UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE des SCIENCES de la VIE et de la SANTÉ

UPR9022 CNRS, IBMC

Thesis presented by

Jelena Patrnogić

Public defense on September 26th, 2014

In order to obtain a PhD degree in Life Sciences
(Molecular and cellular aspects of biology)

Serine proteases and serine protease homologs : genetic analysis of their involvement in immune response activation in *Drosophila*

Thesis supervisor :
Prof. Jean-Marc Reichhart — University of Strasbourg, France

Reporters :
Dr. Michèle Crozatier — University of Toulouse, France
Dr. Petros Ligoxygakis — University of Oxford, UK

Other members of the jury :
Prof. Jean-Luc Imler — University of Strasbourg, France

Acknowledgements

One of my favorite quotes says that the only real mistake is the one from which we learn nothing. I've learned plenty. I am still learning.

These past few years passed so quickly that the notion of leaving and starting anew is still hard to grasp. These past few years will carry the most wonderful memories of my life. There are so many people that come to mind now, people that shaped my time here. I owe my gratitude to them.

My PhD supervisor, Prof. Jean-Marc Reichhart, thank you for giving me the opportunity to be a part of your team, to be a part of this lab. Thank you for all the support and help you have given me when I needed it.

Dr. Vincent Leclerc, thank you for your guidance, for all the discussions we had, for all the advice. Most importantly, thank you for making me find my own way.

Members of my team, thank you for all the help and support.

Prof. Jean-Luc Imler, thank you for having your door always open for me, for all the support and advice you have given me.

Dr. Eric Marois, thank you for being there to answer all my questions.

All the members of UPR9022, I am very grateful.

My friends, thank you for always being by my side.

You. I wish you were here.

My family, my biggest support, my biggest comfort. Thank you for never questioning my decisions. Thank you for believing in me. Thank you for allowing me to pursue my dreams. This is for you.

Table of Contents

Abbreviations

aa	Amino acid
AMP	Antimicrobial peptide
BAEE	Nα-benzoyl-L-arginine ethyl ester
Cas9	CRISPR-associated protein 9
CCP	Complement Control Protein
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
Cys	Cysteine
DAMP	Damage (Danger) Associated Molecular Pattern
DAP	*meso*-Diaminopimelic acid
ds	Double strand
EGF	Epidermal Growth Factor
FLP	Flippase
FRT	Flippase Recognition Target
FXa	Factor Xa
GD	Gastrulation Defective
Gla	γ-carboxyglutamic acid
GlcNAc	N-Acetylglucosamine
GNBP	Gram-Negative Binding Protein
Grass	Gram Positive Specific Serine Protease
HR	Homologous Recombination
IL-1	Interleukin-1
IMD	Immune Deficiency
IRC	Immune Responsive Catalase
ISPL5	Immune-related Serine Protease Like sequence 5

LDLa	Low Density Lipoprotein-receptor class A
ModSP	Modular Serine Protease
MP	Melanization Protease
MSP	Modular Serine Protease
MurNAc	N-Acetilmuramic acid
Nec	Necrotic
NF-κB	Nuclear Factor kappa B
NHEJ	Non-Homologous End Joining
PAM	Protospacer Adjacent Motif
PAMP	Pathogen Associated Molecular Patterns
PAP	proPO Activating Enzyme
PGN	Peptidoglycan
PGRP	Peptidoglycan Recognition Protein
PO	PhenolOxidase
PPAE	proPO-Activating Enzyme
PPAF	proPO-Activating Factor
PPO	ProPhenolOxidase
PRR	Pattern Recognition Receptor
PrtA	Protease A
Psh	Persephone
PVS	PeriVitelline Space
PZ	Protein Z
RCL	Reactive Center Loop
RIP	Receptor Interacting Protein
ROS	Reactive Oxygen Species
Ser	Serine
serpin	Serine Protease Inhibitor
sgRNA	Single guide RNA

SP	Serine Protease
SPCLIP1	Serine Protease CLIP1
SPE	Spaetzle Processing Enzyme
SPH	Serine Protease Homolog
Spn27A	Serpin 27A
Spn28D	Serpin 28D
Spn77b	Serpin 77b
Spz	Spaetzle
TALEN	Transcription Activator-Like Effector Nuclease
TCT	Tracheal Cytotoxin
TLR	Toll-Like Receptor
TNF-R	Tumor Necrosis Factor Receptor
tracrRNA	Trans-activating CRISPR RNA
ZPI	PZ-dependent Protease Inhibitor

Introduction

1. Innate vs. adaptive immunity

All multicellular organisms have at their disposal an array of mechanisms to detect and fight pathogens found in their environment. These mechanisms rely on the innate and the adaptive immune responses.

Vertebrates have both innate and adaptive immunity systems, whereas invertebrates solely rely on innate immunity.

Adaptive immunity develops as a response to infection. It is mediated by humoral (B lymphocytes) and cellular (T lymphocytes) mechanisms, using different components to eliminate different types of pathogens. Through somatic recombination, B and T lymphocytes express a wide array of specific receptors capable of recognizing different pathogens. An innate immune system is present in all organisms and is phylogenetically older than the adaptive immune system (Hoffmann and Reichhart, 2002). Many aspects of innate immune responses are evolutionarily conserved (Dushay and Eldon, 1998; Mushegian and Medzhitov, 2001). The innate immune system provides early defense responses against invading pathogens via mechanisms that include phagocytosis, proteolytic cascade activation, and synthesis of potent antimicrobial peptides.

In 1989, Charles Janeway proposed that innate immune mechanisms are essential for the earliest detection and defense against infection in mammals (Janeway, 1989). These mechanisms rely on discrimination between self and microbial non-self. Detection is achieved by germ-line encoded receptors, Pattern

Recognition Receptors (PRRs), that recognize conserved signature molecules expressed by the pathogens, namely Pathogen Associated Molecular Patterns (PAMPs) (Medzhitov and Janeway, 2000). Recognition of microbial non-self by PRRs is a universal strategy of innate immunity. It has been found in all studied multicellular organisms (Mushegian and Medzhitov, 2001). Polly Matzinger put forward a "danger signal" theory, stating that the immune system is more concerned with the danger and potential destruction rather than non-self. This model proposes that the central role of innate immune mechanisms is not the discrimination between self and non-self, but rather the detection of and protection against danger by recognizing pathogens or alarm signals produced by the host's damaged cells (Matzinger, 1994). Both of these models emphasize the importance of early response provided by the innate immunity. However, the difference lies in the initiation of the response. While Janeway's model relies on the exogenous signals and foreign entities for triggering the immune response, Matzinger's danger model uses endogenous signals coming from stressed or injured cells. These signals either already exist in the cells or are inducible (Matzinger, 2002).

The discovery that the Toll receptor is required for the antifungal defense response in *Drosophila* has been important for our understanding of immunity (Lemaitre et al., 1996). It has demonstrated a highly effective immune response in the absence of the adaptive immunity. This response does not rely only on phagocytosis, but also on the systemic immune response. However, the most important finding, leading to the award of 2011 Nobel Prize in Physiology or Medicine, has been the identification of mammalian Toll-like receptors (TLRs). These

receptors, involved in the innate immune response in mammals, also control immune recognition required for the initiation of adaptive immune response (Medzhitov et al., 1997; Poltorak et al., 1998).

There are nine members of the Toll receptor family in *Drosophila*, whereas in mice and humans there are 12 and 10, respectively. The best-characterized Toll receptor in *Drosophila* is Toll-1, shown to be essential in both development (Anderson, 2000; Hashimoto et al., 1988) and immune response of the fly (Lemaitre et al., 1996). In both processes, the activation of Toll is achieved by binding of proteolytically processed cytokine-like molecule Spaetzle, indicating that Toll-1 does not function as a PRR (Levashina et al., 1999). In mammals, however, TLRs are involved in the first step of innate immune response and work directly as PRRs. Each TLR has a distinct function in the process of recognizing different PAMPs and inducing immune response. PAMPs recognized by TLRs come from viruses, bacteria, fungi, and protozoa, and include lipids, lipoproteins, proteins and nucleic acids (Akira et al., 2006). Among TLRs with different specificities, TLR2 recognizes bacterial lipoproteins and peptidoglycan, TLR4 recognizes lipopolysaccharide (LPS), TLR5 - bacterial flagellin, and TLR9 - unmethylated CpG DNA typical for bacteria (Aderem and Ulevitch, 2000; Akira et al., 2001). No Spaetzle homologs have been found in the human genome, and TLRs have no developmental function in mammals (Akira et al., 2001).

The role of the innate immunity in detection of pathogens has been demonstrated in many *Drosophila* studies, highlighting the importance of this model organism.

Drosophila as a model for innate immunity

The fruit fly *Drosophila melanogaster* is a powerful model organism; its whole genome has been sequenced (Adams, 2000), and a wide array of genetic tools is available to study its gene functions *in vivo*. The extent of gene redundancy is lower than in mammals and the signaling pathways are evolutionarily conserved. The absence of adaptive immunity makes it possible to examine the signaling pathways of the innate immune system without interference. The key role of *Drosophila* as a model for studying innate immunity has been illustrated by the initial genetic identification of signaling pathways mediating antimicrobial peptide (AMP) gene expression (Lemaitre et al., 1995a, 1996).

Drosophila spends its entire life in a hostile environment, living on rotten and decaying fruit, where it coexists with different parasites and pathogens. Throughout its life cycle, *Drosophila* is exposed and can be infected by different viruses, parasites, fungi, and bacteria. In the absence of adaptive immunity, innate immune system has developed mechanisms to recognize potential pathogens and keep them at bay (e.g., antimicrobial peptides that are constitutively expressed in the epithelial tissues), as well as mechanisms for defense against invading pathogens (induced upon injury or infection). *Drosophila* has an open circulatory system, in which the blood (hemolymph) circulates throughout the whole body allowing easy spread of mediators and effectors of the immune response.

Drosophila can also be infected by vertically transmitted endosymbionts such as *Wolbachia* and *Spiroplasma*. There is no evidence that these endosymbionts

trigger the immune responses or are affected by them, but it has been shown that they promote symbiont-mediated defense. *Wolbachia* protects against some viruses whereas *Spiroplasma* protects against parasite wasps and nematodes (Chrostek et al., 2013; Xie et al., 2013).

Drosophila immune response depends on the mode of infection, the type of pathogen, route of challenge, the tissue(s) affected, developmental stage, genotype, and many other physiological parameters, including the presence of symbiotic bacteria (Neyen et al., 2014).

Drosophila defense responses

Drosophila mounts a multifaceted immune response, where physical barriers are the first line of defense. When these barriers are breached, introduction of microorganisms into the body cavity activates a strong inducible immune response that relies on three major mechanisms: the activation of proteolytic cascades and cellular and humoral reactions.

Proteolytic cascades are activated at the site of the injury, as well as around the invading microorganisms, inducing blood clotting and melanization and some immune responses.

Different types of blood cells (hemocytes) mediate the cellular response, neutralizing microorganisms by phagocytosis or encapsulation.

The humoral, or systemic, immune response involves the challenge-induced synthesis and secretion of a large set of effector molecules, including antimicrobial peptides (AMPs).

Epithelial immunity

Drosophila lives in a hostile environment, on decaying matter and fermenting fruit, where it coexists with different microorganisms. The first lines of defense are the barrier epithelia, which are in constant contact with the environment. These barriers include the external cuticle, effective against penetration, as well as chitinous matrix lining the gut and the trachea – two main routes of infection. Insect epithelia are more than passive physical barriers; tissues that are in direct contact with microorganisms actively synthesize and release AMPs. In *Drosophila*, several epithelial tissues, such as respiratory tract, oral region, digestive tract, Malpighian tubules, and male and female reproductive tracts constitutively express AMPs (Ferrandon et al., 1998; Tzou et al., 2000).

There are two complementary inducible mechanisms in the epithelia – the local production of AMPs and production of Reactive Oxygen Species (ROS). ROS are induced by bacterial infections and are directly toxic against the pathogens. If ROS are produced in excessive amounts, they are deleterious to the host. Therefore, it is necessary to maintain a balance between ROS production and elimination. This fine redox balance is achieved by the components of the antioxidative system – Duox proteins and Immune-Regulated Catalase (IRC) (Ha et al., 2005a, 2005b).

