types of ligands in invertebrates. However, further paleogenomic analyses using mostly synteny analyses have shown that the scenario is much more complex (Yogorov & Good, 2012) and new data and further sampling revealed that the co-evolution of RLN3 and RXFP3 was not exactly as previously reported (Wilkinson *et al.*, 2005; Bathgate & Wilkinson, 2007). These studies show strong evidence that the signaling of the ancestral relaxin peptide in the vertebrate (and chordate) ancestor occurred via the RXFP1 and RXFP2-type receptors rather than RXFP3 or RXFP4

receptors (Yogorov & Good, 2012). They also suggest that the last common ancestor to all vertebrates probably had two relaxin-like peptides, one similar to RLN and the other most similar to RLN3 (Yogorov & Good, 2012).

Another important fact to consider is the recent discovery of an insulin/relaxin-like peptide in the invertebrate starfish (reviewed in Mita, 2013). This peptide has a gonadotropin-like activity that is necessary for oocyte maturation and ovulation. Even though the receptor for this starfish insulin/relaxin-like peptide is still unknown, it clearly acts via G-protein and adenylyl cyclase (reviewed in Mita, 2013).

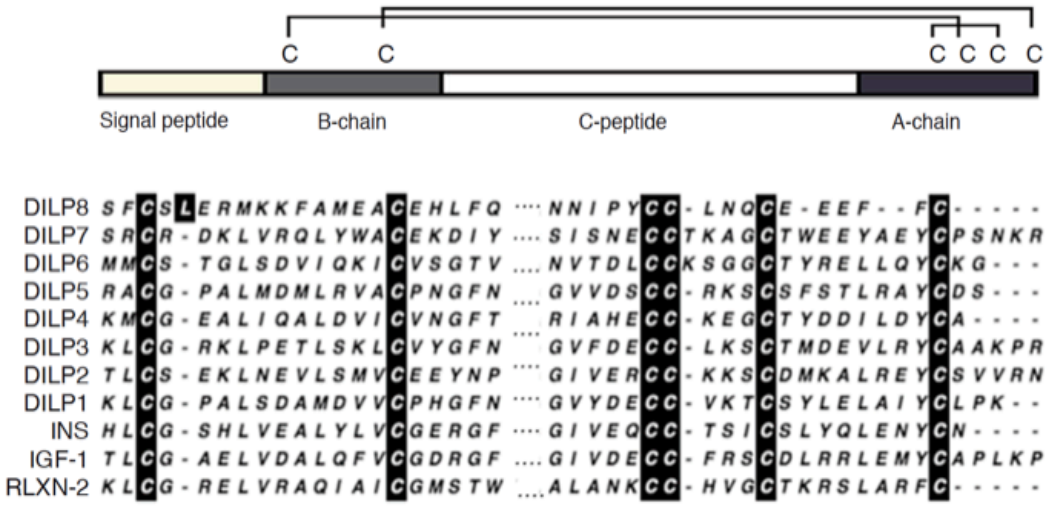


Figure 10. General Ilp scheme and protein alignment of the B and A chains of the eight *Drosophila* Ilps and comparison to three representatives of the human Ilp subfamilies. Image from Garelli *et al.* (2012). Absolutely conserved amino acids are highlighted in dark.

As described above, *D. melanogaster*, a model insect clade and genetic powerhouse (further described below), encodes two clear, yet orphan, relaxin-like receptors. It is unknown whether these receptors can be stimulated by any of the eight *D. melanogaster* Ilps (Dilp1-8; Brogiolo *et al.*, 2001; Ikeya *et al.*, 2002; Yang *et al.*, 2008; Okamoto *et al.*, 2009; Slaidina *et al.*, 2009; Gronke *et al.,* 2010; Garelli *et al.,* 2012) (Figure 10). Some Dilps (*e.g.*, Dilp2 and Dilp5) are strongly expressed in the so-called "insulin-producing neurons" in the brain, and are the major source of IIS-stimulatory signals (Miron *et al.*, 2001; Junger *et al.*, 2003; Puig *et al.*, 2003; Colombani *et al.*, 2005). Others, like Dilp6-8 have strikingly different expression patterns (Yang *et al.*, 2008; Okamoto *et al.*, 2009; Slaidina *et al.*, 2009; Colombani

et al., 2012; Garelli *et al.*, 2012). The only IGF-like Ilp in the fly genome is Dilp6 (Okamoto *et al.*, 2009; Slaidina *et al.*, 2009). All other Dilps are predicted to encode a C-peptide, suggesting they are either insulin or relaxin-like (Gronke *et al.*, 2010; Garelli *et al.*, 2012).

The binding of insect Ilps to insect InRs has only recently been demonstrated for mosquito Ilp3/4 (Brown *et al.*, 2008; Wen *et al.*, 2010) and the Drosophila Dilp5 (Sajid *et al.*, 2011). Furthermore, Dilp2 and 5 can induce tyrosine autophosphorylation of the Drosophila InR (Rulifson *et al.*, 2002). Other Ilps are usually assumed to activate IIS based either on sequence information indicating appurtenance to the Ilp family and/or their ability to modulate InR-dependent growth in gain and loss-of-function studies (Brogiolo *et al.*, 2001; Ikeya *et al.*, 2002; Rulifson *et al.*, 2002; Yang *et al.*, 2008; Okamoto *et al.*, 2009; Slaidina *et at.*, 2009; Zhang *et al.*, 2009; Gronke *et al.*, 2010). In Drosophila, this has been assumed for Dilp1, 3, 4, 6 and 7. Indeed, the genetic evidence for Dilp1-5 acting via an InR is indeed quite strong [*e.g.*, *dilp1-5* occur in a chromosomal cluster (Brogiolo *et al.*, 2001)], Dilp5 has been shown to bind InR (Sajid *et al.*, 2011), and Dilp2 overexpression can rescue the growth defect of animals lacking the *dilp1-5* cluster (Zhang *et al.*, 2009). However, the evidence is weaker, contradictory or absent for Dilp6-8, respectively (see further details below for each Ilp; Brogiolo *et al.*, 2001; Ikeya *et al.*, 2002; Yang *et al.*, 2008; Okamoto *et al.*, 2009; Slaidina *et al.*, 2009; Zhang *et al.*, 2009; Colombani *et al.*, 2012; Garelli *et al.*, 2012).

Dilp6 is a fat body-derived IGF-like peptide that regulates postfeeding growth in Drosophila (Okamoto *et al.*, 2009; Slaidina *et al.*, 2009). It is the sole IGF-like Ilp encoded in the Drosophila genome. Genetic evidence suggests Dilp6 works via IIS (Okamoto *et al.*, 2009; Slaidina *et al.*, 2009), but there is no biochemical data demonstrating the direct activation of the InR by Dilp6. The biochemical identification of the Dilp6 receptor awaits further study and should bring important contributions to the understanding of Dilp6-dependent activities.

Dilp7 shares properties at the sequence and regulatory level with members of the insulin/relaxin subfamily of Ilps (Yang *et al.*, 2008; Gronke *et al.*, 2010). Dilp7 is produced in ventral cord neurons (Brogiolo *et al.*, 2001; Miguel-Aliaga *et al.,* 2008;

Castellanos *et al.*, 2013), which are involved in food-based decision-making (Yang *et al.*, 2008), and modulate tracheal growth in response to nutrition in the posterior hindgut (Linneweber *et al.*, 2014). A previous report had found that Dilp7 overexpression induced organismal growth (Ikeya *et al.*, 2002), and hence it was considered to be an InR agonist. Yet more recent experiments using updated sequence annotations have failed to reproduce this finding (Dr. I. Miguel-Aliaga, ICL, UK, personal communication). Accordingly, Dilp7 deletion mutations show that it is not essential for growth and development (Gronke *et al.*, 2010). Therefore, the biochemical identification of a receptor for Dilp7 -be it the InR or not- shall certainly contribute to elucidate its molecular mechanism of action.

Dilp8 is an insulin/relaxin-like Ilp (Garelli *et al.,* 2012) that couples tissue growth with developmental timing (Colombani *et al.*, 2012; Garelli *et al.*, 2012). Dilp8 is produced and secreted from abnormally-growing imaginal discs (the larval precursors of most adult appendages) and transiently delays the onset of metamorphosis by inhibiting the biosynthesis of the major insect molting hormone (20-hydroxyecdysone (20E)). Concomitantly, Dilp8 slows-down growth of the unaffected imaginal discs via an unknown mechanism. How Dilp8 achieves this is unknown. Interestingly, in the adult stage, Dilp8 is strongly expressed in the ovary (Garelli *et al.*, 2012), similarly to vertebrate relaxins (Blundell & Humbel, 1980; Bathgate *et al.*, 2013). The biology of Dilp8 is further described below.

The lack of invertebrate Ilps with a clear vertebrate "relaxin signature" [RxxxRxxI/V] (Figure 7 and Figure 10) has led the scientific community to think that the ability of relaxin-receptors to be activated by relaxin-ligands was an innovation restricted to vertebrates. The elucidation of the biological function and hopefully the ligand of the Drosophila RXFP1/2-like receptors Lgr3 and Lgr4 should bring important insight into resolving this apparent paradox, apart from contributing to the understanding of the signaling and biology of this conserved family of GPCR receptors.

1.3 Coordinating growth and developmental timing

In many organisms, from mammals to insects, like *Drosophila*, development seems to occur in a perfectly coordinated way, resulting in a rather stable adult phenotype. Formally, developmental stability is the ability of an organism to buffer given traits against environmental and intrinsic perturbations (Denver & Middlemis-Maher, 2010; Minelli & Fusco, 2010). This may involve physiological, temporal or behavioral adjustments to the developmental program (Thornhill & Moller, 1997; Moczek, 2010). The processes leading to developmental stability have been particularly well studied in insects and other arthropods (Shingleton, 2010). For instance, early in 1927, it was observed that x-ray irradiation of *Drosophila* larvae negatively influenced the timing of pupariation, causing a developmental delay (Hussey *et al.*, 1927). Decades of further research suggested that the delay is mediated by a secreted signal, termed inhibitory imaginal signal, which is produced by regenerating damaged imaginal discs following irradiation (Simpson *et al.,* 1980). Hence, the extent of the developmental delay correlates with the amount of regenerative growth (Simpson *et al.*, 1980; Smith-Bolton *et al.*, 2010). The inhibitory imaginal signal communicates the abnormal growth status of the discs with the neuroendocrine centers that trigger the onset of metamorphosis (Simpson *et al.*, 1980). The major effect of the inhibitory imaginal signal is to reduce the production of 20E (Halme *et al.*, 2010; Garelli *et al.*, 2012; Hackney *et al.*, 2012), the critical hormone that triggers metamorphosis (reviewed in Hill *et al.*, 2012; Yamanaka *et al.*, 2013). Ecdysone, the 20E precursor, is biosynthetized in the prothoracic gland, a part of the composite endocrine organ known as the "ring gland", and is secreted directly into the hemolymph. Ecdysone is converted into 20E in the peripheral tissues. By inhibiting Ecdysone biosynthesis, the inhibitory imaginal signal induces a delay in the onset of metamorphosis in the presence of abnormal tissue growth, assuring the coordinated growth of tissues. Such abnormal growth-induced pupariation inhibition can also result from imaginal disc overgrowth (tumors) or slow growth, both of which cause a persistence of dividing mitotic cells (Menut *et al.*, 2007; Parker *et al.*, 2011).

The timing of ecdysone biosynthesis is determined by a pulse of prothoracicotropic hormone (PTTH), which is produced by a pair of neurosecretory

neurons in the brain that directly innervate the ring gland (McBrayer *et al.*, 2007). It is not clear what determines the synthesis of PTTH by the neurons, but a drop in the levels of the Juvenile Hormone (JH), a hormone that is produced in early larval stages and inhibits the larval to pupal transition, might be involved (Edgar, 2006). Halme and colleagues showed that the inhibitory imaginal signal acted upstream of the synthesis of PTTH using a wing disc genetic ablation technique (Halme *et al.*, 2010). Despite these works, the identity of the inhibitory imaginal signal remained a mystery until 2012.

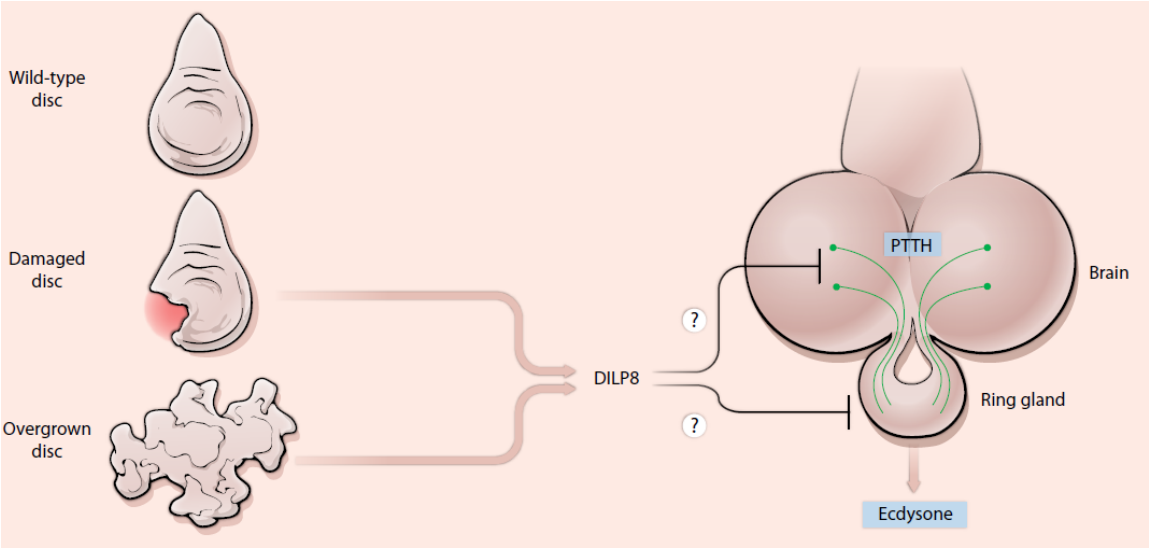


Figure 11. Hypotheses for how Dilp8 could act as an inhibitory imaginal signal. Image from Hariharan (2012).

In 2012, two groups independently reported that the insulin/relaxin-like peptide Dilp8 acted as an inhibitory imaginal signal during tumor growth, tissue regeneration and slow growth (Colombani *et al.*, 2012; Garelli *et al.*, 2012) (Figure 11). Dilp8 was found necessary in the imaginal discs to communicate the presence of local abnormal growth to the ring gland. Furthermore, as predicted, the ecotopic expression of Dilp8 in the absence of tissue damage was sufficient to inhibit ecdysone biosynthesis by the prothoracic gland. In the presence of imaginal disc damage, Dilp8 acted upstream of PTTH, inhibiting its transcriptional upregulation (Garelli *et al.*, 2012). Larvae mutant for Dilp8 were highly sensitive to tissue damage, in that the majority of the larvae died during metamorphosis and/or eclosed with malformed appendages. In the absence of exogenous damaging stimuli, the Dilp8 mutant larvae had an increased level of fluctuating asymmetry, suggesting that Dilp8 is required for the endogenous

coordination of intra-organ communication in face of intrinsic errors occurring during normal development (Garelli *et al.*, 2012).

The precise molecular mechanism by which secreted Dilp8 inhibits ecdysone biosynthesis to cause a developmental delay remains unknown, but there are many hints (Hariharan, 2012; Figure 11). Dilp8, as a member of the secreted Ilp hormone family, is expected to have a receptor. The determination of the Dilp8 receptor could allow the determination of the target tissue of Dilp8 activity. Dilp8 could act either via the sole *Drosophila* InR, the unexplored relaxin-like receptors or other unknown receptors. Here, as part of our objective in understanding the biological role of the *Drosophila* relaxin-like receptors Lgr3 and Lgr4, we plan to use genetics to test the possibility that they could act as receptors for Dilp8.

1.4 Aims

The aims proposed for this project are the following: 1 – Generate mutant alleles for both receptor candidates (*lgr3* and *lgr4*) and characterize them at a molecular level. 2 - To determine the function of both receptors. 3 – Determine their relationship with Dilp8 by performing developmental time assays, and test if this receptors can restore the delay that characterize this peptide.

Chapter 2. Materials and Methods

2.1 *Drosophila melanogaster* as a genetic tool

Drosophila, also known as "vinegar" or "fruit" flies, are a group of small flies (measuring about 3 mm long) that are highly diverse in appearance, behavior and in reproduction habitat. The *Drosophila* group contains around 1500 species (Robert *et al.*, 1999). *Drosophila* are dipteran (*i.e.,* two-winged) insects of the family *Drosophilidae. D. melanogaster*, one of the major model organisms in genetics and development, due to, among other characteristics, its small size, relatively large brood and short generation time (see more details below), belongs to the *Sophophora* subgenus of *Drosophila*.

Since the discovery of the white mutation and its linkage to the X chromosome by T.H. Morgan in the early 1900s, *Drosophila* has been used extensively as an invertebrate model organism, rich in tools and advantages in genetic, evolutionary and developmental research. In the first 50 years of research in *Drosophila*, the attention of the scientific community was focused mainly in the principle of inheritance. Over the years, the development of new molecular tools, concepts and approaches (Figure 12) further revealed the enormous potential of this model, which could allow the understanding of some mechanisms in vertebrates, given the homology between vertebrates and invertebrates in critical aspects (Bellen *et al.*, 2010).

In the mid 2000s, the *Drosophila* genome was completely sequenced and published (Adams *et al.*, 2000). The fly genome represents approximately 5% of the human genome, comprising 4 pairs of chromosomes, which encode nearly 125 million base pairs (bp) of DNA and approximately 13600 gene products (Bier & Bodmer, 2004; Wolf & Rockman, 2008, Wolf & Rockman, 2011). About 75% of the human genes responsible for diseases have homology in the *Drosophila's* genome (Reiter *et al.*, 2001), which makes it a valuable model to better understand the role of some proteins in humans.

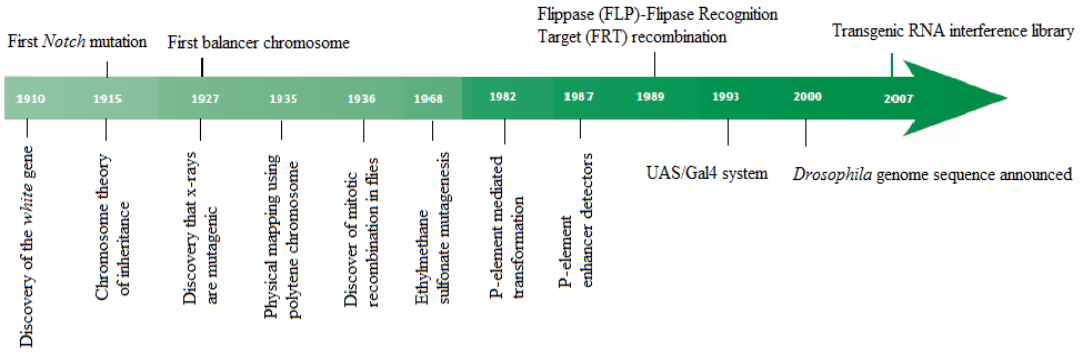


Figure 12. Some key discoveries and methodologies used in fruit flies over the years, from "100 years of Drosophila research and its impact on vertebrate neuroscience: a history lesson for the future", by Bellen *et al.*, 2010.

Drosophila has many other advantages over other invertebrate model organisms to study biological processes *in vivo*: flies are easy and cheap to obtain and maintain in laboratories due to their small size; their short generation time, about 10 days at room temperature (~25ºC; development may be shorter or longer in higher or lower temperatures, respectively); the small number of chromosomes: only four pairs, making genetics very simple; the relatively early availability of a very well annotated sequenced genome; most of the genetic mechanisms, such as gene transmission, linkage, sex determination, genetic interactions, chromosomal aberrations, penetrance and expressivity, are shared between the majority of organisms; it is easy to study molecular, biochemical and developmental mechanisms as well as evolutionary changes in the fruit flies, because this type of mechanisms are well conserved through the evolution; sophisticated genetic tools, such as P-elements, transgenic fly mutants carrying RNA interference (RNAi), single nucleotide polymorphism (SNP) maps, and tissue and temporal specific transgene expression through UAS-GAL4 system; the presence of balancer chromosomes that can be followed by easily observed physical characteristics; and the ability to perform large-scale forward genetic screens, to cite just a few, are some of the great advantages of *Drosophila* as a model organism.

Each fertilized female can lay approximately 50-70 eggs per day. An average egg is about 0.5 mm long and its formation occurs in the *ovaries*. Once the formation

of the mature oocyte is complete, it is kept in the ovaries until it is ovulated one at a time following fertilization. Fertilized females store sperm in the *sperm receptaculum* and *spermatheca*. The ovulated oocyte then passes through the *oviducts* and reaches the *uterus* where it is exposed to sperm cells. These enter the egg through a micropyle localized in its posterior dorsal side. Once the female finds a suitable place to lay the egg, the egg-laying process starts and another ovulation cycle begins.

In *Drosophila*, all of the embryonic development occurs out of the female. Embryogenesis lasts about one day at room temperature (around 25ºC) (Figure 13). The embryos then hatch into first instar larvae, 24 hours later, and begin feeding in the substrate they were laid upon. After one day feeding, the larvae progress to the second instar larvae, and then to the third instar larvae after another day feeding. The third instar, the last and longest larval stage, takes about two days. The third instar larva stage can be further divided into foraging and wandering stages. During wandering stage, the larva stops feeding and leaves the food towards a dry place to stop moving and start pupariation [a 12-h period that is marked by a strong pulse of Ecdysone (20E)]. During pupariation, the 3rd instar larval cuticle hardens and it becomes the puparia or pupal case. The epidermis of the larvae then detaches from the cuticle, in a process termed apolysis. When this is completed, head eversion takes place and marks the start of the true "pupal" stage of this insect, where metamorphosis occurs (pupation; see further details below). A last pupal to adult ecdysis takes place within the pupal case early in the pupal stage, producing the pharate adult. The adult fly eclodes (hatches) about 4 days after the onset of pupation.

Drosophila larvae are segmented into 14 segments: 3 head segments, 3 thoracic segments and 8 abdominal segments, and contain a certain number of undifferentiated monolayer mitotically-dividing epithelial cell sacks that are called imaginal discs. These distinct and discretely growing sacks bear this name because they will eventually differentiate into many appendages of the adult fly, also known as the *imago* stage (Figure 14).