Coagulation and melanization

At the site of injury, proteolytic cascades are immediately activated to induce blood clotting (coagulation) and melanization.

The process of coagulation is important in organisms with an open circulatory system to prevent or limit the loss of hemolymph and initiate the process of wound healing (Muta and Iwanaga, 1996). Quick sealing of the wounds prevents the spread of infection. Furthermore, the newly formed clot forms secondary barrier to infection, traps and immobilizes the pathogens (Theopold et al., 2004).

In larvae, the rapidly formed clots are composed of different fibers trapping the hemocytes. Proteomic analysis has identified several proteins involved in this process, including Hemolectin, a major component of the clot, and Fondue, a protein involved in the cross linking of the fibers (Lesch et al., 2007; Scherfer et al., 2004, 2006). Hardening of the clot is caused by cross linking enzymes, including prophenoloxidase (proPO) and transglutaminase (TG) (Bidla et al., 2005; Scherfer et al., 2004). TG also enables binding to the surface of the pathogen (Wang et al., 2010). Even though proPO is involved in clotting, the process itself is independent of melanization.

Melanization involves the *de novo* synthesis and deposition of melanin, both at the site of the injury and around the invading organism. It plays a role in wound healing, encapsulation, sequestration of microorganisms, and production of toxic intermediates (Soderhall and Cerenius, 1998). The activation of the melanization

cascade, its components, and regulation will be described in more detail in the second part of this chapter.

Cellular immunity

Another aspect of immune defense mechanism is cellular immunity, which involves different types of hemocytes. Together, they participate in the clearance of invading microorganisms. In *Drosophila* larvae, hemocytes can be categorized into three types on the basis of their function (Meister, 2004). These are plasmatocytes, lamellocytes, and crystal cells; they govern the mechanisms of cellular immunity phagocytosis, encapsulation, and melanization, respectively.

Plasmatocytes are phagocytic cells constituting 90-95% of the mature larval hemocytes. Through the process of phagocytosis, they are involved in the removal of apoptotic cells (Franc et al., 1996) and different microorganisms (Pearson et al., 2003). The process of phagocytosis represents a primordial aspect of innate immunity. Phagocytosis includes the attachment of the phagocyte to the target cell or microorganism, cytoskeletal modifications, internalization, and destruction of the ingested material in the phagosome. In *Drosophila*, the importance of this process has been highlighted by the discovery of the phagocytic receptor Eater, which contains many Epithelial Growth Factor (EGF)-like repeats. Eater is expressed on the surface of plasmatocytes. Upon infection, it binds and helps to internalize a wide range of bacteria (Kocks et al., 2005).

Encapsulation is a defense reaction against invading parasites that are too large to be phagocytized. It is mediated by lamellocytes that are large, flat, adherent cells. It can be induced by parasitic wasps (Russo et al., 1996). Wasp eggs are detected by plasmatocytes, which attach to the egg chorion, inducing strong cellular reactions in the lymph gland. This induction leads to an increase in the proliferation and differentiation of lamellocytes (Jung et al., 2005). Lamellocytes are released from the lymph gland and form capsules around the injected wasp eggs; this process is followed by melanization and blackening of the capsule. Parasites are killed by the local production of cytotoxic ROS and intermediates of the melanization process (Nappi et al., 1995, 2000).

Crystal cells represent 5% of larval hemocytes. They have no phagocytotic properties and are involved in the process of melanization. The activation of melanization reaction, its components, and regulation will be described in more detail in the second part of this chapter.

Antiviral immunity

Besides bacteria and fungi, viruses are important pathogens of insects. Several RNA and DNA viruses belonging to different families can replicate in *Drosophila* and are useful models to decipher antiviral immunity in insects (Merkling and van Rij, 2013; Xu and Cherry, 2014). Viral infection triggers different types of processes including RNA interference (Nayak et al., 2013) and other inducible responses (Lamiable and Imler, 2014).

The small interfering RNA (siRNA) pathway plays a major role in the control of viral infection in *Drosophila*; it is required for restricting both RNA and DNA viruses.

Inducible responses are virus-specific and include several processes such as apoptosis, autophagy, JAK/STAT and NF-*k*B signaling pathways.

Apoptosis is triggered during infections by baculovirus and Flock House Virus, a DNA virus and an RNA virus, respectively (Liu et al., 2013). Autophagy is required to restrict two negative single-stranded RNA viruses, Vesicular Stomatitis Virus, and Rift Valley Fever Virus (Shelly et al., 2009; Moy et al., 2014). JAK/STAT pathway is involved in antiviral response against *Dicistroviridae* (DCV and CrPV), but not against other viruses (Kemp et al., 2013). The IMD and Toll pathway seem to be required; some null mutants for components of these two pathways are sensitive to viruses (Avadhanula et al., 2009; Costa et al., 2009; Rancès et al., 2013; Zambon et al., 2005). In addition to the components of these pathways, hundreds of genes specifically induced during viral infection have been identified using microarrays. For the large majority of this gene set, their function and biological roles are still unclear.

Humoral immunity

The hallmark of the *Drosophila* immune response is the systemic response involving the challenge-induced synthesis and secretion of AMPs. The fat body, which is the analog of the mammalian liver, synthesizes the AMPs that are then secreted into the hemolymph. This antimicrobial activity persists for several days and confers protection against the second challenge (Boman et al., 1972). Large-

scale analyses performed on transcriptome and proteome levels have shown that along with AMPs, many different peptides and proteins are upregulated upon septic injury (Boutros et al., 2002; De Gregorio et al., 2001; Irving et al., 2001; Levy et al., 2004a, 2004b).

Among these various immune effectors, AMPs are the best characterized. AMPs are found in evolutionarily diverse organisms, from prokaryotes and invertebrates to vertebrates and plants. These effectors are small (<10 kDa, except for the 25 kDa Attacin), cationic, and have a broad range of activities against bacteria and/or fungi (Imler and Bulet, 2005). Their mechanism of action involves the disruption of the membrane and/or formation of pores, helping in rapid killing of the pathogens. Seven classes of AMPs have been identified in *Drosophila*; they are divided into three groups based on their activity. Attacins, Cecropins, Diptericins, and Drosocins are effective against Gram-negative bacteria; Defensins work against Gram-positive bacteria, and Drosomycins and Metchnikowins, against fungi. Combined concentration of AMPs in the hemolymph of infected flies can reach 300 M. Unfortunately, loss-of-function mutants for AMP genes are still not available, and the exact contribution of each of these genes cannot be estimated.

NF-κB-dependent activation of systemic immune response

The humoral, or systemic, immune response triggers dramatic changes in the expression of genes responsible for the synthesis of different peptides and molecules, including AMPs. Upon septic injury, the expression of various genes is

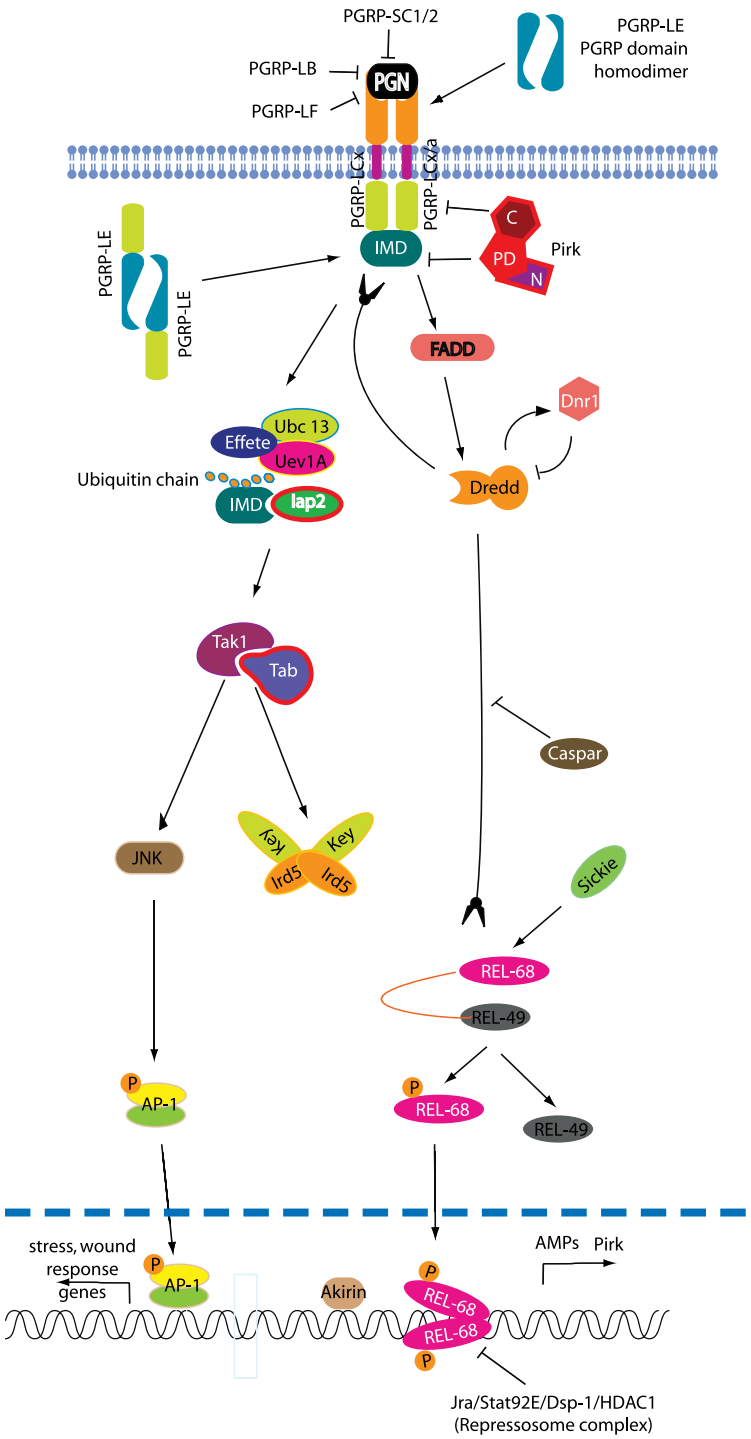


Figure 1 *Drosophila* IMD signaling pathway

Homologies between signaling components are depicted by similar shape. The Imd pathway is activated by DAP-type PGN binding of the PGRP-LC dimer. Other PGRP family members play either negative or positive roles. IMD is connected to the caspase DREDD via the adaptor protein Fas-associated DD protein (FADD). DREDD proteolytically cleaves IMD and Relish. Cleaved IMD associates with the E3-ligase IAP2, E2-ubiquitin-conjugating enzymes UEV1a, Bendless (Ubc13), and Effete (Ubc5) and is K63 polyubiquitinated. This activates the downstream kinase cascade leading to the phosphorylation and activation of Relish and AP-1, which activate the transcription of AMP and stress genes, respectively. Akirin is required for Imd pathway function at the level of Relish. Pirk, Caspar, and Dnr1 are negative regulators of the Imd pathway (adapted from Valanne et al., 2011).

primarily regulated at the transcriptional level and Nuclear Factor kappa B (NF-κB)-binding sites in the AMP gene promoter regions (Engström et al., 1993; Kappler et al., 1993; Meister et al., 1994). NF-κB family of transcription factors have a central role in expression of the genes that control immune responses (Li and Verma, 2002). There are three NF-κB/Rel-like proteins encoded in the *Drosophila* genome – Dorsal, Dif, and Relish (Dushay et al., 1996; Ip et al., 1993; Reichhart et al., 1993; Steward, 1987). Genetic studies have demonstrated the role of these proteins in the expression of AMPs via two distinct signaling pathways, the IMD and Toll pathway.

The identification of the two signaling pathways has demonstrated that *Drosophila* discriminates between different classes of pathogens and develops specific responses. Gram-negative bacteria mostly induce the IMD pathway, whereas Gram-positive bacteria and fungi induce the Toll pathway. There is a correlation between the induced AMPs levels and resistance to infections. The expression of *diptericin* and *drosomycin* transcripts can be used as a read out for IMD and the Toll pathways, respectively.