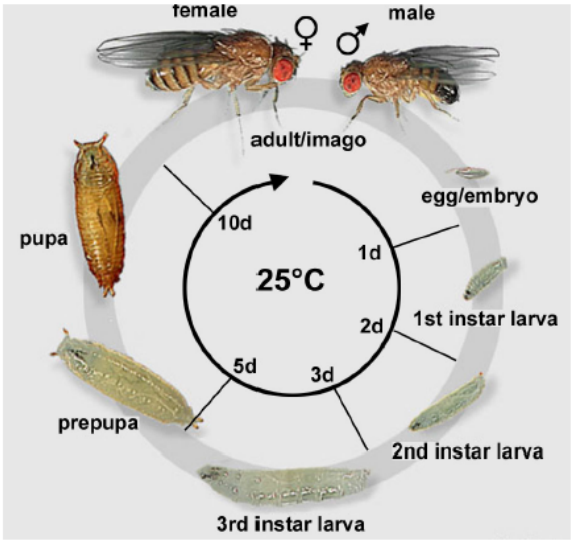


Figure 13. The life cycle of *Drosophila melanogaster*, from Roote *et al.,* 2013.

Metamorphosis consists in the process of transition from larva to adult. All significant developmental changes occur in the pupal stage and all the transformations last about 12 hours (Deepa *et al.*, 2009). This process involves the serial destruction of larval tissues and the organ formation from the imaginal disc cells (Figure 14). Once this process is completed, adult flies are ready to emerge from the pupa. In the first hours, the newly-eclosed flies are light in color, their wings are not yet expanded and they are not sexually mature. This allows the researcher to easily distinguish recently emerged flies from the older ones in order to ensure the collection of virgin females, necessary for a controlled mating (Deepa *et al.*, 2009).

2.2 *Drosophila melanogaster* lines and breeding

Fly stocks used in this work were maintained in vials with standard cornmeal-agar medium at 25°C for experiments, or at 18°C, for stock maintenance and appropriate conditions of humidity for each experiment.

All *Drosophila* stocks used in this work are described in Table 1. These stocks were either obtained from the Bloomington Drosophila stock center (http://flystocks.bio.indiana.edu/), obtained from other laboratories as generous gifts,

or generated in the lab by genetic crosses and/or from genome excisions (as described in Metaxakis *et al.*, 2005).

Fly virgin collections were performed by following the directions stated by Ashburner for all the crosses along the work (Ashburner *et al.*, 2007).

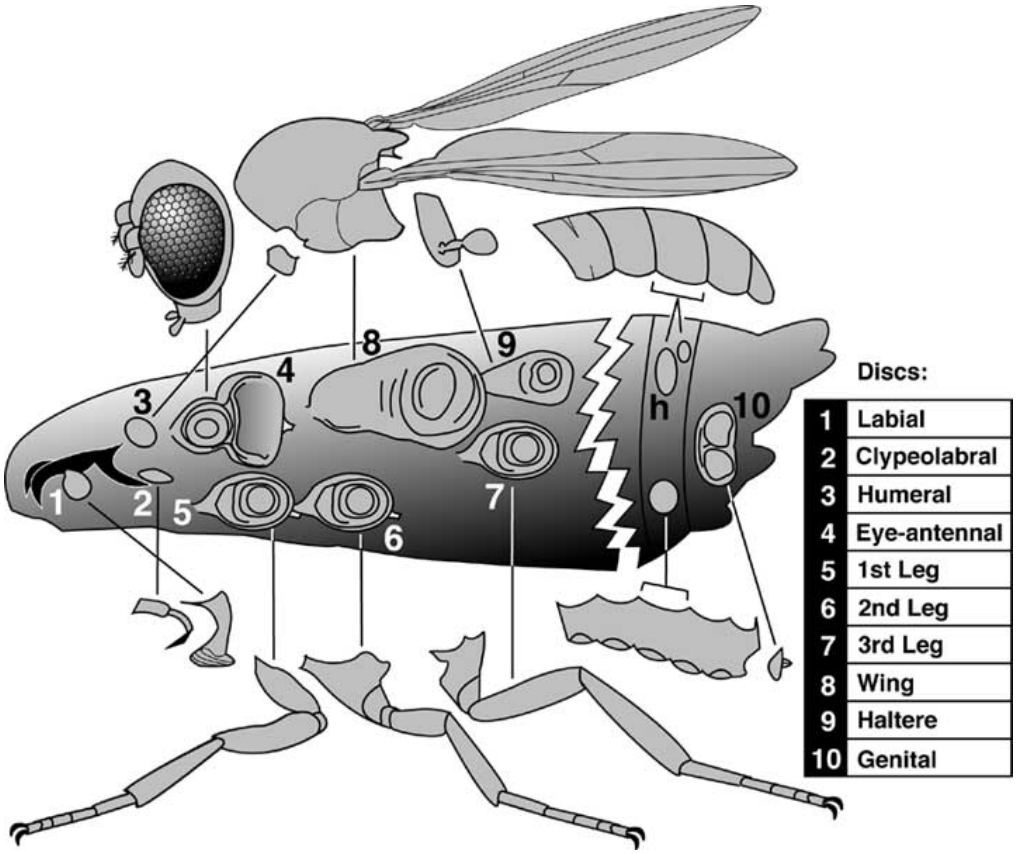


Figure 14. Imaginal discs and the adult appendages that they will form, from Lewis, 2005.

2.3 Excision protocol

In this work we used a classic transposable element remobilization approach to induce imprecise excision mutations in our target genes. Transposable elements are mobile genetic elements that colonize genomes by a copy and paste mechanisms (Rhodes, 1938; McClintock, 1983). During the last decades the *Drosophila* community has generated large collections of fly stocks that carry single genetically modified transposable elements inserted into almost every gene in the *Drosophila* genome. These modified transposable elements usually lack the element-specific enzymatic activity (termed Transposase) that is required for their further mobilization (*i.e.,* remobilization), hence they remain as a single stable insertion in

the genome. These stocks are publically available and provide a valuable tool for many different experimental approaches.

Table 1. *Drosophila st*ocks used in this work.

Stock number	Name	Origin	Genotype
1	*w[1118]*	M. Dominguez's lab	*w[1118]*
2	*iso-1*	Bloomington #2057	*so-1 : y[1]; Gr22b[1] Gr22d[1] cn[1] CG33964[R4.2] bw[1] sp[1]; LysC[1] MstProx[1] GstD5[1] Rh6[1]*
3	*Oregon R*	Bloomington #5	*Oregon-R-C*
4	*if/CyO;MKRS/TM6B*	A Jacinto's lab	*w[1118]; if/CyO; MKRS/TM6B*
5	*FM7*	Bloomington #5708	*Nrg[l4]/FM7C*
6	*MKRS/TM6B*	Generated in the lab	*MKRS/TM6B*
7	*Minus Transposase - if/Cy-Minus*	Bloomington #24613	*w[1118]; if/P{w[+mC]=hsILMiT}2.4 (from 24613); Mkrs/Tm6*
8	*Minus Transposase*	Bloomington #24613	*w[1118]; sna[Sco]/SM6a, P{w[+mC]=hsILMiT}2.4*
9	*tub-dilp8*	M. Dominguez's lab	*tub-dilp8 (1)*
10	*lgr3[MB06848]*	Bloomington #25253	*w[1118];; Mi{ET1}lgr3[MB06848]*
11	*lgr3[ex1]*	Generated in the lab	*w[1118];; lgr3[ex1]*
12	*lgr3[exp1]*	Generated in the lab	*w[1118];; lgr3[exp1]/ TM6B*
13	*lgr3[exp2]*	Generated in the lab	*w[1118];; lgr3[exp2]/ TM6B*
14	*lgr3[exp3]*	Generated in the lab	*w[1118];; lgr3[exp3]/ TM6B*
15	*tub-dilp8; lgr3[ex1]*	Generated in the lab	*w[1118]; tub-dilp8(1); lgr3[ex1]*
16	*lgr4[MB03440]*	Bloomington #23615	*w[1118] lgr4[MB03440]*
17	*lgr4[ex1]*	Generated in the lab	*w[1118] lgr4[ex1]*
18	*lgr4[ex3]*	Generated in the lab	*w[1118] lgr4[ex3]*
19	*lgr4[exp1]*	Generated in the lab	*w[1118] lgr4[exp1]*
20	*lgr4[exp2]*	Generated in the lab	*w[1118] lgr4[exp2]*
21	*lgr4[ex1]; tub-dilp8(1)*	Generated in the lab	*w[1118] lgr4[ex1]; tub-dilp8(1)*
22	*lgr4[ex3]; tub-dilp8(1)*	Generated in the lab	*w[1118] lgr4[ex3]; tub-dilp8(1)*
23	*lgr4[ex1];; lgr3[ex1]/TM6B*	Generated in the lab	*w[1118] lgr4[ex1];; lgr3[ex1]/TM6B*
24	*lgr4[exp1];; lgr3[ex1]/TM6B*	Generated in the lab	*w[1118] lgr4[exp1];; lgr3[ex1]/TM6B*
25	*lgr4[ex1];; lgr3[exp2]/TM6B*	Generated in the lab	*w[1118] lgr4[ex1];; lgr3[exp2]/TM6B*

If crossed to transgenic stocks that express the Transposase enzyme, the remobilization of these transposable elements can be induced in these lines. Usually, the remobilization of the transposable element will not leave any molecular scar in the locus where it was inserted, so that the event is called a "precise" excision. However, in some cases, the remobilization event is not performed and/or repaired correctly by the cellular machinery, producing a deletion mutation in the locus where the transposable element was once inserted. These rarer events are called imprecise excisions. Imprecise excisions can generate deletions of several kilobases, making them an interesting way of producing loss-of-function alleles in the genes that have been targeted with these genetically modificed transposable elements by the fly community over the years. It must be considered that in some cases the presence of a transposable element insertion *per se* within a gene locus can alter the activity of the gene, typically in a negative fashion. This is more common when the transposable elements fall into regulatory regions (*e.g*., promoters) or coding regions (exons). However, some insertions, like intronic insertions, are silent and/or cause weak effects on gene transcription and/or splicing of the gene.

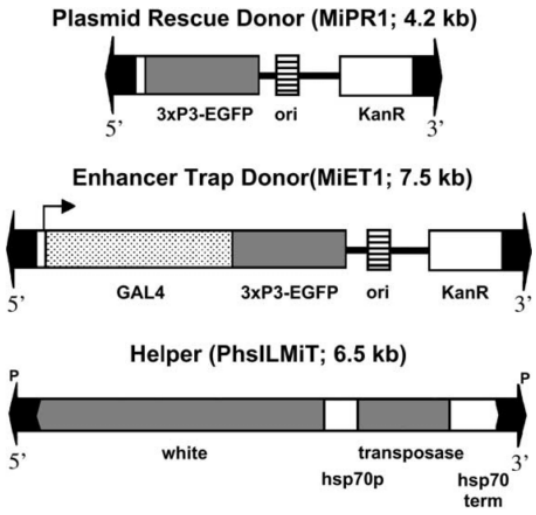


Figure 15. *Minos* donor and helper constructs containing the 3xPax6/EGFP dominant marker. The helper construct expresses *Minos* transposase under heat-shock and is based on pCaSper. Only the transposon regions are shown. Image from Metaxakis *et al.* (2005).

In order to generate excision events we identified lines carrying transposable element insertions in regions in the *lgr3* or *lgr4* genes that would more likely delete coding regions after the imprecise excision of the elements. We identified the *MB*

Minos elements *lgr3[MB06848]* and *lgr4[MB03440]* (Metaxakis *et al.,* 2005; Figure 15), both of which were inserted in intronic regions less than 100 bp from the nearest coding exon (Figure 16 and Figure 17). MB elements carry an eye-specific enhancer driving eGFP, which can be used to determine the presence or not of the insertion (Figure 15).

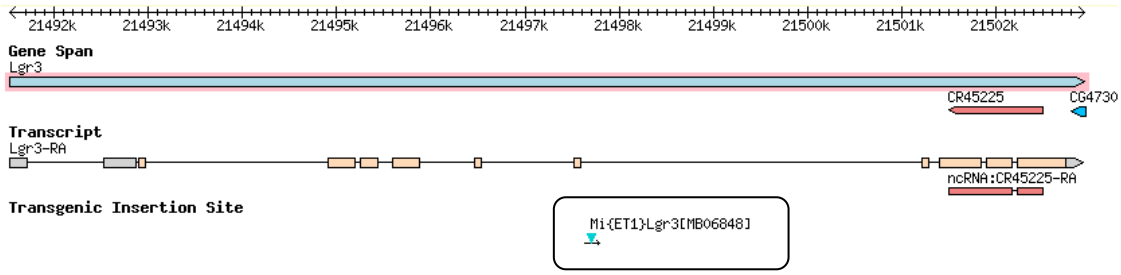


Figure 16. Flybase genome browser snapshot showing the location of the *Minos* element *lgr3[MB06848]* gene insertion (blue triangle) in the *lgr3* locus on the 3rd chromosome of the *D. melanogaster* genome.

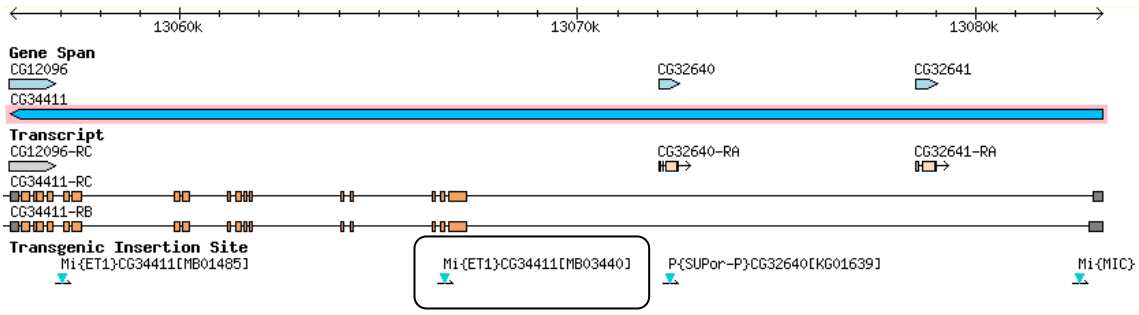


Figure 17. Flybase genome browser snapshot showing the location of the *Minos* element *lgr4[MB03440]* gene insertion (blue triangle) in the *lgr4* locus (*CG34411*) on the X chromosome of the *D. melanogaster* genome.

To generate the excisions we used slightly different crossing schemes for *lgr3[MB06848]* and *lgr4[MB03440]* due to the fact that they are in different chromosomes (III and X, respectively; Figures 18 and 19). In brief, virgin females of the stocks carrying either element, were crossed with males of a stock, carrying the MiET1 transposase insertion on a Cy balancer chromosome (Cy-Minos) as described in Metaxakis *et al.* (2005). The flies were transferred to new vials every day, and two days after, the old vials containing the F1 progeny were heat-shocked daily until pupariation by inserting the vials in a water-bath at 37°C for 1 h. In the heat-shock

process, the remobilization of the transposon occurs and the remobilized transposon can reinsert at different regions or not. If the remobilization event occurred in the germ line and the transposon did not land in any other inherited site, this can lead to the lack of the transposon-driven eGFP-cassette in the progeny. Because os this, we can select eGFP-negative flies in the F2 generation. By selecting only one fly per vial we are reassured that we are dealing with independent events. Then the pupae were allowed to develop until adult phase in normal growth conditions. For *lgr3,* Cy-Minos/+; *lgr3[MB06848]/MKRS* or TM6B male adults were selected and individually crossed to the balancer strain *if/CyO; MKRS/TM6B*. From each vial we then selected a single eGFP-negative and back-crossed them to *if/CyO; MKRS/TM6B* and selected against the balancer chromosomes (Figure 18). The putative *lgr3* excisions were balanced over TM6B to get the following genotypes *w[1118]*; +/+; *lgr3[MB06848]excision/TM6B*. For *lgr4,* Cy-Minos/+ males were selected and crossed to the X chromosome balancer strain *Nrg[l4]/FM7C*. Putative excision lines *lgr4[MB03440]excision/FM7C* were selected and crossed to the *Nrg[l4]/FM7C* strain to select against balancer chromosomes (Figure 19). The final genotype of *lgr4* excision mutations was *lgr4[MB03440]excision /lgr4[MB03440]excision*; +/+; +/+.

2.4 Genomic DNA extraction

In this work, we used a non-lethal PCR genotyping technique to perform genomic DNA extraction from a single fly (Carvalho *et al.*, 2009). Since this protocol is effective using small body parts, this type of procedure allows direct molecular screening of first generation individuals, turning the method faster, cheaper and less laborious. From each established line used in the jump-start protocol, we isolated one to two male flies of the homozygous individuals, which were macerated using pellet pestles and homogenized in 100 µl DNA Extraction Buffer (1 M Tris-HCl at pH 8.2, 0.5 M EDTA, 5 M NaCl). Then, we added 1 µl Protease K 50 ng/µl (Roche), and incubated the mixture at 37°C for 1 h, followed by 95°C for 5 min, to inactive the protease. In the end of the procedure genomic DNA is ready to be used and can be stored at -20°C.

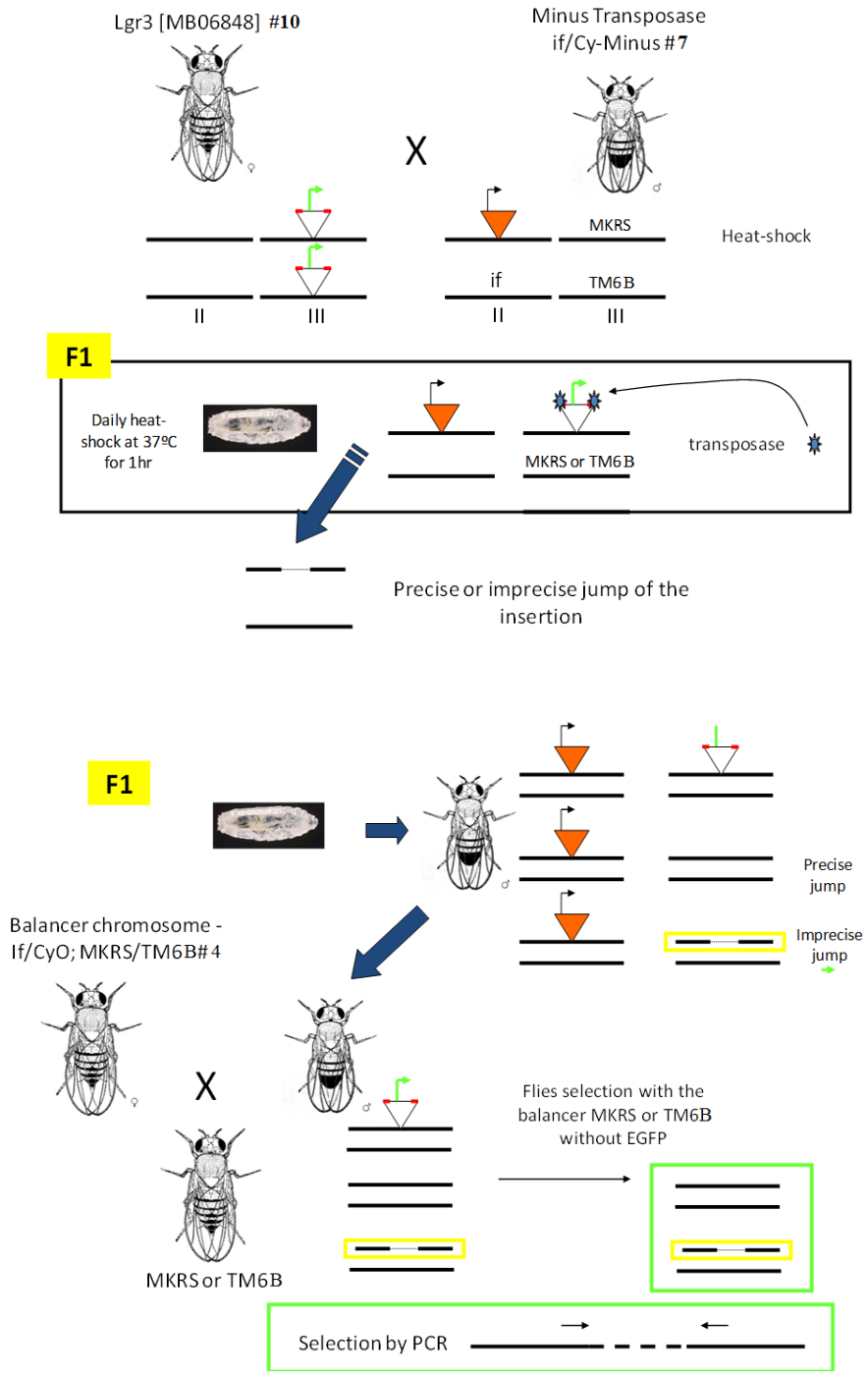


Figure 18. Lgr3 "jump start" experiment. First virgin females of the stocks carrying the *lgr3[MB06848]* element, were crossed with males of a stock, carrying the MiET1 transposon insertion on a Cy balancer chromosome (Cy-Minos). The flies were transferred to new vials every day, and two days after, the old vials containing the F1 progeny were heat-shocked daily until pupariation by inserting the vials in a water-bath at 37°C for 1 h. Cy-Minos/+; *lgr3[MB06848]/MKRS* or TM6B male adults were selected and crossed to the balancer strain *if/CyO; MKRS/TM6B*. From each vial, we then selected a single eGFP-negative fly and back-crossed them to *if/CyO; MKRS/TM6B* and selected against the balancer chromosomes.

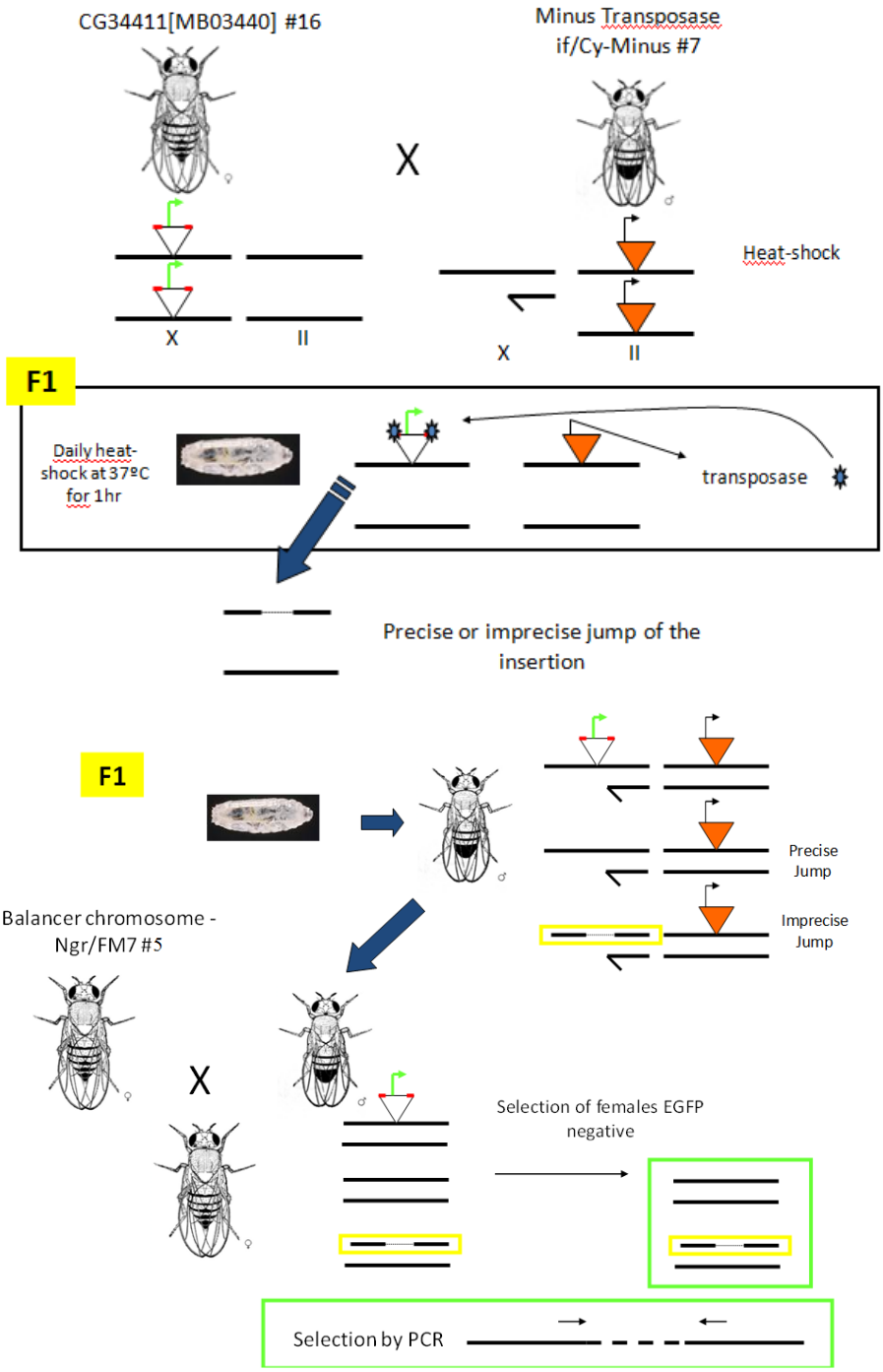


Figure 19. Lgr4 "jump start" experiment. First, virgin females of the stocks carrying the *lgr4[MB03440]* element, were crossed with males of a stock, carrying the MiET1 transposon insertion on a Cy balancer chromosome (Cy-Minos). The flies were transferred to new vials every day, and two days after, the old vials containing the F1 progeny were heat-shocked daily until pupariation by inserting the vials in a water-bath at 37°C for 1 h. Cy-Minos/+ males were selected and crossed to the X chromosome balancer strain *Nrg[l4]/FM7C*. Putative excision lines *lgr4[MB03440]excision/FM7C* were selected and crossed to the *Nrg[l4]/FM7C* strain to select against balancer chromosomes.

2.5 RNA extraction

RNA is the most instable of nucleic acids. Therefore, the extraction procedure had to be carefully performed to avoid degradation of the RNA samples from animal tissues and prevent gene expression artifacts. Extraction and purification of the RNA was made using the Direct-zol™ RNA MiniPrep (Zymo Research). We collected 5 virgin females and 10-15 virgin males of the desired genotypes aged between 3-7 days after eclosion, which gives the equivalent of a maximum of 50 mg of tissue. More than 50 mg of tissue can exceed the RNA binding capacity of the spin column, decreasing largely the efficiency of the reaction. Firstly, flies were macerated using pellet pestles, homogenized in 500 µl TRI Reagent® and centrifuged at 12000 *g* for 1 min, to lower tissue debris. We added one volume of ethanol at 100% directly to one volume of sample (1:1) and centrifuged the mixture at 16000 *g* for 1min. Then the supernatant was carefully transferred into a Zymo-Spin™ IIC Column and again centrifuged at 16000 *g* for 1 min. At this point all the centrifugation steps were performed at 16000 *g*. 400 µl RNA Wash Buffer were added to the Zymo-Spin™ IIC Column and centrifuged for 1 min. The supernatant was discarded and 80 µl of DNase I cocktail (DNase I 5 U, 8 µl 10x DNase I Reaction Buffer, 3 µl DNase/RNase-Free Water, and 64 µl RNA Wash Buffer), were added directly to the matrix of the column, which was incubated at 37°C for 30 min and then centrifuged for 30 s. The column was transferred into a new collection tube and 400 µl Direct-zol™ RNA PreWash was added and centrifuged for 1 min. The supernatant was discarded and this step was performed again. Then 700 µl of RNA Wash Buffer were added to the column and centrifuged for 1 min. The last supernatant was discarded and the centrifugation step was performed again for 2 min with the empty collection tube, in order to ensure the complete removal of the wash buffer. Lastly, the column was transferred to a 1.5-ml microcentrifuge tube and 30 µl of DNase/RNase-Free Water was directly added to the column matrix and centrifuged for 1 min. To maximize the RNA yield, the last step was repeated. RNA concentration and quality were evaluated in a Thermo Scientific NanoDrop™ 2000 Spectrophotometer and then the purified RNA was stored at -80°C.

Purified RNA was subjected to a second round of DNAse treatment, in order to reduce contaminations with DNA following RNA preparation. This treatment is

often used in samples that will be analyzed by specific and sensitive methods for RNA analysis, as qPCR. Briefly, the total volume of purified RNA was incubated at 37ºC with the same volume of Turbo DNase Buffer and 1uL of Turbo DNase, for 30 minutes. Then is added 0.1 volume of DNase Inactivation Reagent, incubated for 5 minutes at room temperature. In the end of the procedure, the sample is centrifuged at 10.000 *g* for 1.5 minutes and the RNA is transferred into a new fresh tube.

Further RNA cleaning and concentration was performed using the RNA Clean & Concentrator™-25 kit (Zymo Research). Two volumes of RNA Binding Buffer were added to each volume of RNA sample. We then added one volume of 100% ethanol to this mixture and transferred it to the Zymo-Spin™ IIC Column in a collection tube and centrifuged. The supernatant was discarded and 400 µl of RNA Prep Buffer were added to the column and centrifuged for 1 min. We then added 800 µl of RNA Wash Buffer and centrifuged it for 30 s. This step was repeated by adding another 400 µl of the same buffer. All the centrifugation steps were performed at 12000 *g*. The Zymo-Spin™ IIC Column was transferred to a 1.5-ml microcentrifuge tube and 25 µl of DNase/RNase-Free Water was directly added to the column matrix and allowed to dry at room temperature for 2 min, then centrifuged at 10000 *g* for 30 s. To maximize the RNA yield, the last step was repeated. With the original volume, the RNA concentration was again estimated as previously described and the eluted RNA was stored at -80°C.

2.5.1 cDNA reaction

cDNA synthesis was performed using the Maxima First Strand cDNA Synthesis Kit for RT-qPCR (ThermoScientific), enabling the quantitative analysis of gene expression. All necessary components are shown in Table 2.

Table 2. Mix reagents for cDNA reaction.

5x Reaction Mix	2µL
Maxima Enzyme Mix	1µL
RNA template	2µL
Nuclease-free water	5µL
Total volume	10µL

The reaction mix, enzyme mix, RNA sample and nuclease-free water were all gently mixed in an RNAse-free tube, centrifuged, and incubated for 10 min at 25°C, for 30 min at 50°C and for 5 min at 85°C. To complete the reaction, samples were diluted according to the initial quantification of RNA, to obtain a final concentration of 1-5 ng/µl to avoid an eventual inhibition of downstream qPCR reactions. A reverse-transcriptase negative (RT-) control reaction (*e.g.,* without maxima enzyme mix) was performed in parallel for all cDNA reactions in order to check for DNA contamination in the RNA samples.

2.6 Primer design

For this study, all primers were designed using Primer - BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primer specificity was also checked using the same program. All primers were synthesized by Sigma, USA.

2.7 Polymerase Chain Reaction

Polymerase chain reaction is an *in vitro* technique of DNA amplification, based on repeated thermal cycles that allow the amplification of large amounts of target DNA sequences even with minimal amounts of initial DNA. The thermal cycles consist in four steps: denaturation, annealing, extension, and final elongation. These break the bonds between nucleotides, separating the double-stranded DNA into two single strands, allow the annealing of designed primer pairs that are complementary to specific sequences that flank the DNA fragment of interest, and the extension of the primer-sequence hybrids. DNA polymerase binds to free 3'-OH ends of the primers and synthesize the new complementary strand of the DNA fragment, from 5' to 3' end. The new synthesized strands serve as template for complementary synthesis in the next few cycles.

We used a T100 Thermal Cycler™ (BioRad) for performing the PCR steps. In this type of technique the master mix can be prepared by using all the reagents separately or by using a commercial Enzyme Mix that already has all the necessary reagents, excepting the gDNA and the specific primers. The PCR reaction conditions were optimized depending on the aim (reaction final volume) and on the fragment size of the product (extension time). Tables 3-6 give standard reaction recipies and primer pairs that were used in this study.

Table 3. Standard PCR reaction reagents.

Reagent	Concentration	Volume
Supreme 2x Green Master Mix (NZYTech)	Enzyme concentration: 0.2 U/µL	
Forward primer	1µM	1µL
Reverse primer	1µM	1µL
DNAse-free Water	not applicable	2µL
gDNA	not applicable	1µL
	Reaction final volume= 10µL	

Table 4. PCR conditions for amplification of interest genes.

Cycles	Cycle step	Temperature	Time
1x	Denaturation	95°C	5 min
34x	Annealing	95°C	30 s
		60°C	30 s
	Extension	72°C	1-3 min*
1x	Final elongation	72°C	5 min
	Final hold	12°C	∞

*depending on the fragment size, 1min per 1000bp.

Table 5. Lgr3 Primers used for Polymerase Chain Reaction.

Primer name	Gene	Primer sequence		Primer size (bp)	Fragment size (bp)
Lgr3_salto_fw	Lgr3	Fw	CCGACGCCTTGCTGCTAACT	20	866
		Rv	TTTATGGAGCGGGCGTGGTC	20	
Lgr3_exonshort	Lgr3	Fw	CCGACGCCTTGCTGCTAACT	20	331
		Rv	GTGCGTTATGAGGTTGTGCTG		
Lgr3_exon3p	Lgr3	Fw	CGCCTTGTCGGTAATCCCAT	20	240
		Rv	GTGGCTCCATTAAACTGCTGC	21	
Lgr3_exons	Lgr3	Fw	CCGACGCCTTGCTGCTAACT	20	5307
		Rv	CAAAGACCACCAACCAGGCGTA	22	

Table 6. Lgr4 Primers used for Polymerase Chain Reaction.

Primer name	Gene	Primer sequence	Primer size (bp)	Fragment size (bp)
Dellgr4_short	Lgr4	Fw GTTCACTGAAGAAAGTTTCGTGGAT	25	761
		Rv GCAATTGAAGCTATACGAACAGGAA	25	
Dellgr4	Lgr4	Fw TTCGGTAGTCGGGCAAAGAC	25	5502
		Rv GCAATTGAAGCTATACGAACAGGAA	25	
cDNA_L4	Lgr4	Fw GTTCACTGAAGAAAGTTTCGTGGAT	25	275
		Rv AAAACTTTTCCTGTCCCTGTCGAG	24	
Lgr4_exon3-4	Lgr4	Fw TCGCCAATCCTGAGATTATCGTGT	24	347
		Rv GCAATTGAAGCTATACGAACAGGAA	25	

2.7.1 Electrophoresis on agarose gel

Electrophoresis is a technique based on the movement of charged particles in a solid matrix by applying an electric field across the mixture, as for example a 1.2% agarose gel that supports the molecules during migration. The migration of these molecules occurs according to the ratio of molecular charge and weight. The larger the molecules, the slower they migrate. All agarose gels used in this experimental work were prepared in a concentration of 1.2% by dissolving 0.6 g agarose, NZYTech, in 50 ml of TAE 1x buffer. A molecular marker was always migrated along with the samples in order to obtain to identify the molecular weight of the migrating DNA bands. For this we used the following molecular weight markers: NZYDNA Ladder III (NZYTech), 1-kb DNA Ladder (Invitrogen) and 50 bp DNA Ladder (ThermoScientific) (Figures 20-22). Electrophoresis was performed in TAE 1x buffer with a constant voltage of 100V in a PowerPac 300® system (BioRad).

For each 50 ml of agarose gel, we added 2 µl of Green Safe Premium (NZYTech) to allow the visualization of the DNA products subjected to gel electrophoresis under UV light in a Molecular Image Chemidoc™ XRS+ (BioRad).

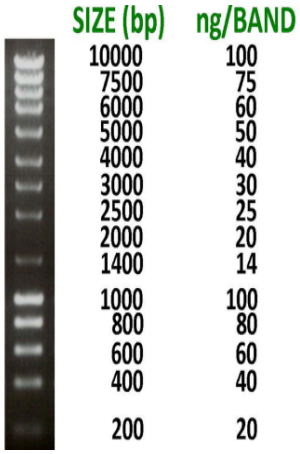


Figure 20. NZYDNA Ladder III, NZYTech.

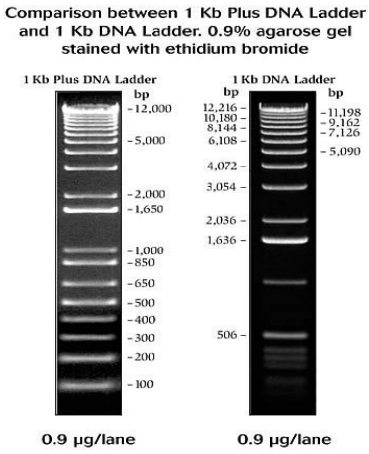


Figure 21. 1kb DNA Ladder, Invitrogen.

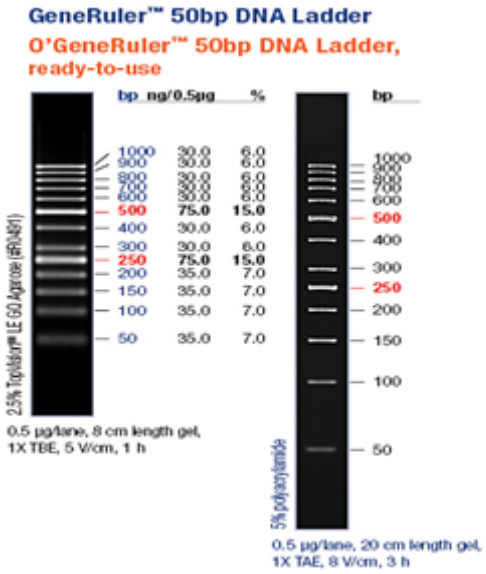


Figure 22. GeneRuler 50bp DNA Ladder, ThermoScientific.

2.7.2 DNA purification and sequencing

In order to send DNA samples for classical Sanger sequencing, we performed a DNA purification step using the NZYGel pure Kit (NZYTech) following PCR amplification and confirmation of product band size by gel electrophoresis. The desired PCR product was transferred into a 1.5-ml microcentrifuge tube, to which were added five volumes of Binding Buffer (NZYGel pure Kit, NZYTech). Then the mixture was transferred into the NZYTech spin column, incubated for 2 min at room temperature, and centrifuged for 1 min at 13000 rpm. Next, we added 600 µl of Wash Buffer to the column and centrifuged it for 1 min, discarding then the supernatant. To further remove any residual ethanol, the spin column was centrifuged again for 2 min. Then the NZYTech spin column was placed in a clean 1.5-ml microcentrifuge tube, and the DNA sample was eluted in 30 µl of Elution Buffer and centrifuged for 1 min. This step was performed in duplicate, in order to recover the maximum amount of DNA. The DNA concentration and quality were measured by Thermo Scientific NanoDrop™ 2000 Spectrophotometer, and stored at -20°C. Sanger DNA sequencing was outsourced to StabVida, Portugal, by sending 20-µl of DNA samples at a concentration of 30 ng/µl, and primers in a concentration of 10 pmol/µl in a final volume of 10 µl. DNA sequences were analyzed using various software packages, as are example MacVector, UCSC Genome Browser, BLAST, among others.

2.8 quantitative Polymerase Chain Reaction (qPCR)

We used quantitative Polymerase Chain Reaction (qPCR) analyses to quantify gene expression levels. The qPCR experiments were performed using Lightcycler® 96 (Roche) using the FastStart Essencial DNA Green Master dye and polymerase (Roche). The final volume for each reaction was 10 µl, consisting of 5µl of dye and polymerase (master mix), 2 µl of cDNA sample, and 3 µl of the specific primer pairs (1 µM/µl).

The conditions for qPCR primer selection were established accordingly to the following parameters: 1- PCR Amplification product size should be between 100-200 bp; 2- Primer size between 20-25 bp; 3- Avoid whenever possible G-C nucleotides at the 3' ends in order to prevent unspecific primer bonds from forming; 4- G-C content of approximately 40-60%; 5- No primer self complementarity; 6- Avoid series of 3 or more stretches of G-C or A-T nucleotides at the end of the primer.

Selected primers were first tested by PCR to verify the production of a product of the expected size by gel electrophoresis. Then, the primer pairs were tested by qPCR using several gDNA dilutions (10 ng, 1 ng and 0.1 ng) to have a qPCR efficiency between 80-100%, which means that in each PCR cycle the number of target amplicons approximately doubles.

Forward and Reverse primers used for qPCR are showed in Table 7. While the first two primer pairs are specific for the interest genes, the *rp49* (also known as *RpL32*) is a frequently used positive/normalization control gene for RT-PCR and qRT-PCR normalization (Ling & Salvaterra, 2011). *Rpl32* is a housekeeping gene that is required for the maintenance of basic cellular functions and is stably expressed in all cells and all stages of *Drosophila* development. As negative controls, we used reactions where cDNA was substituted with dH_2O to reveal primer dimers and RT-reactions, as described above.

Table 7. Primers used in qPCR assays.

Primer Name	Primer sequence		cDNA Fragment Size (bp)
qPCRlgr4aroundins	Fw	TCACCTCGACAGGGACAGGAA	131
	Rv	ACTGCGTGAACGAGGTGGAC	
Lgr4_p14	Fw	TGCAGCGATAAGCAGACACCAT	119
	Rv	GTCCTACGCCTTCTGCTGTTGT	
qPCRlgr3aroundins	Fw	CCATGGACGAGCCTTTGCGAAT	296
	Rv	GAAAGGAGCTGACTCCATCGGT	
Lgr3_p1	Fw	GCTGGGTGCCCATCATCGTTAT	87
	Rv	CAAAGACCACCAACCAGGCGTA	
rp49	Fw	TTGAGAACGCAGGCGACCGT	91
	Rv	CGTCTCCTCCAAGAAGCGCAAG	

The conditions used for all qPCR reactions are described in Table 8. Since the expected products have more than 200 bp, the extension time increased 10 s per 100 bp. The melting curve step was added aiming to identify any non-specific amplification products.

Table 8. qPCR programmable conditions.

Cycles	Cycle step	Temperature	Time
1x	Denaturation	95°C	600 s
45x	Annealing	95°C	10 s
		60°C	10 s
	Extension	72°C	30 s*
1x	Melting Curve	95°C	10 s
		65°C	60 s
		97°C	1 s
1x	Cooling	37°C	30 s

*in the *lgr4* case were 20 s of extension.

qPCR results were analyzed with the Lightcycler 96 manufacturer's software and in custom spreadsheet in Microsoft Office Excel ® v2010. All the Cq values, obtained for each primer pair were normalized with the Cq values obtained for our

positive control gene (*rp49*). In order to calculate the number of copies, of any target product relative to *rp49*, in percentage, we applied the formula: **Percentage relative to *rp49* = $2^{-(\Delta Cq)}$ x 100,** where **ΔCq** represents the difference between the Cq of the target and the arithmetic mean of Cq obtained from three replicates of the positive control (*rp49* gene). The geometric mean ± the standard deviation (SD) of the experimental replicates (usually n = 3) for each condition were used in the graphs.

2.9 Genetic Assays

2.9.1 Viability and Fecundity assays

In this work, we performed different approaches to study the viability and fecundity of the flies (Figure 23). Fecundity assays consist in counting the number eggs delivered by single female flies per hour. On the other hand, the viability assays consist in counting the number of fertilized eggs by female per hour, that actually are able to hatch into larvae.

For that propose, we established a few guidelines that were followed throughout this work, such as the number of flies (5 females and 5 males), the timeline of observation and the conditions where the flies were maintained during the experiment (Figure 23). To control the relationship between age and viability, all virgin flies used were 3-7 days old.

In order to study fecundity, female and male flies from the genotypes in study were crossed and were kept in laying pots for three periods of two hours. In the course of the experiment, the 10 flies were maintained for 3 days in a laying pot and the protocol was repeated in the same conditions every single day giving a total of n=9 2-h egg layings per genotype. At the end of each day the eggs were transferred into a strip of paper, to a new vial and counted, in order to follow their development. After 2 days, the number of eggs that hatched were scored (Figure 23). Every female fly that died or escaped was discounted for in the egg/female ratio. All experiments were performed protected from the light. The strip of paper was necessary to

improve the transfer of the egg from the plate to the vial. Its influence in the eggs hatching rate was previously studied in our lab and no change was observed.

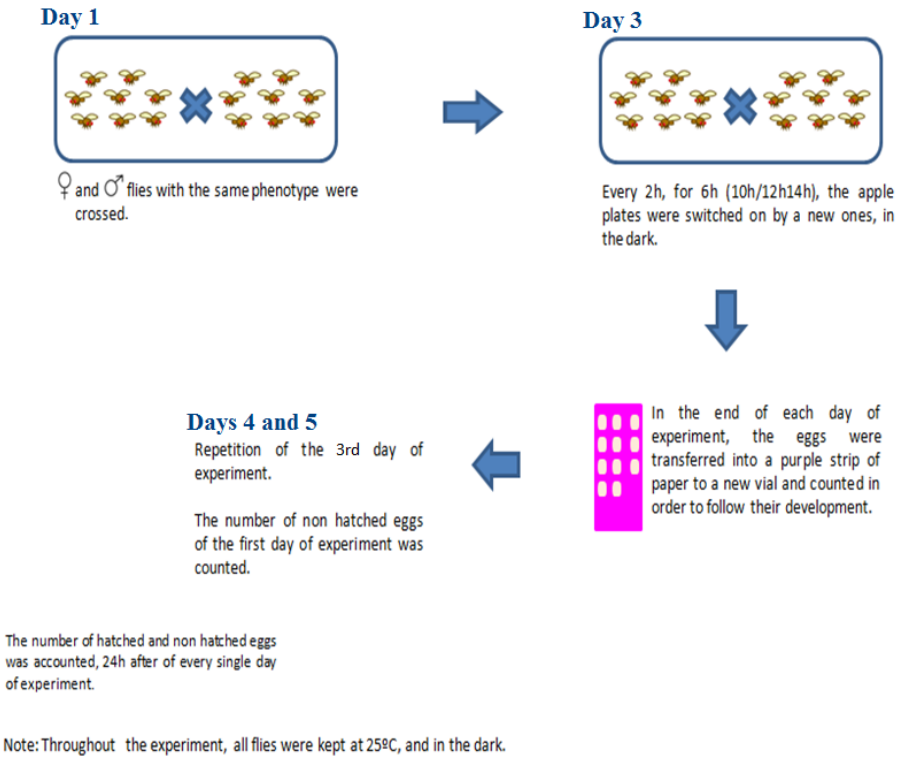


Figure 23. Scheme of fecundity and viability assay.