The IMD pathway was identified by the mutations in the *immune deficiency (imd)* gene, which impair the expression of different antibacterial peptides (Corbo and Levine, 1996; Lemaitre et al., 1995b; Levashina et al., 1998). The IMD pathway shows some similarities to the vertebrate Tumor Necrosis Factor Receptor (TNF-R) pathway. In particular, *imd* encodes a death domain-containing protein that is similar to the Receptor Interacting Protein (RIP) of the TNF-R pathway (Georgel et al., 2001). The IMD pathway and its components are depicted in the Figure 1.

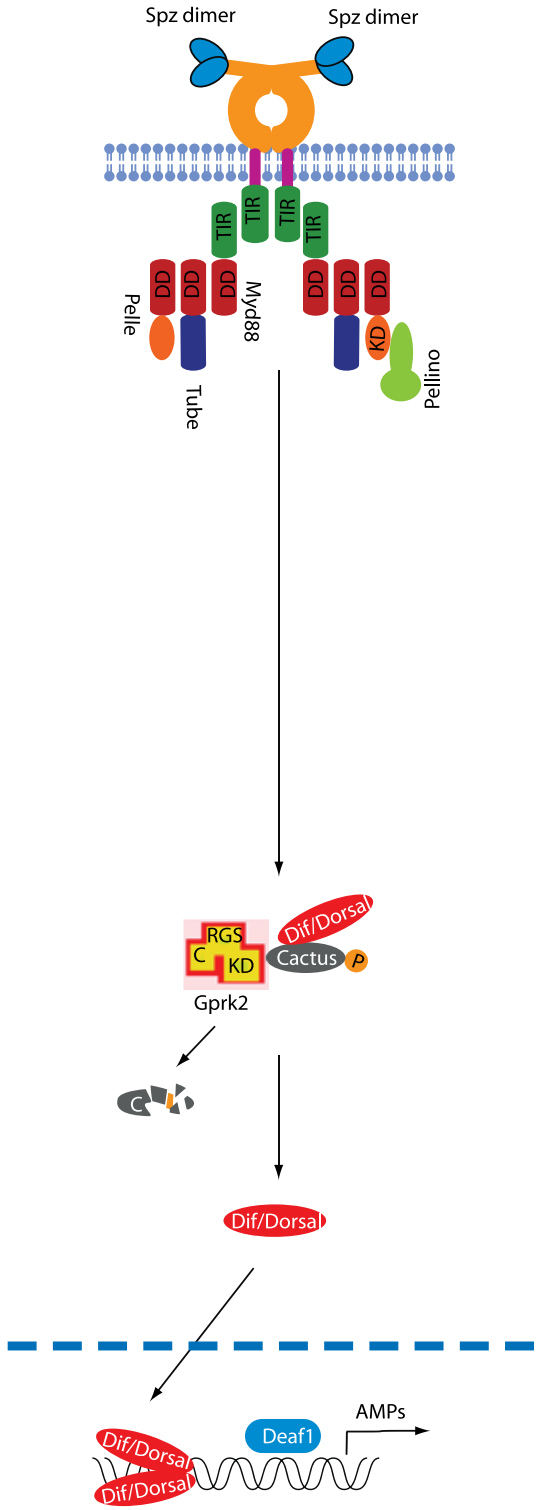


Figure 2 *Drosophila* Toll signaling pathway

Homologies between signaling components are depicted by similar shape. The Toll pathway is activated by Spz. One Spz dimer is depicted to bind the N terminus of Toll and to induce a conformational change leading to the formation of a 4Spz:2Toll complex. Intracellular signaling leads to the phosphorylation and degradation of Cactus, which releases Dif and/or Dorsal to translocate to the nucleus and activate transcription. Gprk2 associates with Cactus in a kinase domain (KD)-dependent manner. DEAF-1 is required to induce Toll pathway target genes at or downstream of Dif/Dorsal. (adapted from Valanne et al., 2011).

The Toll pathway is an evolutionarily conserved signaling pathway involved both in embryonic development and the adult immune response of *Drosophila* (Belvin and Anderson, 1996). Significant similarities have been observed between this signaling cascade and downstream cascade of Interleukin-1 (IL1) and TLRs, implying a common ancestry (Belvin and Anderson, 1996). These similarities and the identification of NF-κB-binding sites in the promoter of *drosomycin* gene prompted genetic studies which demonstrated that some AMPs, such as Drosomycin, are not induced in Toll mutant flies. Resistance to immune challenge requires wild type copies of the signaling components (Lemaitre et al., 1996). The core components consist of a cytokine-like molecule Spaetzle (Spz), the transmembrane receptor Toll, the adaptors Tube and Myd88, the Pelle kinase, NF-κB inhibitor Cactus, and the transactivators Dorsal or Dif (Belvin and Anderson, 1996; Tauszig-Delamasure et al., 2002). The intracellular Toll pathway is depicted in Figure 2.

The Toll receptor does not function directly as a pattern recognition receptor, but needs to be activated by binding of a proteolytically processed ligand, Spz (DeLotto and DeLotto, 1998a; Hu et al., 2004; Lemaitre et al., 1996; Mizuguchi et al., 1998; Weber et al., 2003).

Spz is synthesized and secreted as a pro-protein containing a signal peptide, an N-terminal domain, and a cysteine-rich C-terminal domain. Activation of Spz is achieved by cleavage at a specific site by the terminal protease in the Toll cascade (Easter during embryonic development and Spaetzle Processing Enzyme (SPE) in the adult immune response). This proteolytic activation generates an active fragment, C-106, that binds to The Toll receptor (although the C-106 fragment remains

associated with the Spz N-terminal domain). Binding of Spz-C-106 to the Toll receptor involves formation of a processed Spz dimer, which cross links two Toll receptor molecules (Weber et al., 2003, 2007). Upon binding to the Toll receptor, the N-terminal part of Spz is released, generating C-Spz (activator) and N-Spz (inhibitor). During early development, the Toll pathway is activated ventrally to establish the dorsoventral axis of the embryo. Recent studies based on mathematical models have demonstrated that N-Spz and C-Spz can reassociate, forming a distinct complex that does not activate Toll (Haskel-Ittah et al., 2012). Thus, each processed Spz fragment has a different role in regulating the Toll cascade, creating a self-organized shuttling mechanism during early embryonic development. This mechanism generates a sharp and robust patterning gradient within a uniform region, leading to Toll activation.

The Toll signaling pathway in early development forms a part of the mechanisms for setting up the dorsoventral axis of the embryo, while in the adult this signaling pathway regulates the immune response. Although the processing of Spz is similar in both signaling cascades, different sets of serine proteases are involved.

2. Serine proteases and serine protease homologs

Classification and mode of action

Proteolytic enzymes are found in all organisms; these enzymes are involved in a large number of different physiological processes. They are classified on the basis of the structure and sequence similarities (MEROPS database). This classification distinguishes between clans (based on the protease catalytic mechanism) and families (based on protease common ancestry) (Rawlings and Barrett, 1999). The catalytic mechanism depends on the functional groups occupying the catalytic site for substrate hydrolysis and defines the clans as aspartic, asparagine, cysteine, glutamic, metallo-, threonine, and serine proteases (SP). More than a third of all proteolytic enzymes are SPs (Page and Di Cera, 2008).

In SPs, the functional group in the catalytic site is a nucleophilic serine (Ser) residue that attacks the carbonyl moiety of the substrate peptide bond leading to the formation of an acyl-enzyme intermediate (Page and Di Cera, 2008). The crystal structure of bovine chymotrypsin demonstrates that the catalytic site is composed of a catalytic triad consisting of His57, Asp102, and Ser195, which are responsible for the acyl-transfer mechanism. The catalytic triad operates in concert, where the residues form two dyads, Ser-His and His-Asp (Perona and Craik, 1995). To be hydrolyzed, the scissile bond of the peptide is inserted into the catalytic site of the enzyme. This allows for the nucleophilic attack on the carbonyl carbon in the peptide

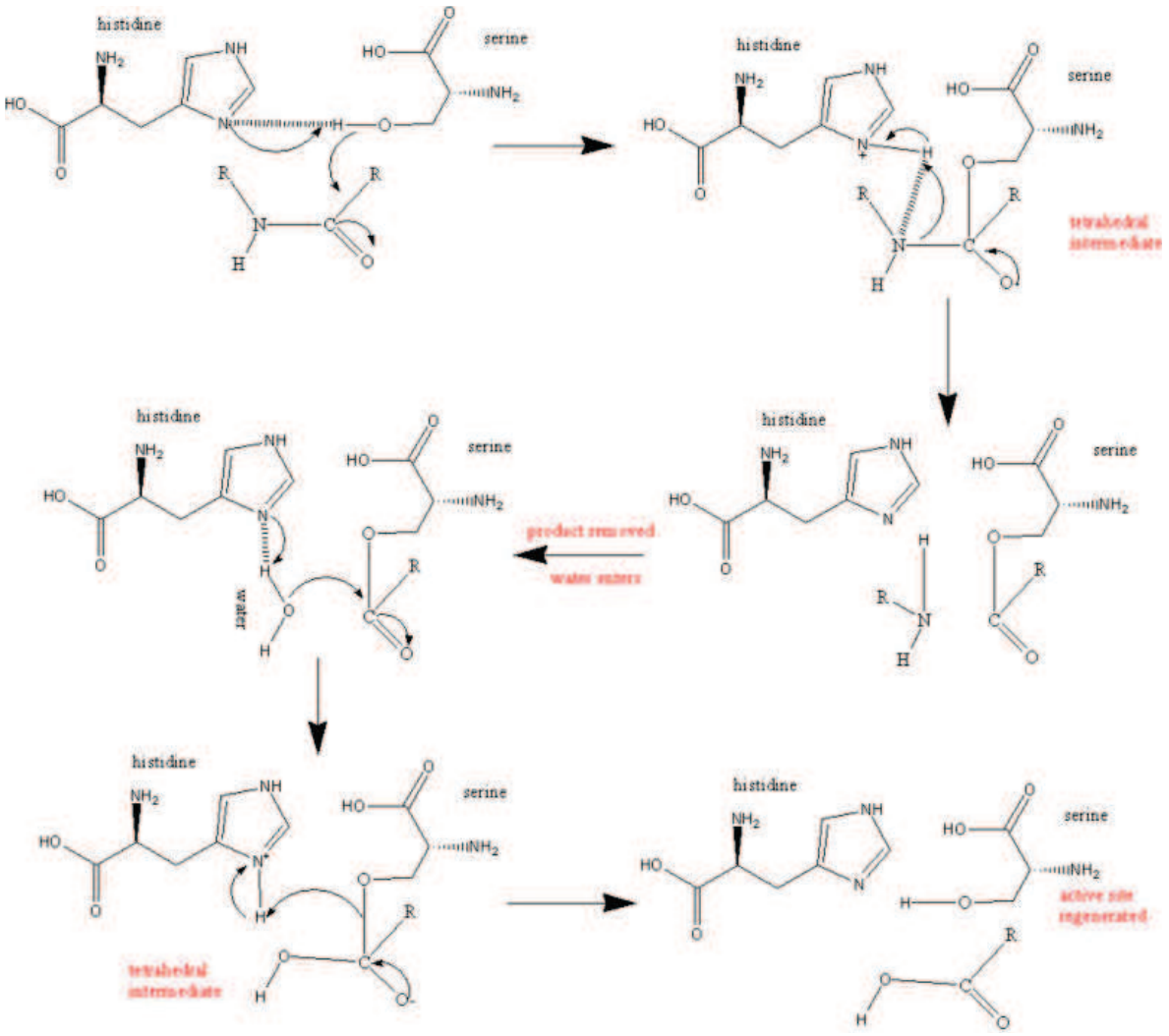


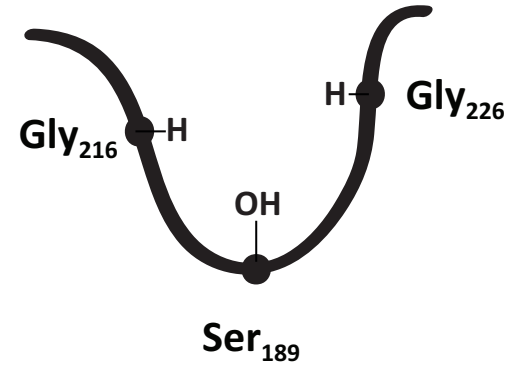
Figure 3 Acyl Transfer mechanism in serine proteases (figure taken from http:// en.wikipedia.org/wiki/Serine_protease)

by the serine hydroxyl (-OH) group. This is assisted by His57 yielding a tetrahedral intermediate. This intermediate then collapses, generating the acyl-enzyme intermediate and stabilizing newly created N-terminus. In the second stage of the mechanism, a water molecule attacks the acyl-enzyme intermediate. This step yields a second tetrahedral intermediate liberating a new C-terminus in the substrate (Hedstrom, 2002; Page and Di Cera, 2008) (Figure 3).