2.9.2 Developmental time assay

To verify the influence of different *lgr3* or *lgr4* genotypes on the *dilp8*-dependent delay in the onset of metamorphosis, we performed pupariation assays. In these assays, we followed the larval development of synchronously growing larvae until their pupariation (Figure 24).

To induce the *dilp8*-dependent delay in the onset of metamorphosis we used a stock carrying a *tub-dilp8* transgene, where *dilp8* is produced constitutively in every cell under the control of the *tubulin* promoter. In normal conditions, this line already shows a delay. For the assays, the flies of the desired genotypes were crossed and maintained in laying pots with apple juice agar plates, with a small amount of yeast-sucrose paste in the borders, for 48 h (Figure 24). Then, the apple plates were exchanged for new ones, where the flies were able to lay eggs for 4 h. This experiment was performed in two days, being the plates switched two times, of 4 h each. Third instar larvae (48 hours after egg laying [AEL]) were transferred into vials

with 5 ml of standard *Drosophila* food. The number of larvae transferred depended on the experiment and ranged from 10-20 in this work. Usually we favored more experimental repeats with 10 larvae rather than less repeats with more larvae. Survey of pupae consist in counting the number of pupae in each time interval. This was done in intervals of 6 and 12 hours (usually at 9 am, 15 pm, and 21 pm). Time "0" was considered as the day when the eggs were layed (Garelli *et al.*, 2012).

The data generated was analyzed in two different forms. First, the average ± standard error of the mean (SEM) of the % of pupae in each time point was obtained and plotted through time to obtain pupariation curves for the each genotype analyzed. Statistical analyses of the data was performed by calculating the median pupariation time for each experimental repeat. One-way Analyses of Variance (One-Way ANOVA) was performed using the open software package SOFA statistics (http://www.sofastatistics.com/), followed by Tukey's HSD test using a custom Excel spreadsheet and the studentized range statistic (q)* critical values obtained at http://web.mst.edu/~psyworld/virtualstat/tukeys/criticaltable.html. Statistical significance was set at alpha = 5%.

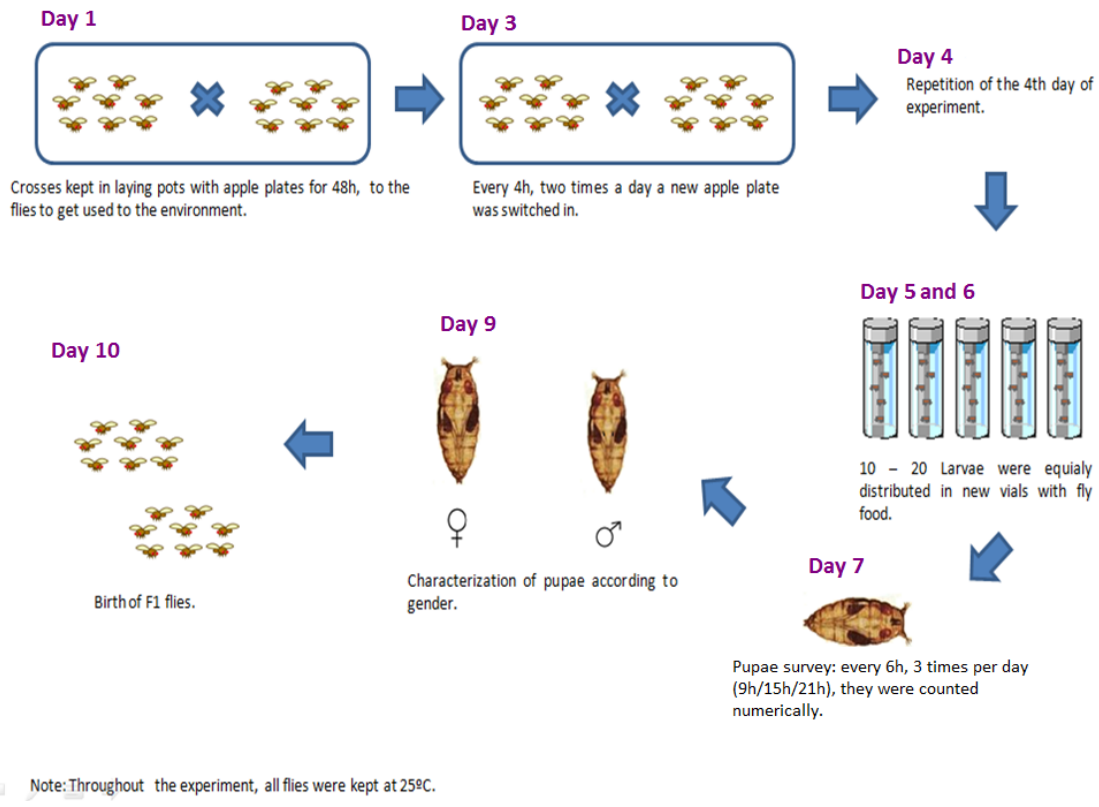


Figure 24. Scheme of development time assay

Chapter 3. Results

3.1 Generation of relaxin-like receptor mutants

3.1.1 *Lgr3* mutants

Initially, we proceeded to the characterization of the *lgr3[MB06848]* element, by measuring the mRNA expression levels by qRT-PCR of both male and female adult flies in different backgrounds. The analysis showed that *lgr3* levels were very low in comparison to the levels of the housekeeping gene *rp49,* but they were slightly higher in males than in females (Figure 25), corroborating a recent report (Hiel *et al.*, 2014). Importantly, the *lgr3* mRNA levels did not seem to be affected by the presence of the *MB06848* element in the *lgr3* locus. Hence, these data suggest that the stocks carrying the *lgr3[MB06848]* insertion *per se* could not be used as a reliable loss-of-function mutation to study *lgr3* function.

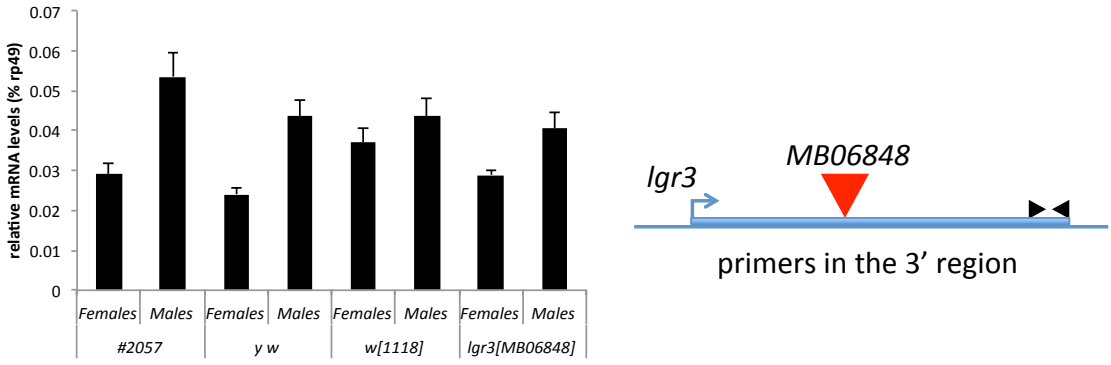


Figure 25. *lgr3* expression levels, quantified by qRT-PCR using primer pairs located in the 3' region of the *lgr3* locus. For these experiments, mRNA was isolated from male and female adult flies from different genetic backgrounds. Results showed that *lgr3* levels were very low in comparison to the levels of the housekeeping gene *rp49* in different backgrounds, but they were slightly higher in males than in females. Shown are the averages of three repeats. Error bars are standard deviations of the means (SDs).

In order to certify ourselves that we had a reliable *lgr3* loss-of-function allele, we remobilized the *lgr3[MB06848]* element to generate a deletion allele. We obtained one imprecise excision mutation, named *lgr3[ex1]*, and three precise excisions, named as *lgr3[exp1], lgr3[exp2]* and *lgr3[exp3]*. These precise excisions serve as genetic background controls for the imprecise excision *lgr3[ex1]*. While *lgr3[ex1]* females were as fecund as controls (Figure 26-A), males from both *lgr3[ex1]* and the control *lgr3[exp1]* stocks were semi-sterile (males did not sire many progeny) when crossed to wild-type *w[1118]* virgin females (Figure 26-B).

Hence these stocks were kept balanced with the TM6B 3rd chromosome balancer. The fact that both *ex1* and control *exp* alleles are male semi-sterile suggests that the male semi-sterility is unrelated to *lgr3*. Apart from this male semi-sterility, *lgr3[ex1]* had no other clear phenotype on reproduction.

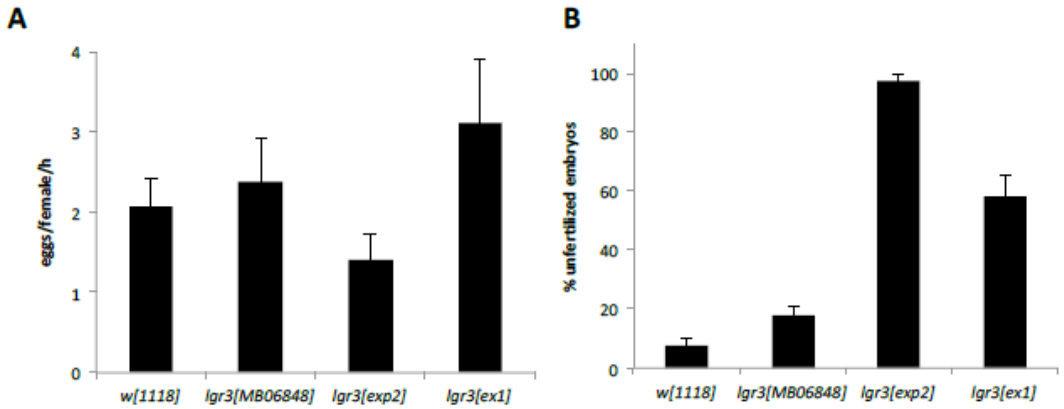


Figure 26. Fecundity (A) and Viability (B) analyses of *lgr3* mutants. *lgr3[ex1]* shows no large differences in fecundity [eggs laid per female per hour (eggs/female/h)] in comparison with the control *lgr3[exp2]* stock. The viability (% of unfertilized embryos) of the embryos laid by the females is nevertheless strongly affected in both *lgr3[ex1]* and *lgr3[exp2],* in comparison with the other stocks. Shown are the averages +/- the standard error of the means (SEMs) of n=9 for each genotype.

To molecularly characterize the *lgr3[ex1]* deletion, we performed a series of PCR assays with primer pairs located around the *lgr3[MB06848]* insertion (Figure 27-A). *lgr3[ex1]* was initially detected by the failure to amplify a 866-bp PCR product flanking the *lgr3[MB06848]* insertion and then a 240-bp PCR product using a pair of primers on exon 7 and 8 to the right of the *lgr3[MB06848]* element (Figure 27-B), indicating that a deletion occurred downstream of the element following its remobilization. A PCR using a primer pair in the exon 10 produced a positive result, indicating that exon 10 was present in *lgr3[ex1]* (not shown). This suggested that while the *lgr3[ex1]* deletion was very large, its breakpoints were confined within the *lgr3* locus. We then tried to amplify a 5.3-kb PCR product with a primer upstream of the *lgr3[MB06848]* position and another on exon 10 (Figure 27-A). Instead of the predicted >5.3 kb product, we obtained a ~1.6 kb product (Figure 27-C), suggesting that *lgr3[ex1]* is indeed a large deletion of approximately 3.8 kb in the *lgr3* locus (Figure 27-D).

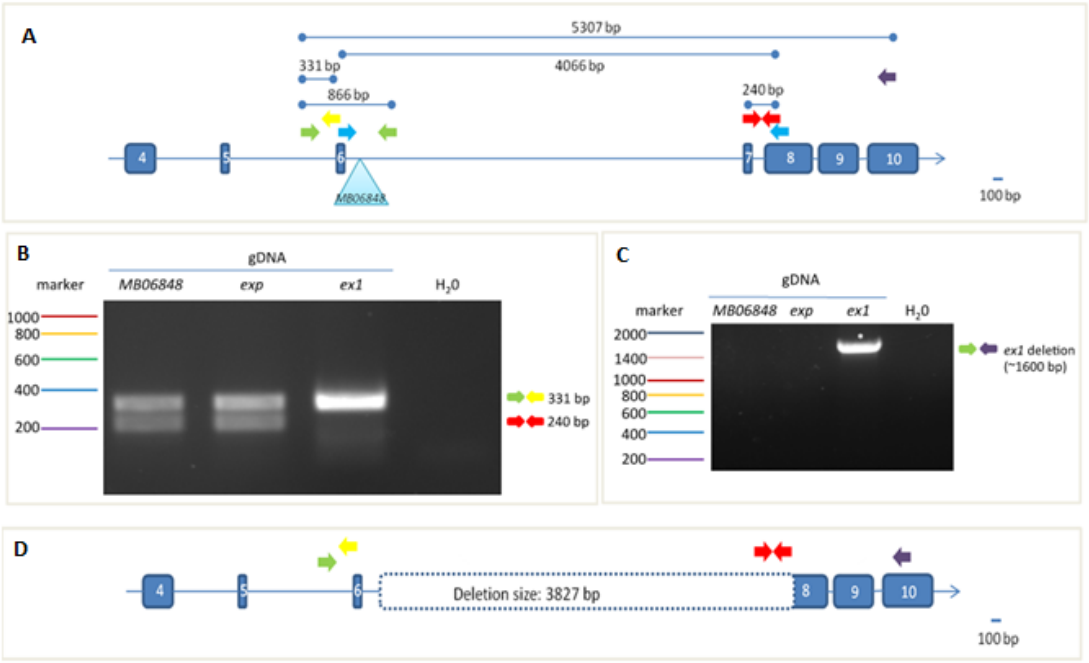
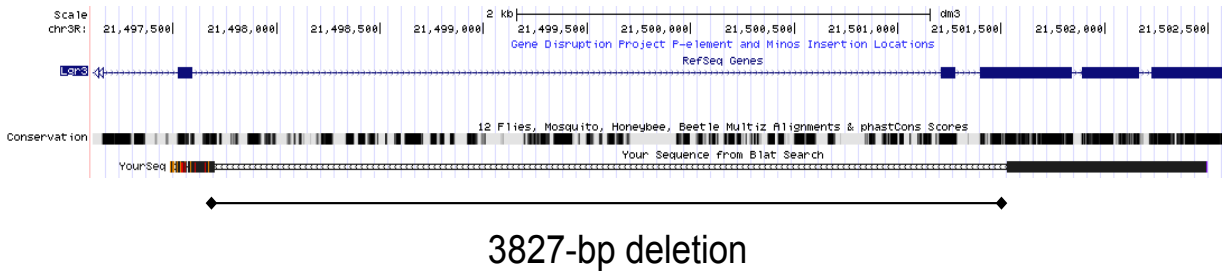


Figure 27. *Molecular characterization of the lgr3[MB06848] remobilization lines*. (A) Scheme with key sets of primers used, as well as the insertion site. In this scheme, the *lgr3* locus transcribes from left to right. Exons are numbered. Colored arrows show the approximate location of the primers used. The blue set of primers was used to identify the deletion breakpoints. (B) Electrophoretic pattern of PCR amplification products of the precise and imprecise excision *lgr3[exp2]* and *lgr3[ex1],* respectively, which shows that the deletion occurred downstream (to the right) of the insertion site. (C) Electrophoretic band of PCR amplification product of *lgr3[ex1]*, from which we later obtained the gDNA sequence and identified a deletion of approximately 1.6 kb. (D) Scheme of the *lgr3[ex1]* deletion. The deletion removes most of the intron and completely removes exon seven and partially exon eight. The negative control used in the different experiments was dH_2O, showing no PCR amplifications, as expected.

Sequencing of the 1.6-kb PCR product confirmed that *lgr3[ex1]* is a 3.8-kb deletion that completely removes the coding exon 7 and partially removes coding exon 8 of *lgr3*, as can be seen in (Figure 27-D/Figure 28).



3827-bp deletion

Figure 28. UCSC genome browser screenshot of the *D. melanogaster lgr3* locus with BLAT results (Your Sequence from Blat Search) obtained by sequencing the 1.6-kb PCR fragment obtained in Figure 29-C. This shows the breakpoints that characterize the 3827-bp deletion in *lgr3[ex1]* mutant.

Next, we performed RT-PCR analyses with mRNA isolated from *lgr3[ex1]* and a series of controls (*lgr3[exp2]* and *lgr3[MB06848]*). *lgr3[ex1]* generated a smear with a major band that was ~200 bp smaller than the other control genotypes, consistent with a mRNA that reads directly from exon 6 into exon 8 without splicing out the remaining intron sequences (Figure 29-A).

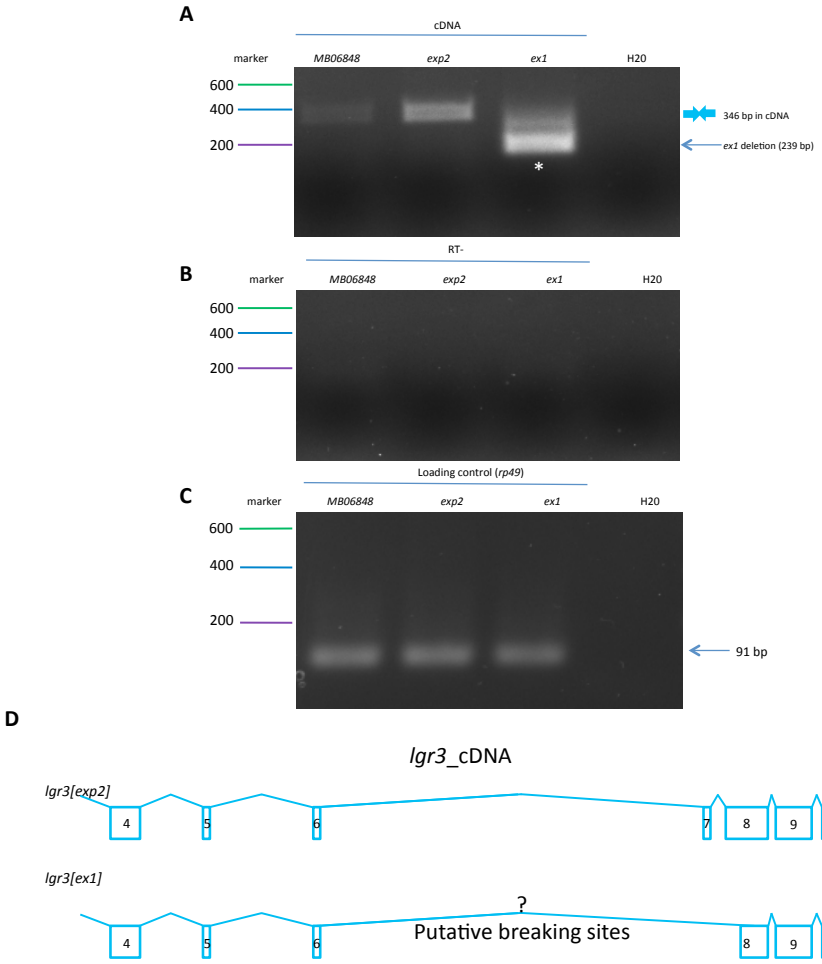


Figure 29. *RT-PCR amplification pattern of lgr3 mutations and controls using mRNA isolated from males.* (A) Electrophoretic pattern of PCR amplification products of *lgr3[exp2]* and *lgr3[ex1]*, with the set of primers *lgr3_around* (set of primers in blue). The *lgr3[ex1]* mutant gave a smear and a major product at around 200 bp, consistent with a non-spliced transcript reading directly through the deletion breakpoint. (B) RT negative control, showing no amplification products, ensuring that any gDNA contamination, if present, was under detection levels. (C) The housekeeping gene *rp49* was used as a positive control for the RT-PCR. The negative control used was dH_2O. (D) Schematic representation of the *lgr3* cDNA in the *lgr3[ex1]* mutant and in the *lgr3[exp2]* control line. Blue, normal splicing and exon reading frames. Red, aberrant splicing and frame shifted exons. We were unable to sequence a specific product from the *lgr3[ex1]* transcript.

To further confirm our predictions that *lgr3[ex1]* produced an aberrant transcript and to certify ourselves that the background control precise excision

lgr3[exp2] allele restored transcripts level to wild-type levels, we also performed qRT-PCR using cDNA primer pairs located in the exons flanking the *lgr3[MB06848]* element (Figure 27-A, blue set of primers). The qRT-PCR results confirmed that *lgr3[ex1]* produced an aberrant transcript, which was actually more abundant than the correctly spliced *lgr3* transcript found in *lgr3[MB06848], lgr3[exp2], and w[1118],* a stock that was used as a further background control for the *lgr3[MB06848]* insertion (Figure 30). We were unable to obtain a sequence for the aberrant transcripts arising from *lgr3[ex1]*, but the prediction is that there is little or no splicing from exon 6 to exon 8. Rather the size of the major transcript is consistent with direct read-through through the intron directly into exon 8.

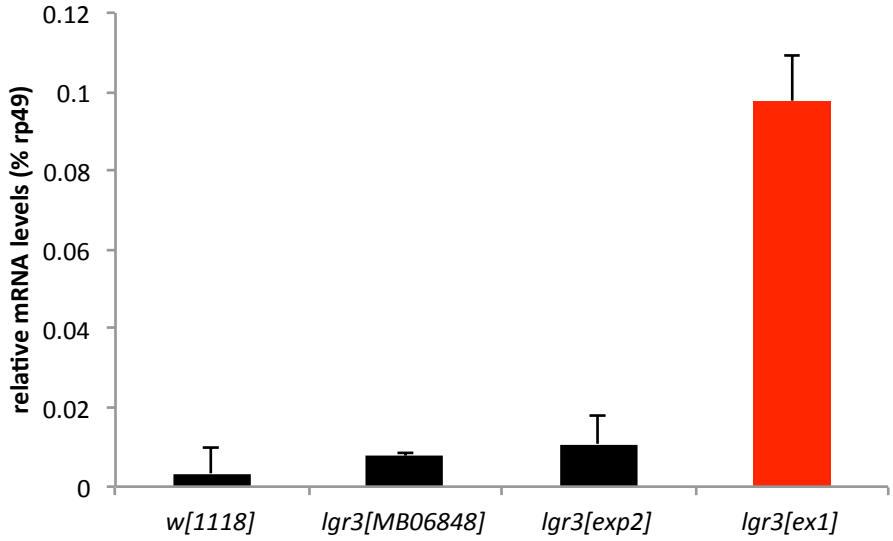


Figure 30. Transcript levels of the mutant *lgr3[ex1]* mutant and controls. An aberrant transcript can be detected in *lgr3[ex1]* which is present at higher levels than that of the other stocks. This could be due to the lack of the large intron deleted in this mutation. The correctly spliced *lgr3* transcript found in *lgr3[MB06848], lgr3[exp2], and w1118* are less abundant. Only male cDNAs were analyzed. Shown are the averages of three repeats +/- SDs.

By comparing the protein product that would be predicted to form in *lgr3[ex1]* to the reference wild-type Lgr3 protein sequence NP_733115.1, we determined that *ex1* would delete at least 69 aminoacids (aa) of the Lgr3 protein from aa I327 to T395. This removes essential aa in the C-terminal part of the Leucine Rich Repeat (LRR) domain, which is associated with ligand binding in the vertebrate relaxin receptors (Figure 31; reviewed in Bathgate *et al.*, 2013). Hence, even if a productive protein product is still made in *lgr3[ex1]* it would be a receptor that could probably not bind a ligand via the LRR domain. However, due to intron read-through of the

transcript following exon 6, the major mRNA produced by *lgr3[ex1]* probably truncates the Lgr3 protein soon after ~aa D326 (Figure 31), yielding a severely truncated protein without the 7TM domains. Hence, *lgr3[ex1]* is predicted to be a strong loss-of-function allele (Figure 31).

WT Lgr3 protein sequence NP_733115.1

```
  1 mvygrsiavg fclmtvvlll aavifylslg pcpaasfacd ngtlcvprrq mcdsrndcad
 61 ssdenpvecg llygskeiad kivrnaiekk qgrlisavsn asgadsttsm vprnqsltln
121 mtcdivtypk acqcgqgtil ycgryaklrr fprlssevtn liiirnnltl rdnifanftr
181 lqkltlkynn isrvplgsfs glfhlerlel shnnvshlph gvflglhslq wlflvnnhlh
241 hlpveqlrff rrlewlvlsr nrltlrnvql pkiptlyevy ldfnrieyig eetfsqldnl
301 hlldlqhnli thihgrafan ltnmrdirlv gnpikelsge tflhntrlea lslalmpihi
361 ssslmeplni sflnltgiry dhidfeains mrnltyiiyd rffycsmtpr vrmckpstdg
421 vssfqdllsk pvlrysawvm atltiagnvl vlwgrfiyrd envavtmvir nlaladmlmg
481 fylvtigvqd yryrneyykv vldwitswqc tligtlavss sevsmlilaf mslerfllia
541 dpfrghrsig nrvmwlalic iwitgvglav apvllwrtst lpyygsysgt cfplhiheaf
601 pmgwlysafv flgvnllllv miamlytall isiwrtrsat pltlldcefa vrfffivltd
661 flcwvpiivm kiwvffnyni sddiyawlvv fvlplnsavn pllytfttpk yrnqiflrgw
721 kkitsrkrae agngnvattt tgtatgssqh pddftifaka amrch
```

Predicted Lgr3[ex1] protein sequence

```
  1 mvygrsiavg fclmtvvlll aavifylslg pcpaasfacd ngtlcvprrq mcdsrndcad
 61 ssdenpvecg llygskeiad kivrnaiekk qgrlisavsn asgadsttsm vprnqsltln
121 mtcdivtypk acqcgqgtil ycgryaklrr fprlssevtn liiirnnltl rdnifanftr
181 lqkltlkynn isrvplgsfs glfhlerlel shnnvshlph gvflglhslq wlflvnnhlh
241 hlpveqlrff rrlewlvlsr nrltlrnvql pkiptlyevy ldfnrieyig eetfsqldnl
301 hlldlqhnli thihgrafan ltnmrdirlv gnpikelsge tflhntrlea lslalmpihi
361 ssslmeplni sflnltgiry dhidfeains mrnltyiiyd rffycsmtpr vrmckpstdg
421 vssfqdllsk pvlrysawvm atltiagnvl vlwgrfiyrd envavtmvir nlaladmlmg
481 fylvtigvqd yryrneyykv vldwitswqc tligtlavss sevsmlilaf mslerfllia
541 dpfrghrsig nrvmwlalic iwitgvglav apvllwrtst lpyygsysgt cfplhiheaf
601 pmgwlysafv flgvnllllv miamlytall isiwrtrsat pltlldcefa vrfffivltd
661 flcwvpiivm kiwvffnyni sddiyawlvv fvlplnsavn pllytfttpk yrnqiflrgw
721 kkitsrkrae agngnvattt tgtatgssqh pddftifaka amrch
```

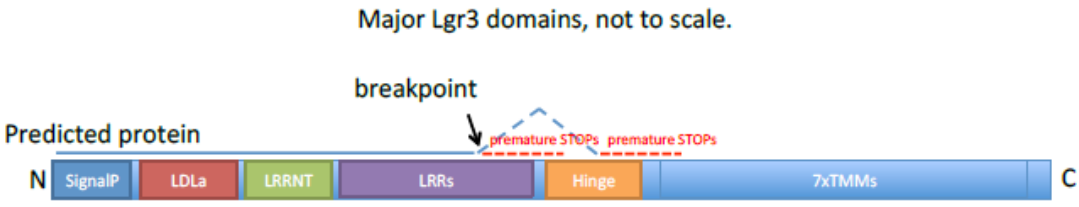


Figure 31. Protein sequence of WT Lgr3 (Reference NP_733115.1) and the predicted Lgr3[ex1] protein. The 69 aa from I327 to T395 in the C-terminal part of the Leucine Rich Repeat (LRR) domain that were certainly removed by the *lgr3[ex1]* deletion are highlighted in red. The bold and underlined aa (D236) marks the end of exon 6, after which the protein most likely soon truncates due to intron readthrough of the transcript. The major domains of the Lgr3 protein are the signal peptide, LDLa domain, LRR N-terminal domain (LRRNT), LRRs, the hinge region (Hinge) and the 7TM domains. This cartoon is not to scale. It depicts the approximate region of the Lgr3[ex1] protein truncation. There could be other protein isoforms if splicing still occurs downstream of the breakpoint, but these are also predicted to encode premature stop codons and hence truncate the protein.

3.1.2 *lgr4* mutants

To start characterizing the *lgr4[MB03440]* element, we measured *lgr4* mRNA expression levels of both male and female adult flies in different backgrounds. The analysis showed that *lgr4* levels were very low in comparison to the levels of the housekeeping gene *rp49,* but they were higher in males than in females (Figure 32), corroborating a recent report (Hiel *et al.*, 2014). Importantly, the *lgr4* mRNA levels did not seem to be strongly affected by the presence of the *MB03440* element in the *lgr3* locus in females and the effect in males was still within the levels of controls. Hence, these data suggest that the stocks carrying *lgr4[MB03440]* insertion *per se* could not be used as a reliable loss-of-function mutation to study *lgr4* function.

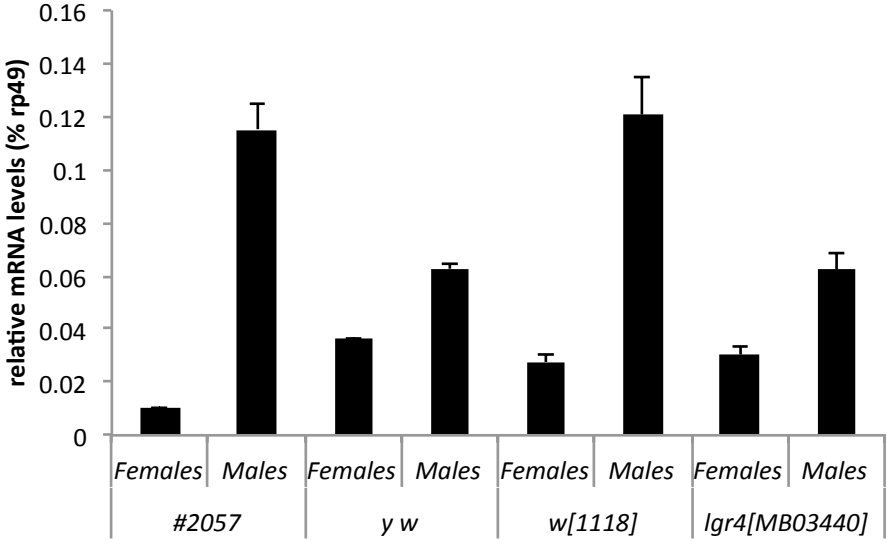


Figure 32. *lgr4* expression levels, quantified by qRT-PCR. For these experiments, mRNA was isolated from male and female adult flies from different genetic backgrounds. Results showed that *lgr4* levels were very low in comparison to the levels of the housekeeping gene *rp49* in different backgrounds, but they were higher in males than in females. Shown are the averages of three repeats +/- SDs.

To generate reliable loss-of-function alleles, we remobilized the *MB03440* element and obtained a total of seven *lgr4* imprecise excisions, named *lgr4[ex1] to [ex7]*. We have successfully characterized at the molecular level two imprecise excisions, *lgr4[ex1]* and *lgr4[ex3]*. We were unable up to this moment to precisely map the breakpoints of the other deletion mutations possibly because some are complex deletions retaining large parts of the *MB03440* element and/or contain inversions. In parallel, we also generated three precise excisions, named *lgr4[exp1], lgr4[exp2]* and *lgr4[exp3]*. All *lgr4* alleles were homozygous viable and fertile, with

no clear phenotype. Indeed, *lgr4* mutation does not seem to affect fecundity or fertility levels (Figure 33).

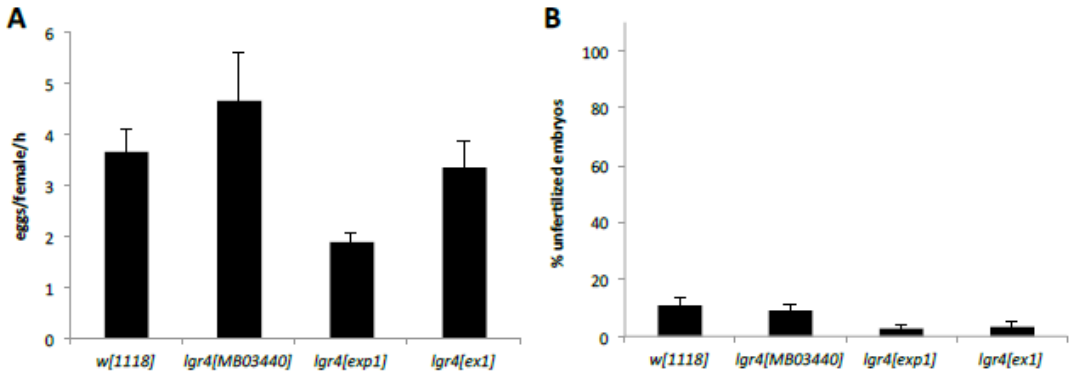


Figure 33. Fecundity (A) and Viability (B) analyses of *lgr4* mutants and controls. (A) *lgr4[ex1]* lay equivalent amounts of eggs as other control lines [eggs laid per female per hour (eggs/female/h)]. (B) The viability (% of unfertilized embryos) of the embryos laid by the females is not negatively affected in *lgr4[ex1]* and *lgr4[exp1],* in comparison with the other stocks. Shown are the averages and the standard error of the means (SEMs) of n=9 for each genotype.

lgr4[ex1] was detected by a faster migrating band than wild-type in a PCR-based screen for deletions following the remobilization of the *MB03440* element (using the green set of primers in Figure 34-A; see also Figure 34-B for gel image). Sequencing of the *lgr4[ex1]* PCR band showed that it is a 159-bp deletion (Figure 34-B-C and Figure 35) that removes part of exon 2 and intron 2 (see sequences in Figure 35). According to the sequence obtained, the deletion causes the formation of a truncated 137-aa protein containing the very N-terminus of Lgr4 up to the LDLa domain (Figure 36).

The molecular nature of *lgr4[ex3]* was much harder to determine. It was detected when primers flanking the remobilized *MB03440* element gave no product (purple set of primers in Figure 34-A). This can occur if one or both of the primer pair sequences have been deleted from the genome, or if part or all of the *MB03440* element has remained in the locus, despite the loss of the internal eGFP expression cassette (loss of eGFP expression was an essential criteria in the screen; see Materials & Methods section). To determine if some *MB03440* element sequences remained in *lgr4[ex3]* we designed primer pairs across the *MB03440* element. The PCR results showed that *lgr4[ex3]* retained a small part of the *MB03440* element of at least 2 kb (in the portion 5' of the insertion , figure 35-D). Using a pair of primers

located in between exon 4 and exon 3 of *lgr4 (*yellow and green primers*)*, we were unable to amplify a 275-bp product in *lgr4[ex3]*, indicating that apart from the loss of almost all the *MB03440* element, a deletion occurred downstream of the element following its "partial" remobilization (Figure 34-D). Attempts to produce a PCR product using gDNA and primer pairs flanking the deletion were unsuccessful, possibly because of the additional *MB03440* element sequence remaining. Further insight into *lgr4[ex3]* came from cDNA analyses (see below).

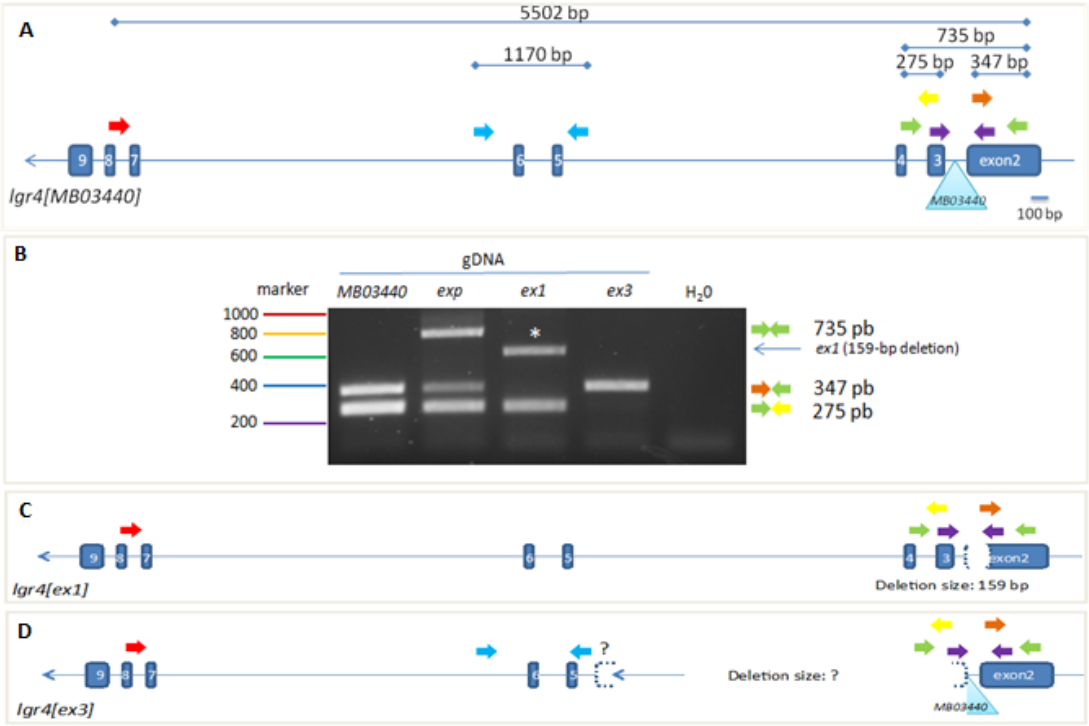


Figure 34. *Molecular characterization of the lgr4[MB03440] remobilization lines*. (A) Scheme with the key sets of primers used, as well as the insertion site. In this scheme, the *lgr4* locus transcribes from right to the left. Exons are numbered. Colored arrows are show the approximate location of the primers used. (B) Electrophoretic pattern of PCR amplification products of the *lgr4[exp1]* precise excision and the *lgr4[ex1]* and *lgr4[ex3]* deletions. This shows that the *lgr4[ex1]* deletion occurred upstream of the insertion site while the *lgr4[ex3]* deletion occurred downstream. This also shows that the *lgr4[exp1]* is indeed a precise excision giving the expected product PCR size with the green primer pair (note that no 735-bp product is produced in the *MB03440* line because of the presence of the several-kb *MB03440* insertion in between the primer sites). The positive control, used to check the deletions, was the original line used in the excision protocol. As expected, the negative control used, dH_2O, showed no PCR amplification. (C) Gene scheme depicting the deletion in *lgr4[ex1],* showing that exon 2 was partially deleted. (D) Gene scheme depicting the deletion in *lgr4[ex3],* showing that the deletion affects introns two, three and four and the exons three and four. Notice that a chunk of the *MB04440* elements still remains in this allele.

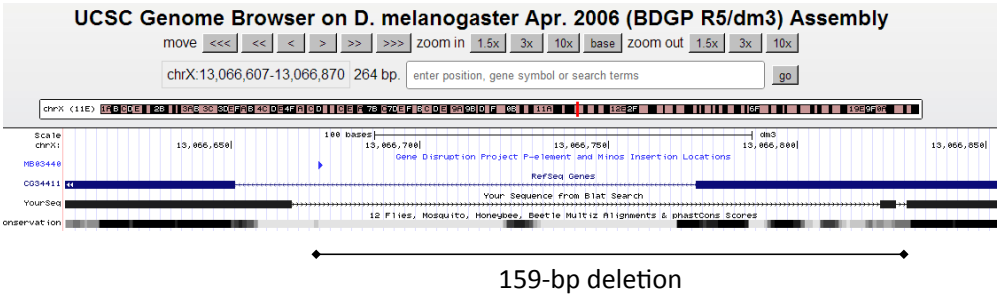


Figure 35. UCSC genome browser screenshot of part of the *D. melanogaster lgr4* locus containing BLAT results (Your Sequence from Blat Search) obtained by sequencing the PCR product with a set of primers upstream and downstream of the *lgr4[MB03440]* position (blue arrowhead). Instead of the predicted 255-bp product, we obtained a product of 96 bp, confirming the *ex1* mutation. Notice that the *lgr4* locus transcribes from the right to the left in this scheme. Hence, the exon to the right is exon 2, and the exon to the left is exon 3. Some of the sequences adjacent to the deletion missalign in exon 2 rather than aligning in the intron. Actually the deletion starts immediately to the right of the *MB03440* element.

Lgr4 protein sequence NP_001096966.2.

```
  1 mciahlpitf tlaillaias negaqgvesa trtaieairt gigtkpetei adateaeapv
 61 revisllgii dgaesdilvp daddkcpggy fhcnttaqcv pqrancdgsv dcddasdevn
121 cvnevdakyw dhlyrkqpfg rhdnlrigec lwpnenfscp crgdeilcrf qqltdiperl
181 pqhdlatldl tgnnfetihe tffselpdvd slvlkfcsir eiashafdrl adnplrtlym
241 ddnklphlpe hffpegnqls ililarnhlh hlkrsdflnl qklqeldlrg nrignfeaev
301 farlpnlevl ylnenhlkrl dpdrfprtll nlhtlslayn qiediaantf pfprlrylfl
361 agnrlshird etfcnlsnlq glhlnenrie gfdleafacl knlssllltg nrfqtldsrv
421 lknltsldyi yfswfhlcsa amnvrvcdph gdgissklhl ldnqilrgsv wvmasiavvg
481 nllvllgryf yksrsnvehs lylrhlaasd flmgiyltli acadisfrge yikyeetwrh
541 sgvcafagfl stfscqsstl lltlvtwdrl msvtrplkpr dtekvrivlr llllwgisfg
601 laaapllpnp yfgshfygnn gvclslhihd pyakgweysa llfilvntls lifilfsyir
661 mlqairdsgg gmrsthsgre nvvatrfaii vttdcacwlp iivvklaals gceispdlya
721 wlavlvlpvn salnpvlytl ttaafkqqlr rychtlpscs lvnnetrsqt qtayesglsv
781 slahlgggvg ggsgrkrmsh rqmsyl
```

Predicted Lgr4 [ex1] protein sequence.

```
  1 mciahlpitf tlaillaias negaqgvesa trtaieairt gigtkpetei adateaeapv
 61 revisllgii dgaesdilvp daddkcpggy fhcnttaqcv pqrancdgsv dcddasdevn
121 cvnevdakyw dqylifi*
```

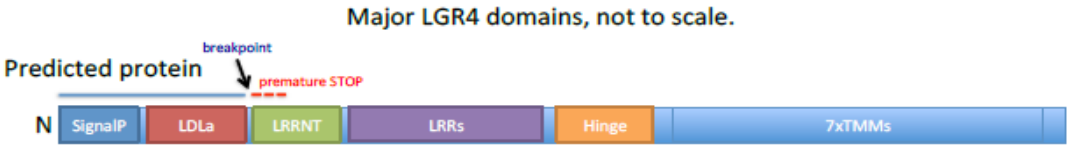


Figure 36 – Protein sequence of WT Lgr4 (Reference NP_001096966.2) and the predicted Lgr4[ex1] protein. According to the sequence obtained, the deletion causes the formation of a truncated 137-aa protein containing the very N-terminus of Lgr4 up to the LDLa domain. In red are the aa that form due to readthrough from exon 2 into the downstream intron. This sequence was confirmed by sequencing the RT-PCR product of this site. The major domains of the Lgr4 protein are the signal peptide, LDLa domain, LRR N-terminal domain (LRRNT), LRRs, the hinge region (Hinge) and the 7TM domains. This cartoon is not to scale. It depicts the approximate region of the Lgr4[ex1] protein truncation. There could be other protein isoforms if splicing still occurs downstream of the breakpoint, but these are also predicted to encode premature stop codons and hence truncate the protein.

Apart from this, by sequencing the *lgr4* locus in the *lgr4[exp]* alleles in parallel, we could also identify four SNPs in the *lgr4* sequence of the iso *w[1118]* background (in which the *MB03440* insertion was generated) compared to the sequenced strain iso-1 (Bloomington stock #2057). One of the SNPs identified is a GAC to CAC (D->A) amino acid change. Interestingly, this looks more like the other *Drosophila spp.* Suggesting that it is the reference iso-1 strain which has an unusual aa in this position. The other SNPs are conservative.