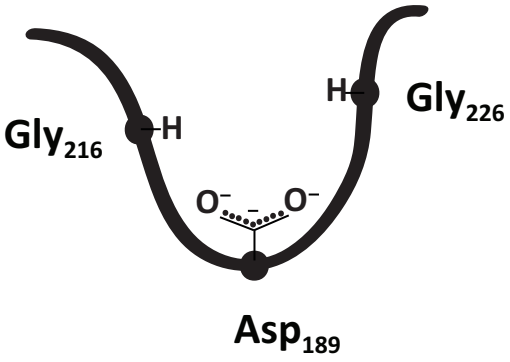
The specificity between the protease and its respective substrate is based on the P1/S1 interaction. This is the Schechter & Berger nomenclature, where P1-P1' represents the peptide residues of the scissile bond that is to be cleaved, and S1, S1', etc., are the corresponding enzyme binding sites. The S1 site is a pocket (specificity pocket) adjacent to Ser195, with residues determining the specificity of the protease (Hedstrom, 2002). SPs are commonly synthesized as inactive zymogens and converted to the active enzyme by a proteolytic cleavage. This cleavage occurs at the P1-P1' scissile bond, releasing the N-terminal and inducing a conformational change leading to the formation of S1 and activation of the protease.

The shape of the binding site and the identity of the peptide residues within S1 specificity pocket, between them determine the substrate specificity of proteases. Four types of SPs can be distinguished (Perona and Craik, 1995) (Figure 4):

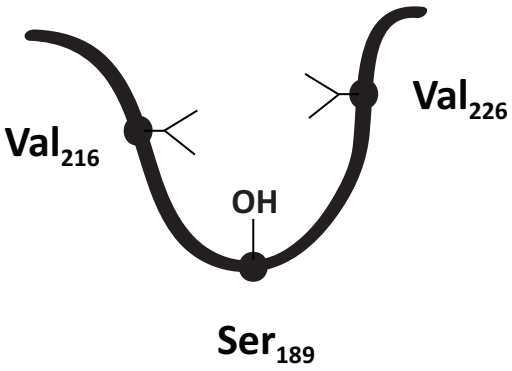
- Chymotrypsin-like SPs have a hydrophobic specificity pocket, giving specificity for hydrophobic amino acid (aa) residues such as phenylalanine, tyrosine and tryptophan.



Chymotrypsin-like SPs



Trypsin-like SPs



Elastase-like SPs

Figure 4 Binding pocket and amino acid residues that determine the substrate specificity of proteases

- The specificity of trypsin-like serine SPs is driven by negatively charged moieties; this results in specificity for positively charged aa residues such as lysine or arginine.
- Elastase-like SPs have much smaller specificity pocket than other proteases, resulting in specificity for aa residues such as alanine, glycine and valine.
- Subtilisin-like SPs have broad specificities and are found only in prokaryotes. They are evolutionarily unrelated to chymotrypsin-like SPs but share the same catalytic mechanism of activation.

In addition to the catalytic domain, many SPs contain different regulatory domains within their N-terminal region (for example, CLIP domain discussed later in this chapter). This region is connected to the catalytic domain via a linker region. During the zymogen activation, a cleavage occurs within the linker domain, liberating the N-terminal domain and releasing the catalytic activity of the protein. In many cases, the N-terminal domain remains attached to the catalytic domain forming covalent disulfide bonds. Through specific protein-protein interactions achieved via N-terminal domains, zymogens can form a cascade in which one protease activates the zymogen of another and mediates a rapid, local reaction (Ross et al., 2003).

Serine protease homologs (SPHs) have similar amino acid sequences as SPs, but they lack the amidase activity – one or more catalytic residues are missing. SPHs are found in both invertebrates and vertebrates, but their physiological functions are still poorly understood (Ross et al., 2003).

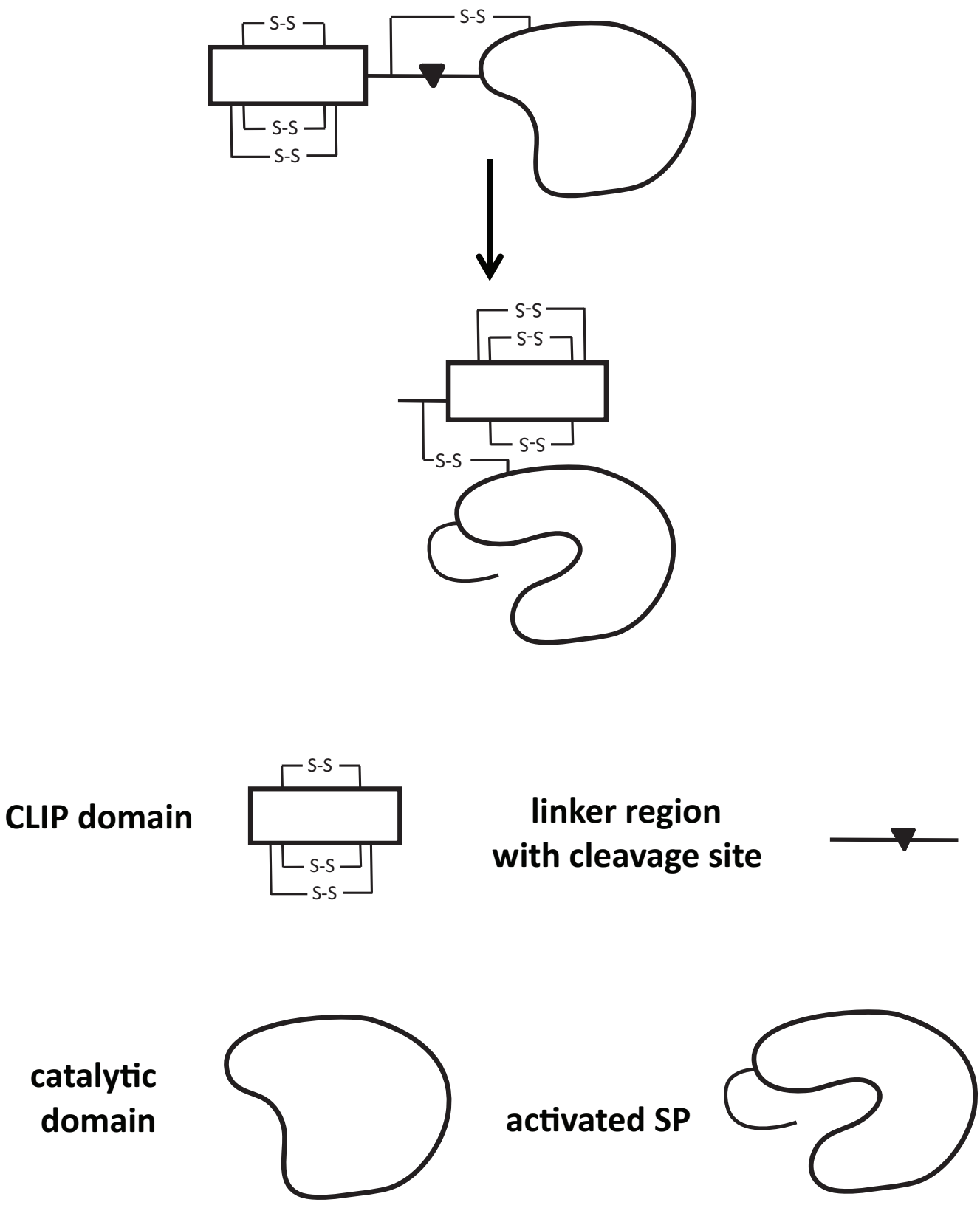


Figure 5 Domain organization and activation of CLIP-domain protease

CLIP domain-containing serine proteases

Some arthropod SPs and SPHs involved in immune defense mechanisms have a disulfide-bridged structure named the CLIP domain within their N-terminal domain (Jiang and Kanost, 2000). The CLIP domain was first described in the study of a pro-clotting enzyme found in the horseshoe crab, *Holotrichia diomphalia* (Muta et al., 1990). The name stems form a "paper clip"-like configuration of this domain, formed by disulfide bridges.

CLIP domains are composed of 30-60 amino acids that are not highly conserved, except the six cysteine (Cys) residues forming disulfide bridges. Three disulfide bridges are formed within the CLIP domain itself. One Cys residue is found in the linker region (sequence connecting the N-terminal and the C-terminal domain), forming an inter-domain disulfide bridge with the Cys residue of the catalytic domain. The CLIP domain is found in many arthropod SPs and SPHs involved in cascade pathways (Jiang and Kanost, 2000). The function of a CLIP domain is still not fully understood. It may be involved in shielding the activation site of the zymogen or anchoring the SP (or SPH) to the surface of the invading organism (Jang et al., 2008; Ross et al., 2003). Furthermore, upon activation of the zymogen, the CLIP domain remains attached to the C-terminal domain, where it can interact with associated proteins through specific protein-protein interactions (Figure 5). In *Drosophila*, there are 24 CLIP domain-containing SPs (c-SPs) and 16 SPHs (c-SPHs). Some of the genes contain multiple or partial CLIP domains.

The involvement of CLIP-SPs in the activation of the Toll pathway was first identified in the protease regulating the formation of the dorsoventral axis in *Drosophila* embryo (Chasan and Anderson, 1989; Chasan et al., 1992; Stein and Nüsslein-Volhard, 1992). Two CLIP domain-containing SPs, Easter and Snake, are involved in the proteolytic processing of the Toll receptor ligand Spz (DeLotto and DeLotto, 1998b). Proteolytic cascade that leads to Spz processing during development will be described in part three of this chapter. Studies of the activation of the Toll pathway in the adult have shown that these embryonic SPs are dispensable during the immune response of *Drosophila.* This result suggested the existence of other SPs in the proteolytic cascade leading to Toll activation during the immune response (Lemaitre et al., 1996). To date, several CLIP domain-containing SPs have been demonstrated to be involved in the proteolytic cascade leading to Toll activation during the immune response of *Drosophila*; they will be discussed in part three of this chapter.

One of the SPHs implicated in the immune response, Spheroide, will be discussed in part two of the Results chapter.

SPs and SPHs in *Drosophila* and other arthropods

Studies of sequence conservation and phylogenetic relationships have established SPs and SPHs as the second largest family of genes encoded in *Drosophila melanogaster* genome (Ross et al., 2003). The number of SP-related

genes in *Drosophila* genome is 211 (148 SPs and 63 SPHs), in comparison with only one in *Saccharomyces cerevisiae* and 13 in *Caenorhabditis elegans* (Rubin, 2000).

SPs and SPHs of *Drosophila* and other arthropods are involved in various defense mechanisms such as hemolymph coagulation, melanotic encapsulation, induction of AMP synthesis, and activation of cytokines (Jiang and Kanost, 2000). SPs involved in the activation of Toll pathway during the development and immune response in *Drosophila* will be discussed in part three of this chapter.

SPHs have been implicated in various physiological processes. In 1991, Hogg et al. reported that a mammalian serine protease homolog, protein Z (PZ), a vitamin K-dependent glycoprotein, binds to thrombin causing its conformational changes. When bound to PZ, thrombin is associated with phospholipid membrane vesicles. This membrane localization is important during coagulation and clotting as it partitions thrombin to the site of an injury (Hogg and Stenflo, 1991). Later studies of protein crystal structure have demonstrated that PZ functions as a cofactor regulating proteolytic activity of Factor Xa (FXa) on phospholipid vesicles (Qureshi et al., 2014). This is achieved through interaction with FXa and PZ-dependent Protease Inhibitor (ZPI), which forms a serpin/protease complex with FXa. PZ has an NH_2-terminal domain-containing γ-carboxyglutamic acid (Gla) domain and two EGF-like domains. PZ and ZPI interact via the sites on the inactive catalytic domain, but also within the EGF2-like domain. Protein-protein interactions between PZ and FXa take place through Gla-domains. This interaction is necessary for the assembly of a protein complex on the phospholipid vesicle surface, which leads to the formation of an effective inhibitory complex containing PZ/FXa/ZPI. One reported serine protease

homolog in *Drosophila*, Masquerade, is necessary during embryonic development to promote and/or stabilize cell-matrix interactions.