To analyze the consequences of the *lgr4[ex1]* and *lgr4[ex3]* deletion mutations on *lgr4* transcripts, we performed RT-PCR assays with one primer located in exon 2 and the other in exon 8. This way, we could detect a single transcript of the expected size in the control *lgr4[exp1]* and aberrant transcripts in *lgr4[ex1]* and *lgr4[ex3]* (Figure 37). *lgr4[MB03440]* showed the correctly spliced transcript and a smaller transcript, which we have not sequenced. The *lgr4[ex1]* animals showed a ~50-bp smaller transcript than the control *lgr4[exp1]* which is consistent with a transcript that reads through exon 2 into intron 3 and exon 3, with no splicing (Figure 37-A). This prediction was confirmed by sequencing (see also Figure 36).

To further confirm our predictions that *lgr4[ex1]* produced an aberrant transcript and to certify ourselves that the background control precise excision *lgr4[exp1]* allele restored transcripts level to wild-type levels, we also performed quantitative qRT-PCR using cDNA primer pairs located in the exons flanking the *lgr4[MB03440]* element (Figure 38-A). The qRT-PCR results confirmed that *lgr4[ex1]* produced an aberrant transcript (evidenced here by an aberrant product melting curve; Figure 38-B). These results also confirmed that *lgr4[MB03440]* induces a small reduction in *lgr4* levels in males (see also Figure 34), which are restored to normal in *lgr4[exp1]*.

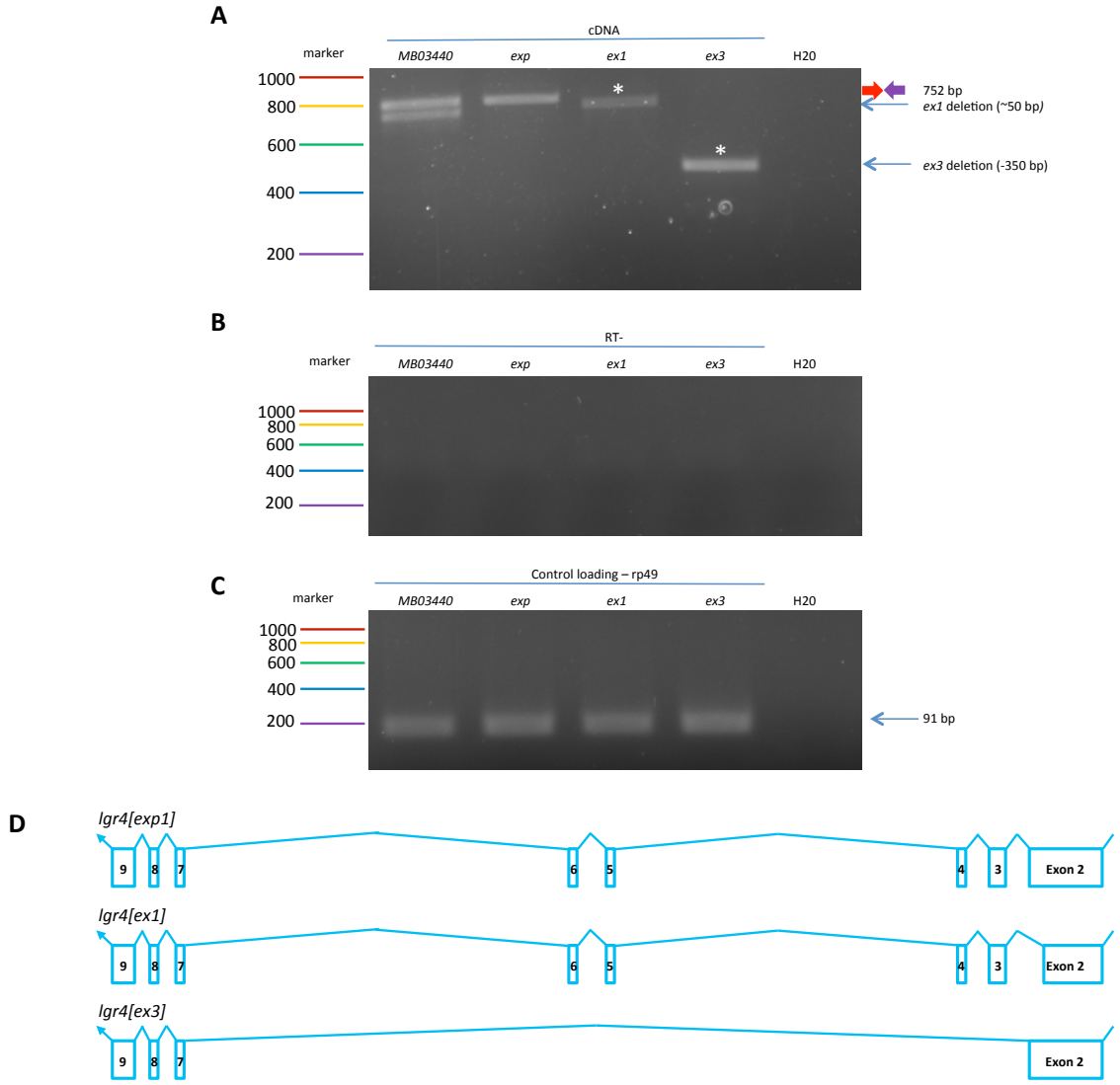


Figure 37. RT-PCR amplification pattern of *lgr4* mutations and controls using mRNA isolated from males. (A) Electrophoretic pattern of RT-PCR amplification products in the depicted genotypes with the set of primers for *lgr4 gene* (set of primers forward in red and reverse in purple). (B) RT negative control, showing no amplification, as expected, ensuring that no gDNA contamination occurred. (C) Positive control, the housekeeping gene *RP49* expression. The negative control used was dH_2O, showing no PCR amplification, as expected. (D) Scheme of the *lgr4* transcripts in the mutants and in the control *exp1* line. Notice that in this scheme the *lgr4* locus is transcribed from the right to the left.

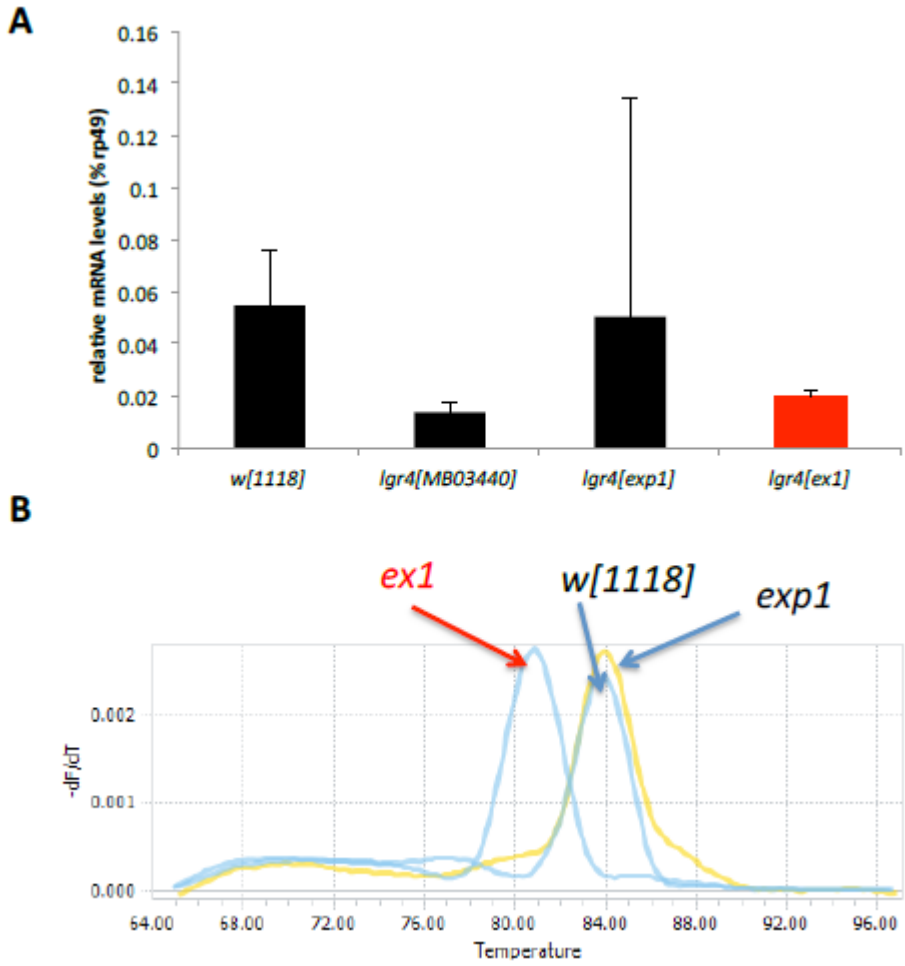


Figure 38. qRT-PCR analyses of transcript levels of the *lgr4[ex1]* deletion mutant in comparison with the control lines, *lgr4[exp1] and w[1118],* and the original insertion line *lgr4[MB03440]*. Primer pairs in exonic sequences flanking the *ex1* deletion were used and normalized to the housekeeping gene *rp49*. A transcript was detected in *lgr4[ex1],* which was nevertheless aberrant and had a distinct dissociation curve (B). One can also observe that the *lgr4[MB03440]* insertion induces a small reduction in *lgr4* levels, as previously detected (see Figure 34), which is restored to normal in *lgr4[exp1]*. Only males cDNA were analyzed. Shown are the averages of three repeats, except for *lgr4[exp1]*. Error bars are standard deviations of the means. (B) The dissociation curves of selected qRT-PCR products shown in (A) for the three depicted genotypes. These results confirm that *lgr4[ex1]* produced an aberrant transcript, different from the *w[1118]* and *lgr4[exp1]* controls.

The *lgr4[ex3]* transcript was ~350 bp smaller than the control *lgr4[exp1]* transcript, consistent with a large deletion (Figure 34–D). Sequencing of the *lgr4[ex3]* cDNA product revealed a transcript that spliced directly from exon 2 to exon 7, skipping 350 bp of *lgr4* mRNA (Figure 34–D). Hence, this confirms that at

least the exons 3 and 4, and likely exons 5 and 6, are deleted from *lgr4[ex3]* (Figure 34–D). Splicing from exon 2 to exon 7 leads to truncated proteins right after the LDLa domain (Figure 39). This splicing causes a frameshift mutation and a stop codon to appear 10 aa into exon 7. The last aa of the wild type Lgr4 sequence is probably C153. Therefore no LRRNT, LRR, Hinge, or 7TM domains are produced in *lgr3[ex3]*. These data indicate that *lgr4[ex3]* is a major loss of function allele.

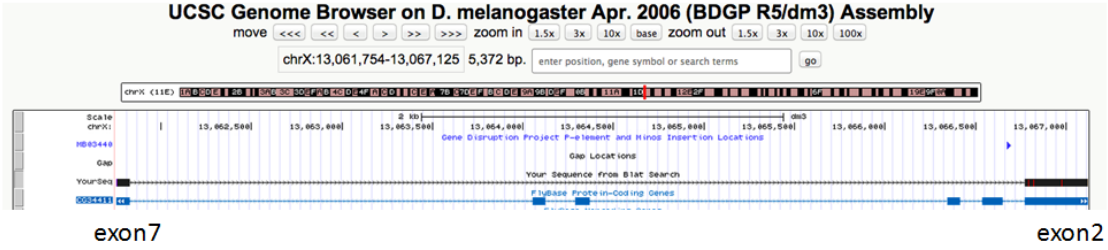


Figure 39. *lgr4[ex3]* cDNA sequence. UCSC genome browser screenshot of part of the *D. melanogaster lgr4* locus containing BLAT results (Your Sequence from Blat Search) obtained by sequencing the RT-PCR product obtained in Figure 39-A for *lgr4[ex3]* with a set of primers in exon2 and in exon7. The *lgr4[MB03440]* position is depicted with a blue arrowhead. Notice that the *lgr4* locus transcribes from the right to the left in this scheme. The *lgr4[ex3]* transcript splices perfectly from exon two to exon seven, consistent with the prediction of the deletion of all introns two, three and four, and all exons three and four. Exons 5 and 6 were also probably deleted.

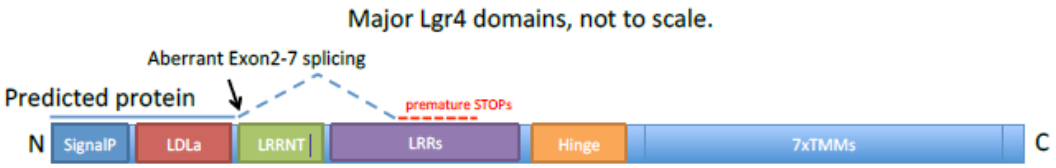


Figure 40. Protein sequence of WT Lgr4 (Reference NP_001096966.2) and the predicted Lgr4[ex3] protein. According to the sequence obtained, the deletion causes the formation of a truncated ~163-aa protein (153 aa of Lgr4 sequence and 10 aa of framshifted sequence from exon 7) containing the very N-terminus of Lgr4 up to the LDLa domain. In red are the aa that form due to out-of-frame splicing from exon 2 into the exon 7. The major domains of the Lgr4 protein are the signal peptide, LDLa domain, LRR N-terminal domain (LRRNT), LRRs, the hinge region (Hinge) and the 7TM domains. This cartoon is not to scale. It depicts the approximate region of the Lgr4[ex3] protein truncation. There could be other protein isoforms if splicing still occurs downstream of exon7, but most of these are also predicted to encode premature stop codons and hence truncate the protein.

3.2 Developmental time assays

We tested the hypothesis that one or both of the *Drosophila* relaxin-like receptors, *lgr3* or *lgr4,* acted in the *dilp8* pathway. *dilp8* functions in a pathway that coordinates growth with developmental timing (Colombani *et al.*, 2012; Garelli *et al.*, 2012). Ectopic expression of *dilp8* under the control of the *tubulin* promoter (*tub*) is sufficient to cause reproducible and robust delay in the onset of metamorphosis (Garelli *et al.*, 2012). If *dilp8* has a receptor, it should be required for this *dilp8*-dependent developmental delay. Hence, our strategy was straightforward: to assay for the timing of the onset of metamorphsosis (pupariation) in larvae ectopically expressing *dilp8* in the absence of either *lgr3* or *lgr4*.

To test if *lgr3* is required for the *dilp8*-dependent delay we first removed either one or two *lgr3* copies in a normal (+/+) or developmentally-delayed background (*tub-dilp8/*+). In this assay, we scored eight tubes per genotype, but animals that were heterozygotes for *lgr3[ex1]* (*lgr3[ex1]*/+) appeared in two different crosses, hence the double amount of repeats (n=16) were available for these genotypes: +/+*; lgr3[ex1]*/+ and *tub-dilp8/+; lgr3[ex1]*/+ (Figure 41). Firstly, one may notice that removing one or two copies of *lgr3* has no effect on developmental timing *per se* (compare +/+*; +/+, +/+; lgr3[ex1]*/+ and +/+*; lgr3[ex1]*/*lgr3[ex1]*). In contrast, we found a complete and statistically significant suppression of the *dilp8*-dependent delay when two copies of *lgr3* were removed (*tub-dilp8/+;lgr3[ex1]/lgr3[ex1]*, $p < 0.05$, Tukey's HSD post-hoc Test). Hence, *tub-dilp8/+;lgr3[ex1]/lgr3[ex1]* larvae are indistinguishable from normal (+/+) or *+/+;lgr3[ex1]/lgr3[ex1]* larvae as regards developmental timing. In other words, *dilp8* ectopic expression has no effect on larvae that do not have *lgr3*. Interestingly, the *dilp8*-dependent delay was consistently partially suppressed when one copy of *lgr3* was removed in *tub-dilp8/+; lgr3[ex1]*/+ larvae, although this effect was not statistically significant (Figure 41).

Taken together, this is a very interesting finding, which suggests that *lgr3* plays an important role in the *dilp8* pathway, and that it is required for *dilp8* activity in larvae. Hence, *lgr3* is a good candidate for being the *dilp8* receptor.

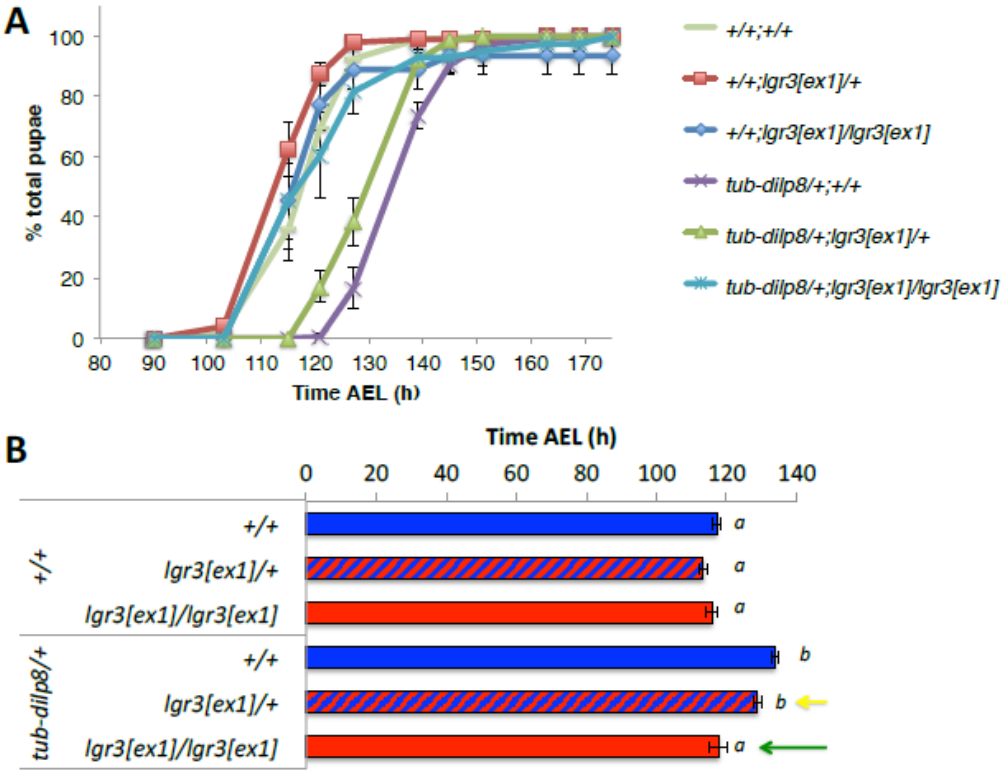


Figure 41. *lgr3* mutation rescues the *dilp8*-dependent delay in the onset of pupariation. (A) Pupariation timing curves for different genotypes. Shown are averages +/- SEMs of 8-16 experiments. (B) Bar graph showing the median pupariation time of the depicted genotypes. ANOVA: $p<0.001$. Groups sharing the same letters are not statistically significantly different at alpha=5% according to the Tukey's HSD post-hoc test. Shown are averages +/- SEMs of 8-16 experiments. Bar colors depict the dosage of *lgr3* (blue = two wild-type alleles, red = two loss-of-function alleles; hashed bars = heterozygotes). Green arrow depicts full suppression. Yellow arrow depicts partial suppression.

In the second set of experiments we used a chromosome deficiency (Df) line, *Df(3)BSC321*, where at least 60 Kb of chromosome III are deleted, removing the *lgr3* gene completely as well as many other genes (Cook *et al.,* 2012). By placing the *lgr3[ex1]* mutation over *Df(3)BSC321*, we can formally test whether or not the *lgr3[ex1]* is indeed a loss-of-function mutation. In this experiment we also included the control *lgr3[exp2]* chromosome, which is important to exclude the possibility that any linked mutations could be the cause of the suppression seen in the first experiment, rather than the *lgr3* mutation itself. As in the first experiment described

above, these mutations were placed in a normal (+/+) or developmentally-delayed background (*tub-dilp8/+*). Results show that there was a similar full rescue of the *dilp8*-dependent delay in transheterozygote *tub-dilp8/+;lgr3[ex1]/Df(3)BSC321* larvae as there was in *tub-dilp8/+;lgr3[ex1]/lgr3[ex1]* larvae (Figure 42).

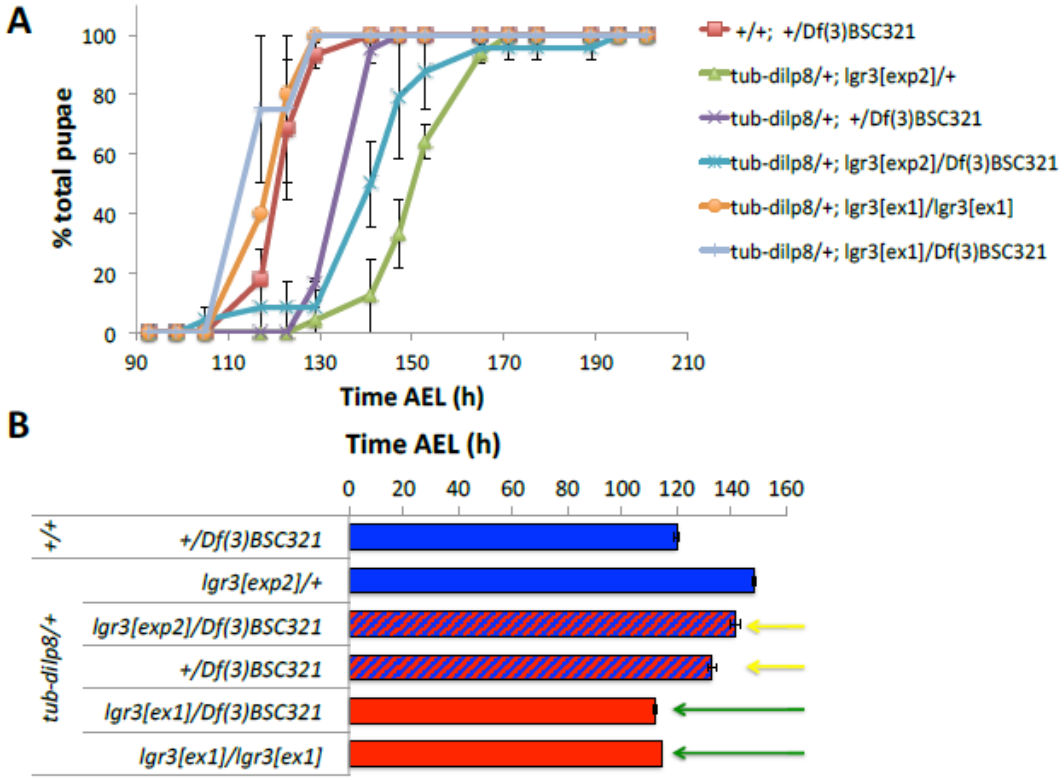


Figure 42. Preliminary evidence that *lgr3[ex1]* is a null mutation and that *dilp8* activity is sensitive to *lgr3* dosage. (A) Pupariation timing curves for different genotypes. Shown are averages +/- SEMs of 3 experiments except for *tub-dilp8/+;lgr3[ex1]/Df(3)BSC321* and *tub-dilp8/+;lgr3[ex1]/lgr3[ex1]*, which had two and one repeats, respectively. The genotype *+/+; +/Df(3)BSC321* served as a negative control. (B) Bar graph showing the median pupariation time of the depicted genotypes. Shown are averages +/- SEMs of 3 experiments except for *tub-dilp8/+;lgr3[ex1]/Df(3)BSC321* and *tub-dilp8/+;lgr3[ex1]/lgr3[ex1]*, which had two and one repeats, respectively. ANOVA was not performed due to the lack of at least three repeats in the above-mentioned genotypes. Bar colors depict the dosage of *lgr3* (blue = two wild-type alleles, red = two loss-of-function alleles; hashed bars = heterozygotes). Green arrows depict full suppression. Yellow arrows depict partial suppression.

These results again demonstrate that *lgr3* is required for *dilp8* activity and further confirm our hypothesis that *lgr3[ex1]* is a loss-of-function mutation, most-likely a null. In addition, by removing one copy of *lgr3* by placing *Df(3)BSC321* over two different control chromosomes III, "+" or *lgr3[exp2]*, we observed partial

rescues of the *tub-dilp8* delay when compared to control +/*lgr3[exp2]* animals, which have two functional copies of *lgr3*. These results are consistent with the partial rescue we found in *tub-dilp8;lgr3[ex1]/+* animals in the first experiment, and again suggest that *dilp8* activity is sensitive to *lgr3* dosage (Figure 42). Unfortunately, statistical analyses of these results was not possible because two critical crosses did not generate enough larvae to allow at least three repetitions (namely, there were two repetitions for the *tub-dilp8/+;Df(3)BSC321/lgr3[ex1]* and one for the positive control *tub-dilp8/+;lgr3[ex1]/lgr3[ex1]*). Therefore, despite the consistent results and similar trends to the first set of experiments, the data of this second set of experiment should formally be considered preliminary.

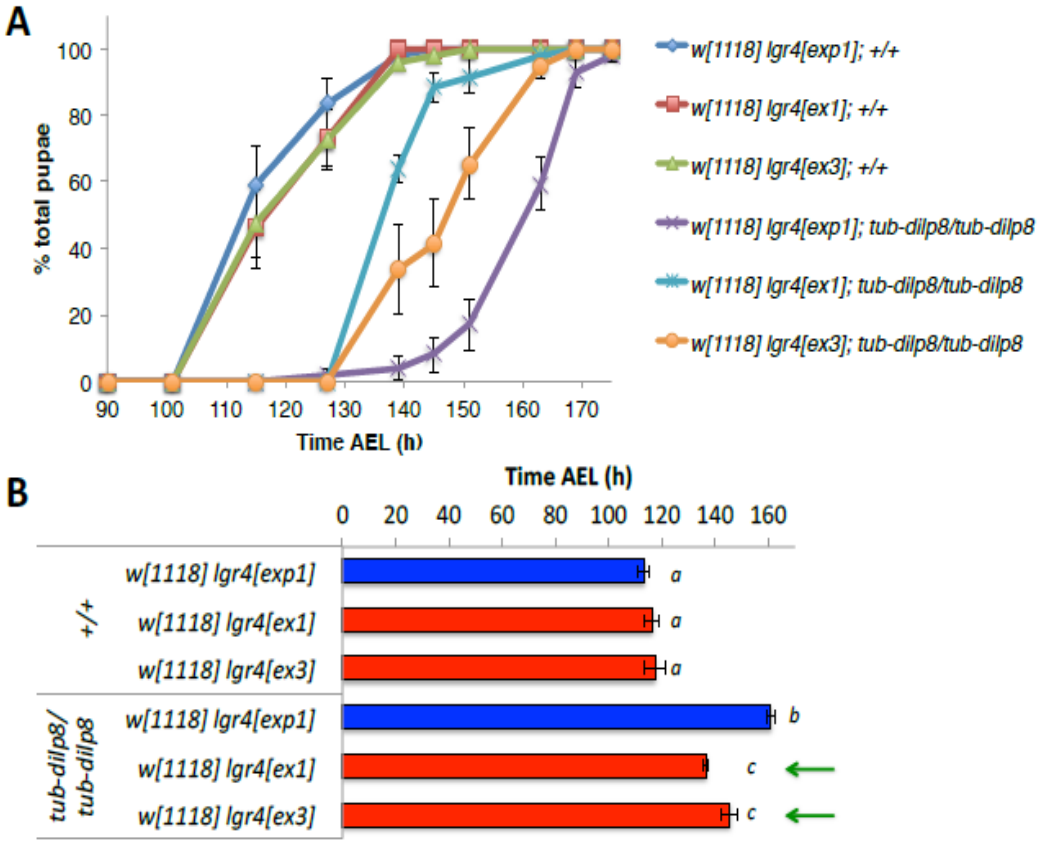


Figure 43. *lgr4* mutation partially rescues the dilp8-dependent delay in the onset of pupariation. (A) Pupariation timing curves for different genotypes. Shown are averages +/- SEMs of 4-6 experiments. (B) Bar graph showing the median pupariation time of the depicted genotypes. ANOVA: $p<0.001$. Groups sharing the same letters are not statistically significantly different at alpha=5% according to the Tukey's HSD post-hoc test. Shown are averages +/- SEMs of 4-6 experiments. Bar colors depict the dosage of *lgr4* (blue = two wild-type alleles, red = two loss-of-function alleles). Green arrows depict full suppression.

As regards *lgr4,* we placed the control *lgr4[exp1]* or the deletion mutations *lgr4[ex1]* or *lgr4[ex3]* into a wild-type (+/+) or developmentally-delayed background (*tub-dilp8/tub-dilp8*). Results show that, in the absence of *dilp8* ectopic activation, larvae carrying either *lgr4[ex1]* or *lgr4[ex3]* pupariated at the same time as control *lgr4[exp1]* larvae (Figure 43-A). These results suggest that *lgr4* does not strongly influence pupariation time during normal development (*e.g.,* in the absence of an ectopic source of *dilp8*). In contrast, when *dilp8* was ectopically activated, larvae carrying either *lgr4[ex1]* or *lgr4[ex3]* pupariated earlier than larvae carrying the control chromosome *lgr4[exp1]* ($p < 0.05$ for each *ex1* or *ex3* mutation, Tukey's HSD post-hoc Test; Figure 43-B). *lgr4[ex1]* seems to suppress the *dilp8*-dependent activity more strongly than *lgr4[ex3],* despite the difference not being statistically significantly different. These results suggest that *lgr4* is partially required for the *dilp8*-dependent delay in the onset of pupariation.

Chapter 4. Discussion and conclusion

In this project, we proposed to begin to characterize the biological function of both *Drosophila* Lgr3 and Lgr4 receptors. We wished to clarify their role during *Drosophila* development and test the hypothesis that one or both of them were acting downstream of the developmental-delay-inducing insulin-like peptide, Dilp8, possibly as the receptor for this peptide. For this, we began by using a transposable element remobilization strategy to generate stocks carrying deletion mutations in *lgr3* or *lgr4*. For *lgr3*, we obtained one strong loss-of-function allele, the *lgr3[ex1]* deletion, while for *lgr4*, we obtained two loss-of-function alleles: the *lgr4[ex1]* and *lgr4[ex3]* deletions. For both genes, we also obtained and characterized control precise excision alleles, *lgr3[exp2]* and *lgr4[exp1]*, in which the chromosome carrying the respective *MB* element was reverted back to a wild-type situation. All alleles were PCR-confirmed and their effect on mRNA transcription for both loci was determined. All deletion alleles are predicted to encode aberrant mRNAs that lead to proteins truncated close to the N-terminus of each receptor. The obtained lines were used in several kinds of experiments, such as expression assays, fertility assays and developmental assays, throughout the work.

In order to identify the pattern of expression of *lgr3* and *lgr4* during development we initially performed *in situ* hydribization using different tissues from all stages of development. Unfortunately, our results were not satisfactory because we could not detect any specific pattern that was present in the antisense strand compared to the sense strand control. Therefore, we decided to proceed to the analysis of the expression by qRT-PCR. This is a very sensitive technique, able to detect low transcript levels with a limited amount of sample tissue. Indeed, our qRT-PCR results revealed a very low expression level for the *lgr3* and *lgr4* receptors. Such low levels of expression could explain why we could not detect the tissue expression pattern of these genes by *in situ* hybridization. Although low, the expression levels were found to be slightly higher in males than in females for both receptor genes. During this thesis, a publication came out that independently confirms this finding for both receptors (Hiel *et al.,* 2014). Since we have not been able to detect any clear phenotype in *lgr3* or *lgr4* mutants during normal development, the functional relevance of this differential expression between males and females is yet to be determined.

In the fertility and viability results, we observed that male adults homozygous for *lgr3[ex1]* were semi-sterile. We were initially excited with this phenotype, however the male adults homozygous for the control *lgr3[exp2]* (and other *exp* alleles; not shown) all of which are wild-type for Lgr3 activity, were also semi-sterile to a variable extent and sometimes showed an even stronger effect than *lgr3[ex1]* (as in Figure 25-B). Whereas the *lgr3[MB06848]* element is homozygous viable in the original background and shows no male semi-sterility, our results suggest that a male semi-sterile mutation is only revealed in this background when this chromosome is placed with the chromosomes of the jump protocol. We have not pursued this mutation as it is clearly unrelated to Lgr3 activity. Nevertheless, due to this fact, we kept the *lgr3* excision alleles balanced with *TM6B*. We noticed however that while we were able to make homozygous viable stocks for all *exp* alleles, the *ex1* allele seems sicker. Preliminary observations suggest that there might be a mild pupal lethality phenotype in this stock which needs to be further quantified, controlled for and characterized. As regards *lgr4,* we detected no overt phenotype and all lines were homozygous viable, fertile and fecund.

A clear phenotype for the *Drosophila* relaxin-like receptors *lgr3* and *lgr4* was revealed when we assayed for their requirement in the *dilp8* developmental delay pathway. For this test, the mutant lines were crossed with themselves, *w[1118]* and with the developmentally-delayed *tub-dilp8* stocks. Several assays were performed with the *lgr3* mutants either in homozygosis or in trans-heterozygosis with a *Df* that completely removes *lgr3* and the results were very consistent: in all cases *lgr3* is absolutely required for the *dilp8*-dependent delay. Larvae lacking *lgr3* are insensitive to ectopically expressed *dilp8.* Interestingly, in all occasions tested, larvae lacking one of the two *lgr3* copies showed a partial loss of *dilp8* sensitivity, suggesting that the *dilp8*-dependent developmental delay is sensitive to *lgr3* dosage. These findings strongly place *lgr3* in the *dilp8* pathway, which could be consistent with a role as a receptor. However there are other possibilities which will be discussed further below.

When the same tests were performed for *lgr4*, a slight suppression was verified with both mutants assayed, *lgr4[ex1]* and *lgr4[ex3]*. This finding suggests that *lgr4* is partially required for the *dilp8*-dependent delay, and could suggest partial redundancy of both receptors in one or more target tissues. Despite the fact that *lgr3*

already completely suppresses the *dilp8*-dependent delay, it would be interesting to assay for the developmental and *dilp8*-dependent effects in animals lacking both receptors. For that, were have made double mutant lines (*lgr4[ex1];;lgr3[ex1], lgr4[ex1];;lgr3[exp2], lgr4[exp1];;lgr3[ex1]*) carrying and not the *tub-dilp8* transgene. These lines took several months to be made are were unfortunately not yet adequately tested, since all the controls are not yet ready, but they already reveal that *lgr4;lgr3* double mutants are viable and fertile. Hence, relaxin-like receptors play no essential role during *Drosophila* development in the laboratory environment in the absence of challenges or stimuli. Nevertheless the availability of these lines with the knowledge provided in this current work should give important information about the relationship that both receptors may have among themselves. For this it will be important to know more about the expression pattern of both receptor genes. Van Hiel and collaborators have recently profiled the expression of *lgr3* and *lgr4* in different tissues of developing larvae and adults using qRT-PCR (Van Hiel *et al.,* 2014). They find broad expression of *lgr3* and *lgr4* in many tissues, but for *lgr3* it is strongest in the central nervous system, midgut and gonad of the larvae, while *lgr4* is most strongly expressed in the larval midgut. Although interesting, these data do not provide subtissue resolution. For this resolution, a technique like immunofluorescence, *in situ* hybridization or genetic reporter systems are required. As there are no antibodies yet available for the *Drosophila* Lgr3 or Lgr4, our *in situ* hybridization attempts have failed, we believe that the next step would be to work towards genetic reporters to localize the cells where these receptors are expressed. The advent of the *CRISPR/Cas9* technique for gene targeting could be an appropriate solution to tag the endogenous *lgr3* and *lgr4* loci with fluorescent reporters. This is a technology that the lab will explore in the near future, and we have hope that it could help in the continuity of this project.

All the work done so far places the two *Drosophila* relaxin-like receptors in the *dilp8* pathway. Despite the fact that the vertebrate homologues of *lgr3* and *lgr4* are receptors for Ilps of the relaxin-like family, from our data alone we cannot yet say with certainty that both genes are also receptors for Dilp8. There are several other hypotheses that could likewise explain our results (Figure 44). For instance, Lgr3 and Lgr4 might not act as direct receptors for Dilp8. This can happen in many ways. One way is that Dilp8 binds its cognate receptor, leading to the secretion of a second

peptide that then activates Lgr3 or Lgr4, leading to a developmental delay. A second way is that Lgr3 and Lgr4 act in parallel to the Dilp8 pathway by being required for the proper expression of the receptor system for Dilp8. This might happen at several levels, such as during cell fate determination of the cells receiving the Dilp8 signal and/or during the expression of the cognate Dilp8 receptor in the right tissues (Figure 46).

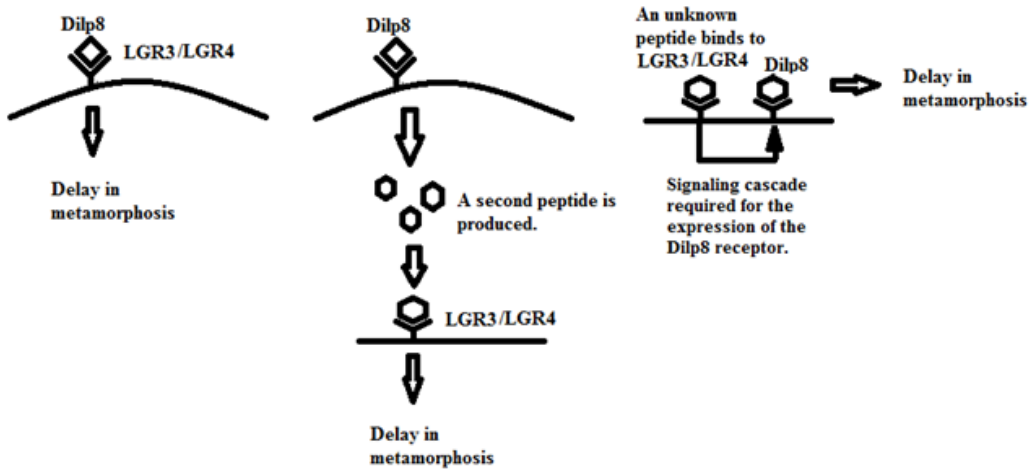


Figure 44. Hypothesis for the role of the Lgr3 and Lgr4 receptors, in the Dilp8 pathway.

The Lgr3 and Lgr4 belong to the subgroup of Rhodopsin-like GCPRs, as already mentioned. As GPCRs are a group of receptors of the Relaxin family hormones, which are implicated in the regulation of primarily reproductive and neuroreproductive processes, they are widely distributed throughout the vertebrate and invertebrate species along the evolution. To better understand the role that these receptors may have, several phylogenetic studies have been carried out (Muffato *et al*., 2008). One of the research groups, which have contributed to most of the advances in this area, is the group of Sara Good and Sergey Yogorov. Traditionally, the field has regarded the vertebrate RLN3 and RXFP3 as being the ancestor ligand x receptor pair for the relaxin signalling system. However, Yogoroc and Good's studies have showed strong evidences that the signaling of the ancestral relaxin peptide in the last common ancestor of vertebrates (and chordates) most-likely occurred via the RXFP1 and RXFP2-type receptors rather than RXFP3 or RXFP4 receptors (Yogorov & Good, 2012). Because Ilps and notoriously the relaxin

subfamily of Ilps evolve very quickly, it is difficult to trace back the evolution of Dilp8 or any other insect or invertebrate Ilp to the ancestral relaxin-like peptide of the last common ancestor of vertebrates. Nevertheless, as the Drosophila Lgr3, a bona fide homologue of the vertebrate relaxin-like receptor, is required for Dilp8 activity, it is very tempting to speculate that Dilp8 could be a ligand for this orphan family of invertebrate relaxin-like receptors. This hypothesis should be tested biochemically. A demonstration of a receptor x ligand interaction between Lgr3/4 and Dilp8 would suggest that the relaxin-like receptors have always been sensitive to Ilps since the last common ancestor to all metazoans alive. This would show once again the power of *Drosophila* as a relevant model organism to study human development and disease mechanisms. The insulin field benefited enormously from fly genetics with many of the key players in the pathway having been identified in *Drosophila.* Hopefully our work can help initiate a similar drive to dissect at the molecular genetic level the genes that participate and/or modulate conserved aspects of relaxin-like receptor signaling.

In conclusion, we show that the Drosophila relaxin-like Leucine-rich repeat-containing G-protein coupled receptors Lgr3 and Lgr4 are required and partially-required, respectively, for the activity of the developmental delay inducing insulin-like peptide, Dilp8.

4.1 Future perspectives

After characterizing the biological function of each *Drosophila* relaxin-like receptor separately, our work will move forward for the exploration of the relaxin-like receptor double mutants. As both *lgr3* and *lgr4* mutants are viable and fertile, it could be stated that these genes are not essential for *D. melanogaster* survival. A possible explanation is the fact that the *lgr3* and *lgr4* receptors could potentially be functionally redundant. Even though some of our double-mutant stocks already hint toward the idea that double-mutants are viable and fertile, this possibility has to be formally tested and controlled to detect more subtle phenotypes and also needs to be

explored by performing the developmental assays with the double mutant lines in a *tub-dilp8* context, which have already been constructed.

Until now, all the assays were made with lines that have Dilp8 constitutively expressed in every tissue. In order to produce the same results, the mutant lines for both receptors, will be subjected to imaginal disc damage, so that the endogenous Dilp8 is activated. We could thus test whether the mutations in the relaxin-like receptors rescue an endogenously-produced developmental delay. For that, we will be performed assays with *ethyl methanosulfonate* (EMS), a genotoxic agent that produce strong caspase activation in imaginal discs, leading to strong tissue damage and tissue regeneration, together with a strong transcriptional activation of *dilp8*, which leads to a devleopmental delay (Garelli *et al.*, 2012).

It will also be important to rescue the phenotype of our mutants, by replacing the normal receptor, in order to verify if the delay is restored. For that, we need to cross our mutants for both receptors, with lines that have the receptor constitutively active. We can develop and characterize UAS-Lgr3-eGFP and UAS-Lgr4-eGFP transgenes for this purpose. Now that we have a phenotype for each receptor (*dilp8*-delay rescue), we can formally test whether or not the transgenes are functional.

Finally, it remains important to understand more about the tissue and cellular expression pattern of the receptors, in order to help realize their specific function. This will surely be one way through which this work will continue.

Chapter 5. Bibliography

Adams M.D, et al. (2000). The genome sequence of drosophila melanogaster. *Science* 287, 2185-2195.

Aliaga IM, Thor S, Gould AP. (2008). Postmitotic Specification of Drosophila Insulinergic Neurons from Pioneer Neurons. *PLoS Biology* 6(3): e58.

Arroyo JI, Hoffmann FG, Opazo JC. (2014). Evolution of the Relaxin/Insulin-Like Gene Family in Anthropoid Primates. *Genome Biology Evolution* 6(3):491–499.

Ashburner M, Roote J. (2007). Maintenance of a Drosophila laboratory: general procedures. *CSH Protocols* 2007, pdb.ip35.

Audet M, Bouvier M. (2012). Restructuring G-Protein-Coupled Receptor Activation. *Cell* 151(1) p14–23.

Bailit HL, Workman PL, Niswander JD, Maclean CJ. (1970). Dental asymmetry as an indicator of genetic and environmental conditions in human populations. *Human Biology* 42, 626-638.

Baker JD, Truman JW. (2002). Mutations in the Drosophila glycoprotein hormone receptor, rickets, eliminate neuropeptide-induced tanning and selectively block a stereotyped behavioral program. *The Journal of Experimental Biology* 205(17): 2555-2565.

Ballen KK, Shrestha S, Sobocinski KA, Zhang MJ, Bashey A, Bolwell BJ. (2010). Outcome of transplantation for myelofibrosis. *Biol Blood Marrow Transplant* 16: 358–367.

Bargmann CI. (1998). Neurobiology of the Caenorhabditis elegans genome. *Science* 282, 2028–2033.

Bathgate RAD, Halls ML, Van der Westhuizen ET, Callander GE, Kocan M, Summers RJ. (2013). Relaxin family peptides and their receptors. *Physiological Reviews* 93:405-480.

Bathgate RAD. (2006). International union of pharmacology (IUPHAR); Recommendations for the nomenclature of receptors for relaxin family peptides. *Pharmacological Reviews* 58:7-31.

Bellen HJ, Tong C, Tsuda H. (2010). 100 years of Drosophila research and its impact on vertebrate neuroscience: a history lesson for the future. *Nature Reviews* Neuroscience 11, 514-522.

Bell RJ, Permezel M, MacLennan A, Hughes C, Healy D, Brennecke S. (1993). A randomized, double-blind, placebo-controlled trial of the safety of vaginal

recombinant human relaxin for cervical ripening. *Obstetrics & Gynecology* 82: 328–333.

Bellen HJ, O'Kane CJ, Wilson C, Grossniklaus U, Pearson RK, Gehring WJ. (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. *Genes & Development* 3(9):1288-300.

Bier E, Bodmer R. (2004). Drosophila, an emerging model for cardiac disease. *Gene* 2004;342:1–11

Bischof J, Basler K. (2008). Recombinases and their use in gene activation, gene inactivation, and transgenesis. Drosophila Methods and Protocols. *Humana Press*. pp. 175-195.

Blundell TL, Humbel RE. (1980). Hormone families: pancreatic hormones and homologous growth factors. *Nature* 287, 781-787.

Bogatcheva NV, Truong A, Feng S, Engel W, Adham IM, Agoulnik AI. (2003). GREAT/LGR8 is the only receptor for insulin-like 3 peptide. *Molecular Endocrinology* 17: 2639–2646.

Bridges TM, Lindsley CW. (2008). G-protein-coupled receptors: from classical modes of modulation to allosteric mechanisms. *ACS Chemical Biology* 3,530-541.

Brody T, Cravchik A. (2000). Drosophila melanogaster G protein-coupled receptors. *The Journal of cell biology* 150(2), F83–8.