Masquerade has a CLIP domain in the NH_2-terminal region of the protein. One model for Masquerade function proposes that the CLIP domain mediates protease interactions by promoting cell-substrate adhesion. Another proposed model is based on the fact that Masquerade does not have amidase activity. As an inactive protease homolog, Masquerade might compete for a substrate with the active serine protease, indirectly stabilizing cell-substrate interactions (Murugasu-Oei et al., 1995). A cell-adhesion molecule similar to *Drosophila* Masquerade has been reported in the crayfish, *Pacifastacus leniusculus*, where it is found in adult hemocytes. This molecule also contains a CLIP domain, which might have antimicrobial properties apart from its function in protein-protein interactions. The sequence of the CLIP domain is similar to the motif found in defensin of a horseshoe crab, *Tachypleus tridentatus* (Huang et al., 2000). In a later study, it has been found that this antibacterial property is specific to Gram-negative bacteria (*Escherichia coli*, *Shigella flexneri*) and yeast (*Saccharomyces cerevisiae*) (Lee and Soderhall, 2001). Another serine protease homolog with antibacterial properties is limulus factor D, found in the granular cells of the horseshoe crab hemolymph. This protein is released upon bacterial stimulation; it displays antibacterial activity against Gram-negative bacteria (*E. coli*, *Salmonella typhimurium*, *Salmonella minnesota*). It is similar to the human serine protease homolog azurocidin, found in azurophil granules, which also has antibacterial properties (Kawabata et al., 1996). Two SPHs taking part in the immune response of mosquito *Anopheles gambiae* have been

described. Dimopoulos et al. have shown that an immune-responsive marker, Immune-related Serine Protease-Like sequence 5 (ISPL5), is induced upon challenge with Gram-negative bacterium (*E. coli*) (Dimopoulos et al., 1997). However, biochemical and cellular functions of ISPL5 have not been elucidated. Another SPH found in mosquito is Serine Protease CLIP1 (SPCLIP1). Upon infection of the mosquito with *Plasmodium berghei*, complement C3-like protein TEP1 is required for recruitment of SPCLIP1, enabling its binding to the surface of ookinetes. Furthermore, SPCLIP promotes the activation of proPO cascade that leads to the melanization of the parasite (Povelones et al., 2013).

A recent report has described a serine protease homolog, SPH3, without a CLIP domain that is involved in the immune response of the moth *Manduca sexta* against infection with the Gram-negative bacterium *Photorhabdus luminescens*. SPH3 was initially identified as a target for Repeats-in-Toxin (RTX) metalloprotease, protease A (PrtA), secreted by the bacterium. Upon infection, SPH3 is upregulated in the fat body and hemocytes. Moths with RNAi-mediated knockdown of SPH3 succumb to infection to a higher extent than wild type animals. Furthermore, in such moths, the levels of effector molecules such as antimicrobial peptides and prophenoloxidase (PPO) are reduced. Conversely, mRNA and protein levels of recognition molecules are not changed in these animals. This implies the existence of two signaling pathways, one that governs the recognition, and the other the transcription of effector molecules. On the basis of their findings, the authors have proposed that SPH3 controls the effector molecules in the signaling pathway (Felföldi et al., 2011).

Regulation by Serine Protease Inhibitors (SERPINs)

Proteolytic cascades are involved in a wide range of physiological processes mediating rapid local responses. A proteolytic cascade is formed via sequential activation of zymogens. To prevent unnecessary activation, serine proteases need to be tightly regulated.

Proteases are regulated by protease inhibitors. These inhibitors can have either a broad inhibitory effect or can be protease-specific. There are two major types of protease inhibitors, tight-binding and trapping inhibitors (Rawlings et al., 2004).

Tight-binding inhibitors undergo strong but reversible interactions with proteases via a lock-and-key mechanism. Trapping inhibitors undergo a conformational change upon peptide cleavage of the protease. There are two families of trapping inhibitors, the Serine Protease Inhibitors (Serpins) and the macroglobulins. The serpins form irreversible, covalently linked complexes with their target proteases. In the case of macroglobulins, the catalytic site of the target protease is sequestered within the inhibitor and remains intact.

In the *Drosophila* genome, there are 29 genes encoding serpins that are involved in the regulation of different proteases. Of particular interest are the serpins that have immune-related functions, such as Spn43Ac (Necrotic (Nec)) controlling Toll pathway activation (Levashina et al., 1999), Spn27A and Spn28Dc implicated in PO activation cascade (Ligoxygakis et al., 2002a; Scherfer et al., 2008),

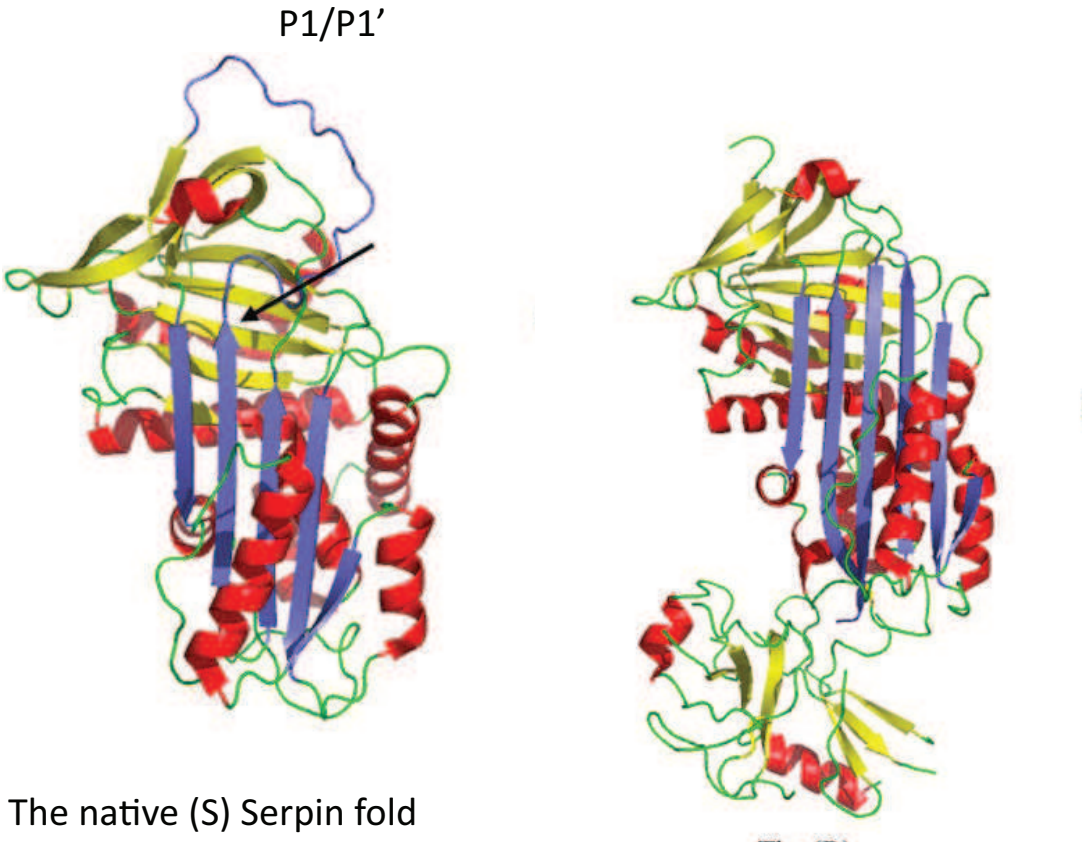


Figure 6

The native (S) serpin structure carries an exposed RCL, which is cleaved between the P1 and P1’ sites. Inhibitory serpins have a flexible hinge region, consisting of amino acids with short side-chains (arrow). Following cleavage, the protease is translocated and crushed against the bottom of the (R) serpin as a denatured covalent complex as the RCL inserts between β-sheet A (purple). (Gubb et al. 2010)

and Spn77Ba, involved in the tracheal melanization (Tang et al., 2008). Spn27A is also involved in the proteolytic cascade leading to Toll activation during embryonic development (Hashimoto et al., 2003; Ligoxygakis et al., 2003).

Serpins consist of 3 β-sheets with 8-9 α-helical linkers and an exposed Reactive Center Loop (RCL). RCL is composed of 20 aa and is bait for the target protease, providing specificity. In the native state, serpins are in a metastable or stressed conformation (S). Upon cleavage by the protease at the P1-P1' position of the RCL the serpin conformation changes to its relaxed conformation (R), which traps the protease in a covalent complex with the serpin. The protease translocates from the top to the bottom pole of the serpin and becomes denaturated. This denaturated serpin/protease complex is targeted for degradation (Huntington et al., 2000). (Figure 6)

Investigating potential target SPs of the serpins will provide us with better understanding of the mechanisms and interactions that underlie proteolytic cascades involved in the activation of the Toll pathway.

Melanization cascade

One of the major proteolytic cascades involved in the immune response in arthropods is the process of activation of prophenoloxidase (proPO, PPO), leading to melanization (Soderhall and Cerenius, 1998). The melanization reaction is the result of the oxidation process converting monophenols or diphenols to quinones, which

polymerize to form melanin. Melanin is deposited at injury sites and around intruding microorganisms, forming a physical barrier. A key enzyme in melanin biosynthesis is phenoloxidase (PO), synthesized in the form of zymogen (proPO, PPO), and activated by a proteolytic cascade. This cascade consists of CLIP domain SPs, called proPO-Activating Factors or Enzymes (PPAFs or PPAEs), also known as proPO-Activating Proteins (PAPs) (Jiang and Kanost, 2000; Lee et al., 1998). The cascade is triggered by the recognition of microbial cell wall components.

The melanization cascade is regulated by several serpins. The first serpin to be identified in the regulation of PO activation was Serpin 27A (Spn27A). Flies with loss-of-function mutations in *Spn27A* show spontaneous melanization, whereas over expression of *Spn27A* suppresses PO activation induced by microorganisms (De Gregorio et al., 2002; Ligoxygakis et al., 2002a). Another serpin involved in the melanization reaction is Serpin 28D (Spn28D). Flies mutant for *Spn28D* show spontaneous melanization in various tissues, especially the tissues in contact with air, such as tracheae (Scherfer et al., 2008). In the presence of microorganisms, the process of melanization is also activated in tracheae and regulated by another serpin, Serpin 77Ba (Spn77Ba) (Tang et al., 2008).

There are three genes in *Drosophila* genome encoding for PPO. PPO1 and PPO2 are found in crystal cells, whereas PPO3 is found in lamellocytes (and functions during encapsulation) (Irving et al., 2005). PPO1 and PPO2 are synthesized in their inactive zymogen form. Upon injury and microbial recognition, crystal cells are ruptured, and inactive zymogens are released into the hemolymph (Bidla et al., 2007). A recent study using deletion mutants for PPO1 and PPO2 has demonstrated

that PPO2 is a component of the crystal cells and is stored within those cells, whereas PPO1 is probably released into the hemolymph (Binggeli et al., 2014). Further analysis following the kinetics of melanization has demonstrated reduced melanization in *PPO1* mutant larvae, whereas *PPO2* mutants have developed stronger melanization in comparison with wild type controls. No melanization has been observed in double mutants, indicating that both PPO1 and PPO2 contribute to the melanization process. The authors suggest that PPO1 is required for the rapid delivery of phenoloxidase and PPO2 found in crystal cells is in storage, necessary for the second phase (Binggeli et al., 2014).