Broeck JV. (1996). G-protein-coupled receptors in insect cells. *International Review of Cytology* 164: 189–268.

Brogiolo W, Stocker H, Ikeya T, Rintelen F, Fernandez R, Hafen E. (2001). An evolutionarily conserved function of the Drosophila insulin receptor and insulin-like peptides in growth control. *Current Biology - Cell* 11(4):213-21.

Brown MR, Clark KD, Gulia M, Zhao Z, Garczynski SF, Crim JW, Suderman RJ, Strand MR. (2008). An insulin-like peptide regulates egg maturation and metabolism in the mosquito Aedes aegypti. *PNAS* 105(15):5716-21.

Carvalho GB, Já WW, Benzer S. (2009). Non-lethal PCR genotyping of single Drosophila. *BioTechniques*, 46(4), 312–4.

Castellanos MC, Tang JC, Allan DW. (2013). Female-biased dimorphism underlies a female-specific role for post-embryonic Ilp7 neurons in Drosophila fertility. *Development* 140(18): 3915-3926.

Chassin D, Laurent A, Janneau JL, Berger R, Bellet D. (1995). Cloning of a new member of the insulin gene superfamily (INSL4) expressed in human placenta. *Genomics* 29: 465–470.

Colombani J, Andersen DS, Léopold P. (2012). Secreted peptide Dilp8 coordinates Drosophila tissue growth with developmental timing. *Science* 336(6081), 582–5.

Colombani J, Bianchini L, Layalle S, Pondeville E, Villemant CD, Antoniewski C, Carré C, Noselli S, Léopold P. (2005). Antagonistic Actions of Ecdysone and Insulins Determine Final Size in Drosophila. *Science* Vol. 310 no. 5748 pp. 667-670.

Conklin D, Lofton-Day CE, Haldeman BA, Ching A, Whitmore TE, Lok S, Jaspers S. (1999). Identification of INSL5, a new member of the insulin superfamily. *Genomics* 15;60(1):50-6.

Cook RK, Christensen SJ, Deal JA, Coburn RA, Deal ME, Gresens JM, Kaufman TC. (2012). The generation of chromosomal deletions to provide extensive coverage and subdivision of the Drosophila melanogaster genome. *Genome Biology* 13:R21.

Denver RJ, Middlemis-Maher J. (2010). Lessons from evolution: developmental plasticity in vertebrates with complex life cycles. *Journal of Developmental Origins of Health and Disease* 1(5):282-291.

Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, Gasser B, Kinsey K, Oppel S, Scheiblauer S, Couto A, Marra V, Keleman K, Dickson BJ. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* 448(7150):151-6.

Dixon RFA, Kobilka BK, Strader DJ, Benovic JF, Dohlman HG, Frielle T, Bolanowski MA, Bennett CD, Rands E, Diehl RE, Mumford RA, Slater EE, Sigal IS, Caron MG, Lefkowitz RJ, Strader CD. (1986). Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. *Nature*, 1-7;321(6065):75-9.

Edgar BA. (2006). How flies get their size: genetics meets physiology. *Nature Reviews* 7(12): 907--916.

Elliott DA, Brand AH. (2008). The GAL4 system: a versatile system for the expression of genes. Drosophila Methods and Protocols. *Humana Press*. pp. 79-95.

Ellis C. (2004). The state of GPCR research in 2004. *Nature Reviews Drug Discovery* 3, 577-626.

Eriksen KW, Kaltoft K, Mikkelsen G, Nielsen M, Zhang Q, Geisler C, Nissen MH, Röpke C, Wasik MA, Odum N. (2001). Constitutive STAT3-activation in Sezary syndrome: tyrphostin AG490 inhibits STAT3-activation, interleukin-2 receptor expression and growth of leukemic Sezary cells. *Leukemia.* 15(5):787-93.

Fredriksson R, and Schioth HB. (2005). The repertoire of G-protein–coupled receptors in fully sequenced genomes. *Molecular Pharmacology* 67:1414–1425.

Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular Pharmacology* 63(6):1256-72.

Garber SL, Mirochnik Y, Brecklin CS, Unemori EN, Singh AK, Slobodskoy L, Grove BH, Arruda JA, Dunea G. (2001). Relaxin decreases renal interstitial fibrosis and slows progression of renal disease. *Kidney International, Nature* 59, 876–882.

Garelli A, Gontijo AM, Miguela V, Capattos E, Dominguez M. (2012). Imaginal Discs secrete insulin-like peptide 8 to mediate plasticity of growth and maturation. *Science*, 336, 579.

Gronke S, Clarke DF, Broughton S, Andrews TD, Partridge L. (2010). Molecular Evolution and Functional Characterization of Drosophila Insulin-Like Peptides. *PLoS Genetics* 6(2): e1000857.

Hackney, JF, Zolali-Meybodi, O and Cherbas P. (2012). Localized Tissue Damage Disrupts Developmental Progression and Ecdysteroidogenesis in Drosophila. *PLoSONE*. 7(11):e49105.

Haley J, Hudson P, Scanlon D, John M, Cronk M, Shine J, Tregear G, Niall H. (1982). Porcine relaxin: molecular cloning and cDNA structure. *DNA* 1: 155–162.

Halme A, Cheng M, Hariharan IK. (2010). Retinoids Regulate a Developmental Checkpoint for Tissue Regeneration in Drosophila. *Current Biology*. 20(5): 458--463.

Hariharan IK. (2012). How growth abnormalities delay "puberty" in Drosophila. *Science Signal* 5(229):pe27.

Harmar AJ.(2001). Family-B G-protein-coupled receptors. *Genome Biology* 2(12):reviews3013.1–3013.10.

Hill RJ, Billas IML, Bonneton F, Graham LD, Lawrence MC. (2012). Ecdysone Receptors: From the Ashburner Model to Structural Biology. *Annual review of entomology* 58 251-71.

Hisaw FL. (1926). Experimental relaxation of the pubic ligament of the guinea pig. *Proceedings of the Society for Experimental Biology and Medicine* 23:661-663.

Hsu SY, Nakabayashi K, Nishi S, Kumagai J, Kudo M, Sherwood OD, Hsueh AJ. (2002). Activation of orphan receptors by the hormone relaxin. *Science* 295: 671–674.

Hudson P, Haley J, Cronk M, Shine J, Niall H. (1981). Molecular cloning and characterization of cDNA sequences coding for rat relaxin. *Nature* 291: 127–131.

Hudson P, Haley J, John M, Cronk M, Crawford R, Haralambidis J, Tregear G, Shine J, Niall H. (1983). Structure of a genomic clone encoding biologically active human relaxin. *Nature* 301: 628–631.

Hudson P, John M, Crawford R, Haralambidis J, Scanlon D, Gorman J, Tregear G, Shine J, Niall H. (1984). Relaxin gene expression in human ovaries and the predicted structure of a human preprorelaxin by analysis of cDNA clones. *EMBO* J 3: 2333–2339.

Hussey RG, Thompson WR, Calhoun ET. (1927). The influence of X-rays on the development of Drosophila larvae. S*cience* 66, 65–66.

Ibegbu AO, Mullaney I, Fyfe L, Macbean D. (2012). The roles of G-protein coupled receptors in health and disease conditions. *International Journal of Medical and Health Sciences* 1(2): 2277-4505.

Ikeya T, Galic M, Belawat P, Nairz K, Hafen E. (2002). Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in Drosophila. *Current Biology – Cell 12*(15): 1293-1300.

Insel AP, Tang CM, Hahntow I, Michel MC. (2007). Impact of GPCRs in clinical medicine: genetic variants and drug Targets. *Biochim Biophys Acta* 1768(4): 994–1005.

Ivell R, Bathgate RA. (2002). Reproductive biology of the relaxin-like factor (RLF/INSL3). Biology of Reproduction 67: 699–705.

Junger MA, Rintelen F, Stocker H, Wasserman JD, Vegh M, Radimerski T, Greenberg ME, Hafen E. (2003). The Drosophila Forkhead transcription factor FOXO mediates the reduction in cell number associated with reduced insulin signaling. The Journal of Biological Chemistry 2(3): 20.

Katritch V, Cherezov V, Stevens RC. (2013). Structure-function of the G protein-coupled receptor superfamily. Annual review of pharmacology and toxicology 53, 531–56.

Koman A, Cazaubon S, Couraud PO, Ullrich A, Strosberg AD. (1996). Molecular characterization and in vitro biological activity of placentin, a new member of the insulin gene family. The Journal of Biological Chemistry 271: 20238–20241.

Kong RC, Shilling PJ, Lobb DK, Gooley PR, Bathgate RA. (2010). Membrane receptors: structure and function of the relaxin family peptide receptors. Molecular Cell Endocrinology 320:1–15.

Kumagai J, Hsu SY, Matsumi H, Roh JS, Fu P, Wade JD, Bathgate RA, Hsueh AJ. (2002). INSL3/Leydig insulin-like peptide activates the LGR8 receptor important in testis descent. *The Journal of Biological Chemistry* 277: 31283–31286.

Liebmann C. (2004). G protein-coupled receptors and their signaling pathways: classical therapeutical targets susceptible to novel therapeutic concepts. *Current Pharmaceutical Design*, 10:1937–58.

Ling D, Salvaterra PM. (2011). Robust RT-qPCR data normalization: validation and selection of internal reference genes during post-experimental data analysis. *PloS one.* 6: e17762.

Linneweber GA, Jacobson J, Busch KE, Hudry B, Christov CP, Dormann D, Yuan M, Otani T, Knust E, de Bono M, Miguel-Aliaga I. (2014). Neuronal Control of Metabolism through Nutrient-Dependent Modulation of Tracheal Branching. *Cell* 156(1-2): 69--83.

Liu C, Chen J, Sutton S, , Roland B, Kuei C, Farmer N, Sillard R, Lovenberg TW (2003b). Identification of Relaxin-3/INSL7 as a Ligand for GPCR142. Journal of Biological Chemistry 278: 50765–50770.

Liu C, Eriste E, Sutton S, Chen J, Roland B, Kuei C, Farmer N, Jörnvall H, Sillard R, Lovenberg TW (2003a). Identification of Relaxin-3/INSL7 as an Endogenous Ligand for the Orphan G-protein-coupled Receptor GPCR135. Journal of Biological Chemistry 278: 50754–50764.

Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. (2001). Methods San Diego Calif 25, 402–408.

Lok S, Johnston DS, Conklin D, Lofton-Day CE, Adams RL, Jelmberg AC, Whitmore TE, Schrader S, Griswold MD, Jaspers SR. (2000). Identification of INSL6, a new member of the insulin family that is expressed in the testis of the human and rat. Biology of Reproduction 62:1593–1599.

Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, et al. (2005) Modeling gene and genome duplications in eukaryotes. PNAS 102: 5454–5459.

Makino T, Hokamp K, McLysaght A. (2009). The complex relationship of gene duplication and essentiality. Trends in Genetics 25(4):152-5.

McBrayer Z, Ono H, Shimell MJ, Parvy JP, Beckstead RB, Warren JT, Thummel CS, Dauphin-Villemant C, Gilbert LI, O'Connor MB. (2007). Prothoracicotropic hormone regulates developmental timing and body size in Drosophila. *Development - Cell* 13(6): 857--871.

Menut L, Vaccari T, Dionne H, Hill J, Wu G, Bilder D. (2007). A mosaic genetic screen for Drosophila neoplastic tumor suppressor genes based on defective pupation. *Genetics* 177, 1667-1677.

Metaxakis A, Oehler S, Klinakis A, Savakis C. (2005). Minos as a genetic and genomic tool in Drosophila melanogaster. *Genetics* 171:571-581.

Minelli A, Fusco G. (2010). Developmental plasticity and the evolution of animal complex life cycles. Philosophical Transactions of the Royal Society B. doi: 10.1098/rstb.2009.0268.

Miron M, Verdu J, Lachance PED, Birnbaum MJ, Lasko PF, Sonenberg N. (2001). The translational inhibitor 4E-BP is an effector of PI(3)K/Akt signalling and cell growth in Drosophila. *Nature Cell Biology* 3(6): 596--601.

Mita M. (2013). Relaxin-like gonad-stimulating substance in an echinoderm, the starfish: A novel relaxin system in reproduction of invertebrates. General and *Comparative Endocrinology Volume* 181, Pages 241–245.

Moczek AP. (2010). Phenotypic plasticity and diversity in insects. In From polyphenism to complex metazoan life cycles, edited by A. Minelli, G Fusco *Philosophical Transactions of the Royal Society of London, Series B* 365: 593-603.

Møller AP, Thornhill R. (1997). Developmental stability is heritable. *Journal of Evolutionary Biology* 10:69–76.

Moller AP. (1993). Morphology and sexual selection in the barn swallow, Hirundo rustica, in Chernobyl, Ukraine. *Proccedings of the Royal society of London B* 252, 51-57.

Montell DJ. (2003). Border-cell migration: the race is on. Nature Reviews Molecular *Cell Biology* 4, 13-24.

Muffato M, Roest Crollius H. (2008). Paleogenomics in vertebrates, or the recovery of lost genomes from the mist of time. *Bioessays* 30: 122–134.

Nakatani Y, Takeda H, Kohara Y, Morishita S. (2007). Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res* 17: 1254–1265.

Nef S, Parada LF.(1999). Cryptorchidism in mice mutant for Insl3. *Nature Genetics* 22 (3):295-299.

Okamoto N, Yamanaka N, Yagi Y, Nishida Y, Kataoka H, O'Connor MB, Mizoguchi A. (2009). A fat body-derived IGF-like peptide regulates postfeeding growth in Drosophila. *Cell* 17(6): 885-891.

Oldham WM, Hamm HE. (2008). Heterotrimeric G protein activation by G-protein-coupled receptors. *Nature Reviews Molecular Cell Biology* 9, 60-7.

Overbeek PA, Gorlov IP, Sutherland RW, Houston JB, Harrison WR, Boettger-Tong HL, Bishop CE, Agoulnik AI. (2001). A transgenic insertion causing cryptorchidism in mice. *Genesis* 30: 26–35.

Overington JP, Al-Lazikani B, Hopkins AL. (2006). How many drug targets are there? *Nature Reviews Drug Discovery* 5, 993-996.

Parker L, Padilla M, Du Y, Dong K, Tanouye MA. (2011). Drosophila as a Model for Epilepsy: bss Is a Gain-of-Function Mutation in the Para Sodium Channel Gene That Leads to Seizures. *Genetics* 187(2): 523--534.

Perez DM. (2005). From Plants to Man: The GPCR “Tree of Life”. *Molecular pharmacology* 67:1383–1384.

Puig O, Marr MT, Ruhf ML, Tijan R. (2003). Control of cell number by Drosophila FOXO: downstream and feedback regulation of the insulin receptor pathway. *Genes & Development*. 17(16):2006-2000.

Purves D, Augustine GJ, Fitzpatrick D. (2001). G-Proteins and Their Molecular Targets, in Purves D. (2nd edition) *Neuroscience*.

Reiter LT, Potocki L, Chien S, Gribskov M, Bier E. (2001). A systematic analysis of human disease associated gene sequences in Drosophila melanogaster. *Genome research* 11:1114-1125.

Ritter SL, Hall RA. (2009). Fine-tuning of GPCR activity by receptor-interacting proteins. Nature reviews. *Molecular cell biology* 10(12), 819–30.

Robert EK. (1994). Lords of the fly "Drosophila genetics and experimental life". *University of Chicago Press* 1999;23-29.

Rørth P. (1996). A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. *Proceedings of National Academy of Sciences of the United States of America* 93(22): 12418-12422.

Rulifson EJ, Kim SK, Nusse R. (2002). Ablation of Insulin-producing neurons in flies: growth and diabetic phenotypes. *Science* 296(5570).

Sajid W, Kulahin N, Schluckebier G, Ribel U, Henderson HR, Tatar M, Hansen BF, Svendsen AM, Kiselyov VV, Nørgaard P, Wahlund PO, Brandt J, Kohanski RA, Andersen AS, De Meyts P. (2011). Structural and Biological Properties of the Drosophila Insulin-like Peptide 5 Show Evolutionary Conservation. *The Journal of Biological Chemistry* 286(1): 661--673.

Schoneberg T, Schulz A, Biebermann H, Hermsdorf T, Rompler H, Sangkuhl K. (2004). Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacological Therapy*, 104:173–206.

Seafold SC. (2005). G Protein-Coupled Receptors. *Science* STKE, Vol. 2005, Issue 279, p. tr11.

Sherwood OD. (1994). Relaxin. In: The Physiology of Reproduction , edited by Knobil E, Neill JD. *New York: Raven*, p. 861–1009.

Shingleton, A.W. (2010) . The regulation of organ size in Drosophila: Physiology, plasticity, patterning and physical force. *Organogenesis*, 6, 1-12.

Simpson KL, Beck AD, Sorgeloos P. (1980). Report Workshop. Characterization of Artemia strains for application in aquaculture. In: Persoone G, Sorgeloos P,

Roels 0A, Jaspers E. (eds) The Brine Shrimp Artemia, Vol. 3, Ecology, Culturing, Use in Aquaculture. *UniversaPress, Wetteren*.

Slaidina M, Delanoue R, Gronke S, Partridge L, Léopold P. (2009). A Drosophila insulin-like peptide promotes growth during nonfeeding states. Cell 17(6): 874--884.

Smith CM, Shen PJ, Banerjee A, Bonaventure P, Ma S, Bathgate RA, Sutton SW, Gundlach AL. (2010). Distribution of relaxin-3 and RXFP3 within arousal, stress, affective, and cognitive circuits of mouse brain. J Comp Neurol 518: 4016–4045.

Tang CM, Insel PA. (2005). Genetic variation in G-protein-coupled receptors--consequences for G-protein-coupled receptors as drug targets. *Expert Opin Ther Targets*, 9:1247-65.

Tao YX. (2006). Inactivating mutations of G protein-coupled receptors and diseases: Structure-function insights and therapeutic implications. *Pharmacological Therapy*, 111(3):949-73.

Theodosiou NA, Xu T. (1998). Use of FLP/FRT system to study Drosophila *development. Methods* 14(4): 355-365.

Thompson MD, Burnham WM, Cole DE. (2005). The G protein-coupled receptors: pharmacogenetics and disease. *Crit Rev Clin Lab Sci*, 42:311–92.

Thompson MD, Percy ME, Mclntyre BW, Cole DE. (2008). G protein-coupled receptors disrupted in human genetic disease. *Methods in Molecular Biology* 448:109-37.

Thornhill R, Moller AP. (1997). Developmental stability, disease and medicine. *Biological Reviews*, 72(4) 497-548.

Tyndall JD, Sandilya R. (2005). GPCR agonists and antagonists in the clinic. *Journal of Medicinal Chemistr*y 1(4):405-21.

Unemori E N, Pickford L B, Salles A L, Piercy C E, Grove B H, Eriksen M E, and Amento E P. (1996). Relaxin induces an extracellular matrix degrading phenotype in human lung fibroblasts in vitro and inhibits lung fibrosis in a murine model in vivo. *The Journal of Clinical Investigation* 98, 2739–2745.

Unemori EN, Amento EP. (1990). Relaxin modulates synthesis and secretion of procollagenase and collagen by human dermal fibroblasts. *Journal of Biological Chemistry* 265(18), 10681–10685.

Unemori EN, Beck LS, Lee WP, Xu Y, Siegel M, Keller G, Liggitt HD, Bauer EA, and Amento EP. (1993). Human relaxin decreases collagen accumulation in vivo in two rodent models of fibrosis. *Journal of Investigative Dermatology* 101, 280–285.

Van Hiel MB, Vandersmissen HP, Loy TV, Broeck JV. (2011). An evolutionary comparison of Leucine-Rich repeat containing G protein coupled receptors reveals a novel LGR subtype. *Peptides* doi:10.1016/j.peptides.2011.11.004.

Van Hiel MB, Vandersmissen HP, Proost P, Broeck VJ. (2014). Cloning, constitutive activity and expression profiling of two receptors related to relaxin receptors in Drosophila melanogaster. P*eptides* doi: 10.1016/j.peptides.2014.07.014.

Wen X, Li D, Iles RK. (2010). Estradiol, progesterone, testosterone profiles in human follicular fluid and cultured granulosa cells from luteinized pre-ovulatory follicles. *Reproductive Biology and Endocrinology* 8:117.

Westhuizen E, Werry TD, Sexton PM, Summers RJ. (2007). The Relaxin Family Peptide Receptor 3 Activates Extracellular Signal-Regulated Kinase 1/2 through a Protein Kinase C-Dependent Mechanism. *Molecular Pharmacology* vol. 71 no. 6 1618-1629.

Wilkinson TN, Bathgate RA. (2007). The evolution of the relaxin peptide family and their receptors. *Advances in Experimental Medicine and Biology* 612:1-13.

Wilkinson TN, Speed TP, Tregear GW, Bathgate RA. (2005). Evolution of the relaxin-like peptide family: from neuropeptide to reproduction. Annals of the *New York Academy of Sciences* 1041: 530–533.

Williams EJ, Benyon RC, Trim N, Hadwin R, Grove BH, Arthur MJ, Unemori EN, Iredale JP. (2001). Relaxin inhibits effective collagen deposition by cultured hepatic stellate cells and decreases rat liver fibrosis in vivo. *Gut* 49, 577–583.

Wolf MJ, Rockman HA. (2008). Drosophila melanogaster as a model system for genetics of postnatal cardiac function. *Drug Discovery Today* 5:117–123.

Wolf MJ, Rockman HA. (2011). Drosophila, Genetic Screens, and Cardiac Function. *Circulation Research* 109(7): 794-806.

Yamanaka N, Romero NM, Martin FA, Rewitz KF, Sun M, O'Connor MB, Léopold P. (2013). Neuroendocrine control of Drosophila larval light preference. *Science* 341(6150): 1113-1116.

Yang CH, Belawat P, Hafen E, Jan LY, Jan YN. (2008). Drosophila egg-laying site selection as a system to study simple decision-making processes. Science 319:1679–1683.

Yegorov S, Bogerd J, Good S. (2014). The relaxin family peptide receptors and their ligands: New developments and paradigms in the evolution from jawless fish to mammals. General and Comparative Endocrinology, Elsevier, DOI: 10.1016/j.ygcen.2014.07.014.

Yegorov S, Good S. (2012). Using Paleogenomics to Study the Evolution of Gene Families: Origin and Duplication History of the Relaxin Family Hormones and Their Receptors. PLoS ONE 7(3): e32923.

Zhang Z, Qian W, Zhang J. (2009). Positive selection for elevated gene expression noise in yeast. Molecular Systems Biology 5:299.

Zimmermann S, Steding G, Emmen JM, Brinkmann AO, Nayernia K, Holstein AF, Engel W, Adham IM. (1999). Targeted disruption of the INSL3 gene causes bilateral cryptorchidism. Molecular Endocrinology 13: 681–691.