PPO1 and PPO2 require proteolytic cleavage of the zymogen to be activated. CLIP-SPs in the proteolytic cascade activating melanization are Melanization Protease 1 and 2 (MP1 and MP2/sp7/PAE1) (Tang et al. 2006; Leclerc et al. 2006; Castillejo-Lopez and Hacker 2005). Genetic studies have demonstrated that MP2 acts upstream of MP1. However, the response to microbial infection seems to be associated with two cascades, one involving MP2, triggered by fungal infections, and the other, involving MP1 and triggered by both bacteria and fungi (Tang et al., 2006). Indeed, when cotransfected with MP1 zymogen, MP2 does not activate MP1 in *Drosophila* S2 cells. This implies that there is another protease acting upstream of MP1. Biochemical studies using the beetle *Holotrichia diomphalia* have found that a serine protease homolog PPAF-II is required as a cofactor in PPO activation by PPAE (Kim et al., 2002). This result suggests that a similar protein might be involved in melanization cascade in *Drosophila*, acting between MP2 and MP1. It has been

proposed that MP1 might be a PPAE equivalent in *Drosophila*, the terminal protease in the melanization reaction and putative target of Spn27A.

However, a recent study has suggested the existence of multiple PPAEs as MP1 and MP2 are activated by different pathogens. Using recombinant proteins of MP2 and Spn27A, An et al. have demonstrated that MP2 directly cleaves and activates PPO1, indicating its role as a PPAE. Furthermore, they have reported the formation of a covalent complex between MP2 and Spn27A, suggesting that Spn27A might inhibit MP2 activity (An et al., 2013).

3. Activation of the Toll pathway

Activation of the Toll pathway is achieved through proteolytic cascades leading to the processing of Toll ligand Spz.

Activation of the Toll pathway during development

Activation of the Toll pathway in the establishment of the embryonic dorsoventral axis is regulated by binding of proteolytically processed Spz to the Toll receptor as in the immune response. The proteolytic cascade handling this processing in the embryo consists of the SPs Nudel, Gastrulation Defective (GD), Snake, and Easter. Different genetic and biochemical studies have demonstrated the order of the protease in cascade activation. GD acts upstream of Snake, which in turn cleaves Easter. The activated terminal protease Easter then processes Spz into its activated ligand form (Chasan and Anderson, 1989; Chasan et al., 1992; Stein and Nusslein-Volhard, 1992). The cascade is regulated by Spn27A whose target protease is Easter (Ligoxygakis et al., 2003).

Easter and Snake are CLIP domain-containing SPs, structurally similar to trypsin-like SPs. Proteolytic processing occurs at the ventral side of the embryo. This is necessary because Snake, Easter, and Spz are freely diffusible in the extracellular perivitelline space (PVS). This is achieved by the expression of ventrally restricted factor Pipe in the somatic follicle cells. As *pipe* encodes a heparan 2-0-

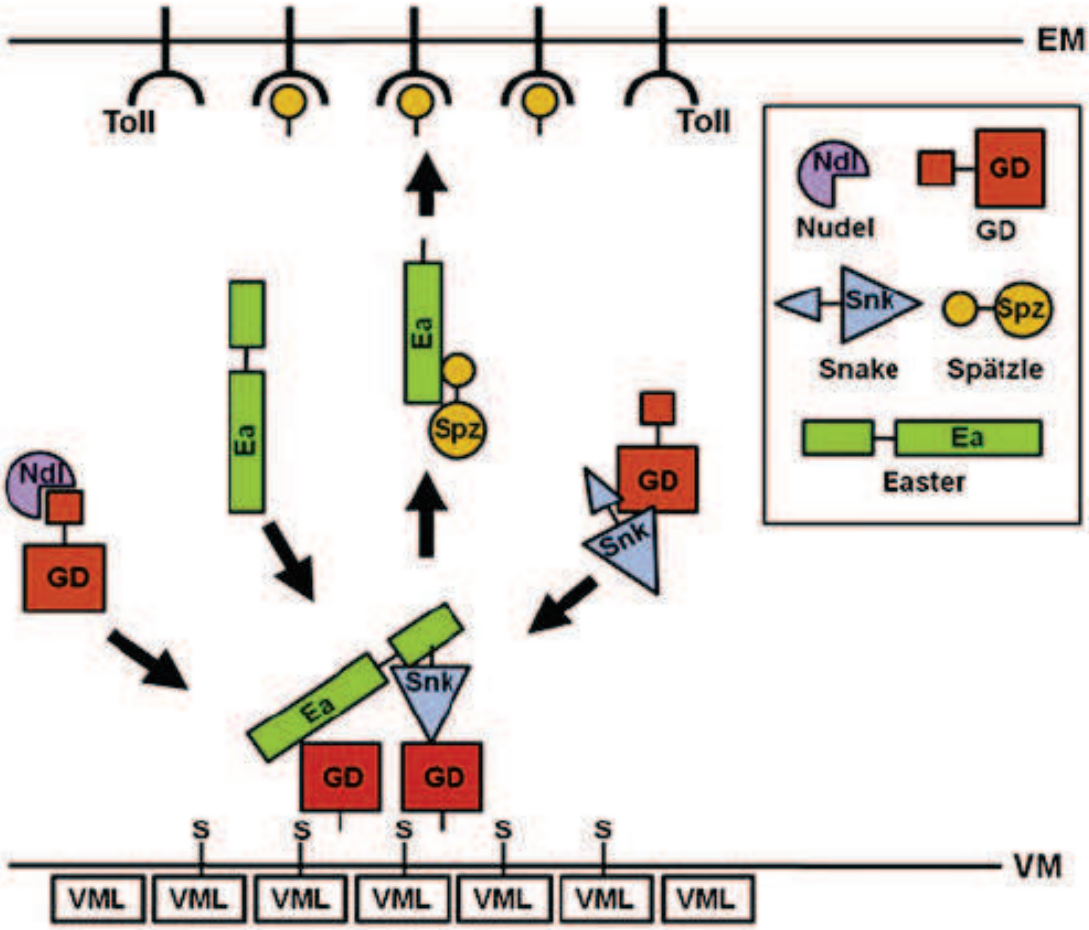


Figure 7 Proteolytic cascade during development

Model for ventral processing of Ea by Snk. GD (red) is processed by Nudel (purple). GD cleaves Snk (blue). Processed GD binds to Pipe-sulfated (S) proteins in the ventral VM, including vitelline membrane-like (VML) [11]. Bound GD recruits and brings together Snk and Ea zymogen (green), resulting in Ea cleavage. Processed Ea cleaves Spz (yellow) to form the active Toll ligand, which binds and activates Toll in the ventral embryonic membrane (EM) (Cho et al., 2012).

sulfotransferase, such sulfate modifications may provide the ventral cue in the embryo. Both GD and Snake bind to heparin-Sepharose, implying that their activity *in vivo* can be regulated by sulfated proteoglycans (Dissing et al., 2001).

Nudel is a modular SP with a central protease catalytic domain combined with other conserved structural motifs. The glycoprotein binding motifs of Nudel suggest that this modular SP might interact with the extracellular matrix and function in establishing the ventral prepattern (LeMosy et al., 1998). However, some studies have demonstrated that Nudel is not required in the ventral follicle cells that express *pipe*, indicating that Nudel is not a target of Pipe (Stein et al., 2008). It is still unclear whether Nudel is directly involved in the cleavage of GD. The processed form of GD is not found in eggs where protease activity of Nudel is compromised (LeMosy et al., 2001), suggesting its involvement in GD cleavage. However, GD does not require processing for activation of its proteolytic function. This has been observed in *gd* mutants with normal processing of Snake, but no activated Easter. These findings imply that there is another catalytic activity of GD. Indeed, GD functions in a complex with Snake and Easter, facilitating Easter processing by Snake (Cho et al., 2012). Furthermore, the necessary ventrally localized Easter processing is achieved by interaction between GD and a sulfated cue on the ventral side of the embryo, provided by *pipe* (Figure 7).

Activation of the Toll pathway during immune response

Activation of the Toll pathway during the immune response in *Drosophila* larvae and adults also requires proteolytically processed Spz. However, the SPs involved in the cascade are different from those involved during embryogenesis; loss-of-function mutants for these SPs show wild type immune response upon infection (Lemaitre et al., 1996). Different studies have demonstrated that two parallel cascades achieve proteolytic processing of Spz. These cascades consist of SPs that converge on SPE, leading to the formation of the ligand form. The cascades are triggered by either recognition of microbial cell wall components or danger signals produced by microorganisms, and are referred to as Recognition Cascade or Danger Signal Cascade, respectively.

Activation by recognition

Differential induction of genes encoding AMPs and their selective activation depend on the immune challenge. Gram-positive bacteria and fungi activate the Toll pathway, whereas Gram-negative bacteria activate the IMD pathway. This implies that microbial recognition mechanisms can distinguish between different classes of microorganisms. Genetic studies have shown that the recognition is achieved by two families of PRRs present in the hemolymph or at the cell membrane, PeptidoGlycan Recognition Proteins (PGRPs) and Gram-Negative Binding Proteins (GNBPs). These two families of receptors were initially identified in larger insects form their ability to

bind microbial ligands and activate the melanization cascade (Kang et al., 1998; Lee et al., 1996; Yoshida et al., 1996).

The family of PGRPs is highly conserved from invertebrates to vertebrates. The PGRP domain, which is common to these receptors, is implicated in the immune response and may function by interacting with or degrading microbial peptidoglycan (PGN). In the *Drosophila* genome, there are 13 genes encoding PGRPs. These receptors are divided into two groups based on their peptide length. Short PGRPs (PGRP-S) have signal peptides and are extracellular proteins. Long PGRPs (PGRP-L) are transmembrane proteins and can be either intracellular or extracellular. PGRPs are expressed in various immune-responsive tissues - the fat body, hemocytes, gut, and epidermis (Werner et al., 2000).

Another classification, based on the sequence of the PGRP domain, distinguishes two groups. Catalytic PGRPs have amidase enzymatic activity capable of hydrolyzing PGN. These include PGRP-LB, SC1, and SC2 (Bischoff et al., 2006; Zaidman-Rémy et al., 2006). Recognition PGRPs lack amino acid residues that are essential for the catalysis, but they can bind PGN. This group includes PGRP-SA, SD (Bischoff et al., 2004; Michel et al., 2001) PGRP-LC, and PGRP-LE (Choe et al., 2002; Gottar et al., 2002; Kaneko et al., 2006; Rämet et al., 2002; Takehana et al., 2002).

The family of GNBPs is related to the group of β-glucanases found in bacteria, fungi, plants, and animals. The insects GNBPs have two conserved domains – an N-terminal domain that binds β(1,3)-glucans and a C-terminal domain, which is homologous to the catalytic domain of β-glucanases. The absence of conserved amino acid residues in this C-terminal domain implies that it is not catalytically

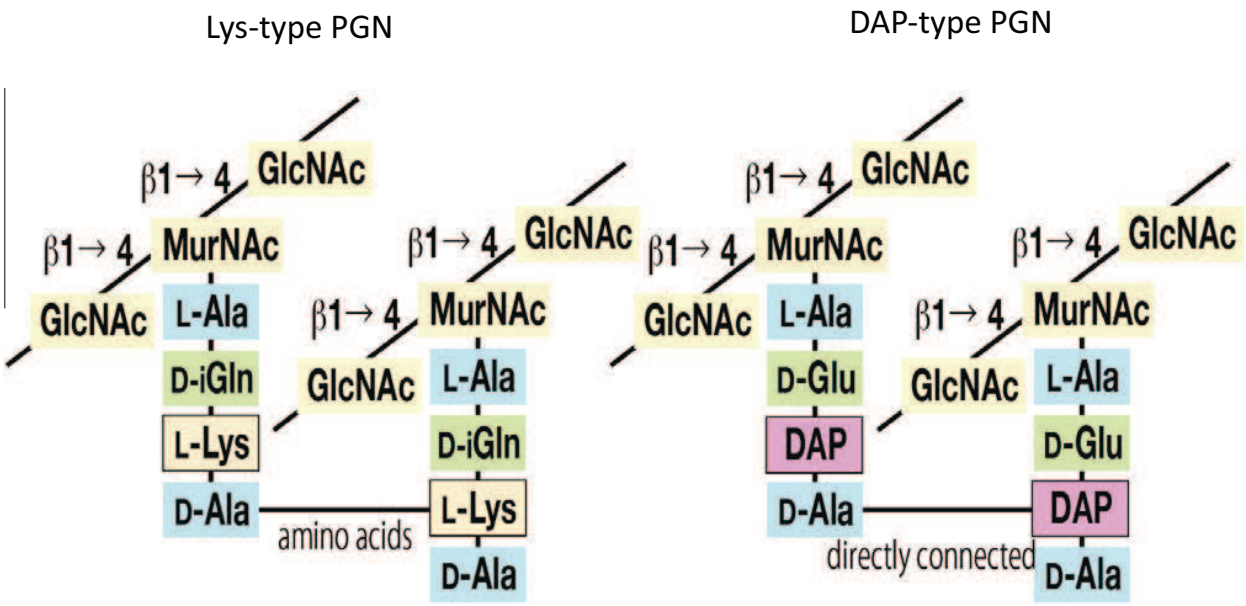


Figure 8 Peptidoglycan structure

Peptidoglycans (PGN) are polymers of β-1,4-linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) cross-linked by short stem peptides, and are categorized into two major types: Lys-type and diaminopimelic acid (DAP)-type, based on the amino acid composition of the stem peptides and the linkage between the stem peptides (Kurata, 2014).

active. There are three genes encoding GNBPs in *Drosophila*, GNBP1-3, and three GNBP-like members of the family that have only an N-terminal domain (Ferrandon et al., 2004). GNBP proteins have a role in the detection of both fungal (Bangham et al., 2006) and bacterial cell wall components (Gobert et al., 2003; Wang et al., 2006).

Bacterial recognition is achieved through the sensing of different forms of PGN, an essential component of the bacterial cell wall. It is found in both Gram-positive bacteria (where it is multilayered and exposed at the cell surface), and Gram-negative bacteria (with single layered PGN, located in the periplasmic space underneath the outer membrane and lipopolysaccharide layers). PGN is a glucopeptidic polymer consisting of long glycan chains of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic (MurNAc) acid cross linked to each other by short peptide bridges (Kang et al., 1998). Besides its localization in the cell wall, the PGN from Gram-negative bacteria differs from the PGN found in most Gram-positive bacteria (Lys-type) by the replacement of lysine with *meso*-diaminopimelic acid (DAP) at the third position in the peptide chain (DAP-type) (Figure 8). Some subclasses of Gram-positive bacteria such as *Bacillus* species also produce DAP-type PGN.

Using highly purified bacterial compounds, Leulier and colleagues have demonstrated that the IMD pathway is activated by DAP-type PGN, whereas the Toll pathway is activated by Lys-type PGN (Leulier et al., 2003). Further studies have shown that a minimal PGN motif capable of IMD induction is a specific monomer known as tracheal cytotoxin (TCT) (Kaneko et al., 2004; Stenbak et al., 2004). The

minimal PGN motif effective in Toll induction is a muropeptide dimer of Lys-type PGN (Filipe et al., 2005).

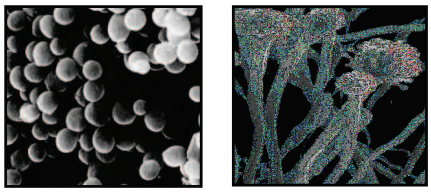
The detection of Gram-negative bacteria and the consequent activation of the IMD pathway are mediated by PGRP-LC and PGRP-LE. PGRP-LC mutant flies show a high susceptibility to some Gram-negative bacteria, such as *Enterobacter cloacae* and *Erwinia carotovora*, which correlates with a decreased AMP expression (Choe et al., 2002; Gottar et al., 2002). Since IMD pathway activation is not completely blocked in PGRP-LC mutants, it has been suggested that there is another co-receptor required for the activation. PGRP-LE binds DAP-type PGN but flies mutant for PGRP-LE do not show any defects in IMD pathway activation (Takehana et al., 2002). A cleaved form of PGRP-LE can be found in the hemolymph, whereas uncleaved full-length PGRP-LE is found in Malpighian tubules and hemocytes. It has been demonstrated that PGRP-LE can activate the IMD pathway in two distinct ways, functioning either as an extracellular or intracellular receptor. Genetic studies analyzing double mutants of PGRP-LC and PGRP-LE have demonstrated that these molecules act in synergy to activate the pathway upon immune challenge with Gram-negative bacteria (Takehana et al., 2004). The extracellular form of PGRP-LE activates the pathway by forming a complex with PGRP-LC. The intracellular PGRP-LE interacts with Imd via the N-terminal region. It is required to sense intracellular bacteria such as *Listeria monocytogenes*.

Two members of the PGRP family (PGRP-SA and SD) and a member of the GNBP family (GNBP1) are involved in the recognition of Gram-positive bacteria and

the subsequent activation of the Toll pathway. They have been predicted to be circulating in the hemolymph (Michel et al., 2001; Wang et al., 2008).

A mutation in the gene encoding PGRP-SA (*PGRP-SA*^{semmelweis}) causes the flies to be susceptible to infection with Gram-positive bacteria; the Toll pathway in these flies is not activated (Michel et al., 2001). A mutation of GNBP1 (*GNBP1*^{osiris}) causes sensitivity to infection with *Enterococcus faecalis*; the mutant flies do not activate the Toll pathway upon challenge with either *E. faecalis* or *Micrococcus luteus* (Gobert et al., 2003). The interaction between PGRP-SA and GNBP1 has been shown on native protein gels, using fly extracts (Gobert et al., 2003). However, some bacterial strains induce the activation of the Toll pathway even in the absence of functional PGRP-SA/GNBP1 complex. This suggested the existence of another alternative receptor, namely PGRP-SD, implicated in the recognition of some Gram-positive bacteria (Bischoff et al., 2004). After infection with either *E. faecalis* or *Staphylococcus aureus*, PGRP-SA mutant flies show reduced immune response, even more pronounced in double PGRP-SA/PGRP-SD mutants. Conversely, infection with *S. saprophyticus* is independent of PGRP-SA. This implies that recognition of different bacteria relies on the formation of different receptor complexes. Indeed, it has been demonstrated that in *S. aureus*, PGRP-SD enhances the binding of PGRP-SA/GNBP1 complex, whereas in *M. luteus*, only PGRP-SA is required (Wang et al., 2008).

The only fungal receptor identified so far is GNBP3. Partial deletion of the gene encoding GNBP3 (*GNBP3*^{hades}) results in the inability to activate the Toll pathway with fungal cell wall components and susceptibility to fungal infections (Gottar et al., 2006).



Gram Positive Bacteria Fungi

PAMPs

Circulating receptors

ModSP

↓

Grass

↓

SPE

↓

Pro-Spätzle

↓

Spätzle

⇣

Toll pathway activation

Figure 9 Recognition cascade

Activation by recognition; proteolytic cascade comprising of serine proteases ModSP, Grass and SPE

Recognition cascade

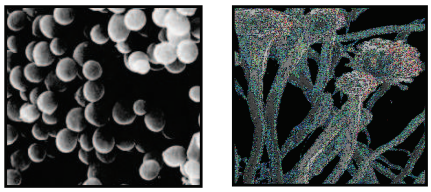
Upon immune challenge, PAMPs expressed by Gram-positive bacteria or fungi, are recognized by PRRs circulating in the hemolymph. This recognition triggers the proteolytic cascade using an unknown mechanism. The recognition cascade consists of 3 SPs: Modular Serine Protease (ModSP), Gram-Positive Specific Serine Protease (Grass), and SPE. Genetic studies have shown that ModSP functions upstream of Grass, which cleaves SPE, which, in turn, processes Spz into its active ligand form (Figure 9).

The complete cascade has been first purified and reconstructed *in vitro* from the hemolymph of infected beetle *Tenebrio molitor* (Kim et al., 2008). In this species, the binding of bacterial PGN to PGRP-SA/GNBP1 complex induces the activation of a modular SP (Tm-MSP). *Drosophila* has a Tm-MSP homolog, ModSP. ModSP is not a CLIP domain SP; it has a large N-terminal domain containing four Low Density Lipoprotein-receptor class A (LDLa) domains and one Complement Control Protein (CCP) module (Sushi domain). The existence of LDL domains suggests that ModSP could associate with extracellular vesicles. Flies that are null mutants for *ModSP* are susceptible to immune challenge with Gram-positive bacteria and show reduced levels of *drosomycin* in comparison with wild type flies (Buchon et al., 2009). ModSP is the initial protease in the recognition cascade, but its interaction with PRRs has not been observed.

Grass was first identified in an *in vivo* RNAi screen as an SP required for Toll activation. RNAi-mediated knockdown gives a hypomorphic phenotype in adult flies

and shows that the SP encoded by this gene is required for resistance to infections only with Gram-positive bacteria (Kambris et al., 2006). Studies performed by our group have used a null allele, *grass*^{Herrade}, and confirmed that Grass is necessary for the response to Gram-positive bacteria. However, we have also shown that Grass is involved in the response to the fungus *Beauveria bassiana* and the yeast *Candida albicans* (El Chamy et al., 2008). Grass functions downstream of ModSP, but it is not cleaved by an activated ModSP *in vitro* (Buchon et al., 2009). It has been predicted that the cleavage requires a trypsin-like SP, and ModSP is a chymotrypsin-like SP. Based on this and other criteria, it has been suggested that several other SPs might function between ModSP and Grass; this proposal will be discussed in the first part of the Results chapter.

SPE has been first identified by homology to SP from another insect, the silkworm *Bombyx mori*. In its hemolymph, a CLIP domain-containing SP called BAEEase (it hydrolyzes the synthetic substrate Nα-benzoyl-L-arginine ethyl ester, BAEE) is activated by PAMPs expressed by both Gram-positive bacteria and fungi. It has been reasoned that the proteolytic cascades downstream of this recognition point in the silkworm and *Drosophila* should also share the downstream protease that cleaves pro-Spaetzle. Indeed, among 24 CLIP domain-containing SPs in *Drosophila*, only one has the identical cleavage site for its activation. This SP, SPE, directly cleaves pro-Spaetzle into its active form *in vitro*. RNAi-mediated knockdown in adult flies has shown that this SP is required for resistance to infections with both Gram-positive bacteria and fungi, as well as for Toll-dependent AMP expression (Jang et al., 2006).



Gram Positive Bacteria Fungi

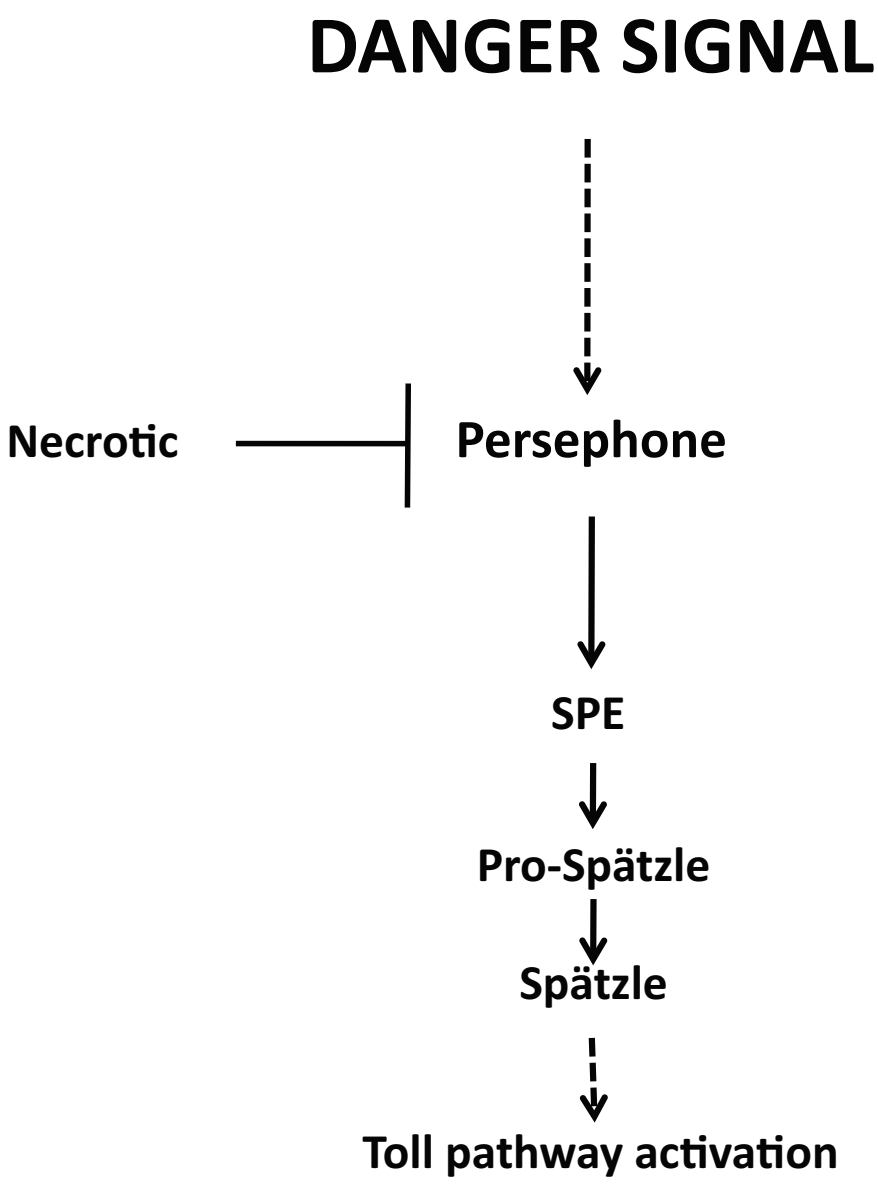


Figure 10 Danger signal cascade

Activation by danger signals; proteolytic cascade comprising of serine proteases Psh.

Danger signal cascade – components and activation

Besides expressing PAMPs, recognized by PRRs in the hemolymph, microorganisms also produce different virulence factors. Both Gram-positive bacteria and fungi produce different proteases as a way to attack the host. These proteases are sensed in the hemolymph and trigger proteolytic cascades leading to the activation of the Toll pathway.

The first evidence that a proteolytic activity can activate the Toll pathway came from the studies of a serpin Nec. Flies that are mutant for the gene encoding Nec show a pleiotropic phenotype that includes constitutive expression of *drosomycin*, spontaneous melanization in the absence of infection, and death in the early adulthood. Furthermore, in these flies, Spz is proteolytically cleaved, which leads to constitutive activation of the Toll pathway (Levashina et al., 1999). The nature of the protein, as well as the observed phenotype, suggests the existence of an SP (or SPs) that would activate the Toll pathway in the absence of Nec. A forward genetic screen using chemical mutagenesis identified a suppressor of Nec phenotype. Flies with this mutation are viable but show susceptibility to fungal infections. The mutation responsible is in a gene encoding the SP Persephone (Psh) (Ligoxygakis et al., 2002a, 2002b). So far, only genetic interaction has been demonstrated between Nec and Psh. Nec has a broad specificity for proteases, and the question remains whether Psh is its direct target or if there are other components involved in the cascade (Figure 10).

As previously mentioned, GNBP3 is the only fungal receptor identified so far. Genetic analyses have shown that it does not function upstream of Psh; however, when *psh* and *GNBP3* mutations are combined, flies are even more susceptible to fungal infections in comparison with single mutants, implying the existence of two parallel pathways. *B. bassiana* and *Metharizium anisopliae* are natural entomopathogens of *Drosophila*; the germinating fungal spores secrete proteases and chitinases to digest the insect cuticle and penetrate the body cavity. Pr1 subtilisins expressed by the spores are major virulence factors of *M. anisopliae*. When Pr1 is ubiquitously overexpressed in flies, it induces constitutive activation of the Toll pathway, dependent on Psh. The pathway mediated by GNBP3 is activated by fungal cell wall components, whereas the pathway mediated by Psh is activated by fungal proteases (Gottar et al., 2006).

Further studies performed by our group have demonstrated that Psh also mediates the Toll pathway activation by bacterial proteases, functioning upstream of SPE and therefore defining a parallel pathway leading to Spz processing (El Chamy et al., 2008).

The studies have demonstrated that Psh acts as a sensor of the danger signals elicited by both Gram-positive bacteria and fungi. When flies mutant for *psh* are injected with diluted commercial proteases (form either *Bacillus sp*., or *Aspergillus oryzae*), a decrease in the Toll pathway activation is observed in comparison with wild type flies. This is not observed in $grass^{Herrade}$ mutants, indicating the existence of two cascades leading to Spz processing (El Chamy et al., 2008). Furthermore, in flies that are double mutants for *psh* and $grass^{Herrade}$, the Toll

pathway activation is completely blocked, whereas in flies that are double mutants for *GNBP3*^{Hades} and *grass*^{Herrade}, the Toll pathway is activated to the same level as in single mutant *grass*^{Herrade} flies. This result has shown that Psh and Grass belong to two different parallel cascades.

Another observation of two parallel pathways has been made during infection with the fungus *B. bassiana*. At the beginning of infection, the Toll pathway activation is blocked in *psh* mutant flies, indicating that the fungal proteases are expressed early in infection. Later, during the course of infection, the Toll pathway is activated as in wild type flies, indicating the recognition of fungal cell wall components (personal observation). This shows the significance of sensing danger signals in the early response of *Drosophila*. Indeed, we can observe that this earlier sensing is vital since *psh* mutant flies that mount a delayed but nonetheless normal immune response die faster than wild type flies.

As mentioned above, the epithelial tissues are in constant contact with the environment and present the first line of defense. One of the epithelial tissues is the trachea. In the presence of microorganisms, the process of melanization in this tissue is activated and regulated by Spn77Ba. It has been reported that tracheal melanization in the absence of Spn77Ba induces a systemic response via Psh-dependent Toll pathway activation in the fat body (Tang et al., 2008). This can be a mechanism that alerts and prepares the host for possible infections. The nature of the product signaling between trachea and the fat body is still unknown. It is probably a diffusible molecule, which can pass the basement membrane to the hemolymph and acts upstream of Psh, leading to the activation of the Toll pathway.

Immune response can be activated by endogenous factors produced by host cells. These signals are termed Damage (or Danger) Associated Molecular Patterns (DAMPs). The release of DAMPs is associated with cellular damage, in particular the loss of membrane integrity or rupture of the plasma membrane, a hallmark of the necrotic forms of cell death. In eukaryotic cells, a family of intracellular proteases, caspases, functions in the apoptotic signaling leading to the removal of unwanted cells without the stimulation of the immune system. In cases of caspase deficiency or suppression, another form of cell death, necrosis, is activated, involving the release of DAMPs. The key initiator of apoptotic caspase cascade is Dronc. Recent studies of *Drosophila* larvae deficient for Dronc have demonstrated that, due to necrosis, endogenous DAMPs are released in their hemolymph. This abnormal proteolytic activity leads to proteolytic processing of Spz, mediated by Psh (Ming et al., 2014).

The mechanisms underlying this activation through Psh by various proteases produced by pathogens, or endogenous signals produced by host's damaged cells or tissues, as well as the existence of other components in this signaling, are not clear. Psh is a CLIP domain-containing SP. The cleavage site associated with zymogen activation is next to a histidine residue, which is very unusual (Rawlings and Barrett, 1999; Ross et al., 2003). Therefore, it is unlikely that many microorganisms producing different proteases with various specificities could cleave and activate Psh at this specific site. It raises the question of other SPs (or SPHs) being involved in this process. This is one of the questions being investigated in our group.

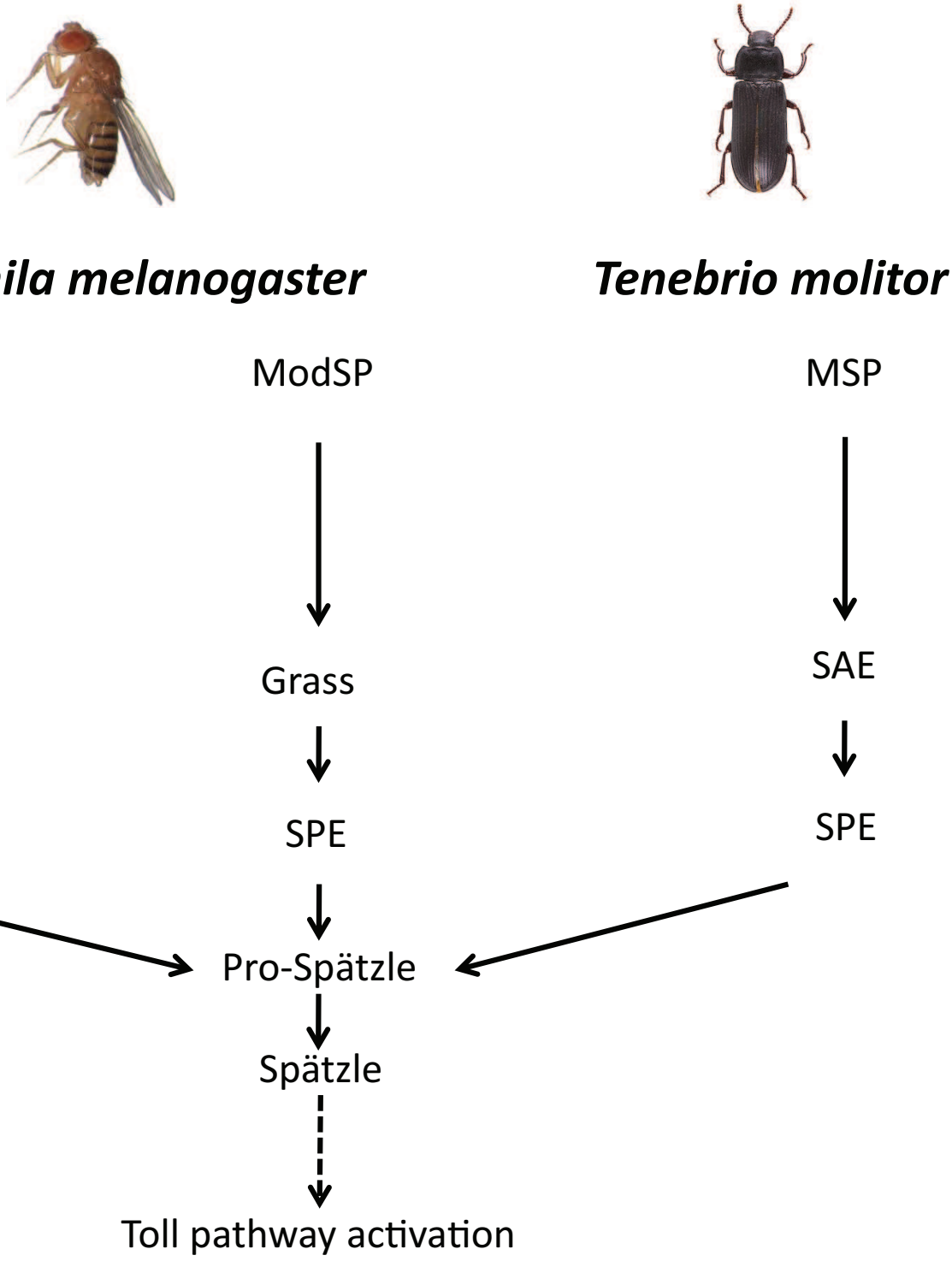


Figure 11 Comparison of proteolytic cascades containing CLIP-SPs

On the left side are the proteolytic cascades involved in Toll pathway activation in *Drosophila melanogaster* (immune response and development); on the right side is the proteolytic cascade involved in the immune response of *Tenebrio molitor*.

4. Thesis project

The aim of my PhD research was to characterize new components involved in the cascades that lead to proteolytical cleavage of Spz and activation of the Toll pathway.

The first part of this study focuses on the components of the recognition cascade and the second part, on the components of the danger signal cascade.

The current model of the recognition cascade consists of ModSP, Grass, and SPE. However, this model is clearly incomplete. Comparing proteolytic cascades that lead to Toll activation during development and immune response in *Drosophila* and immune response in *T. molitor*, we can see structural and functional similarities between ModSP and Tm-MSP, Easter, SPE and Tm-SPE, as well as Tm-SAE and Snake (Figure 11). However, several lines of evidence indicate that Grass is not the functional equivalent of Tm-SAE and Snake:

- The specificity pocket of ModSP is made of hydrophobic residues (Leu, Ala, Thr) and is chymotrypsin-like; therefore, it cannot accommodate Arg109 of the activation site of Grass (R▼VSNG). It has been confirmed that Grass is not cleaved by an activated ModSP *in vitro* (Buchon et al., 2009).
- Activated Grass does not cleave SPE at its normal activation site; rather Grass cleaves SPE several aa upstream of the activation site (Christine Kellenberger, personal communication).
- The structure of Grass is different from Snake or Tm-SAE, which are penultimate proteases in the cascades. Sequence alignment of catalytic

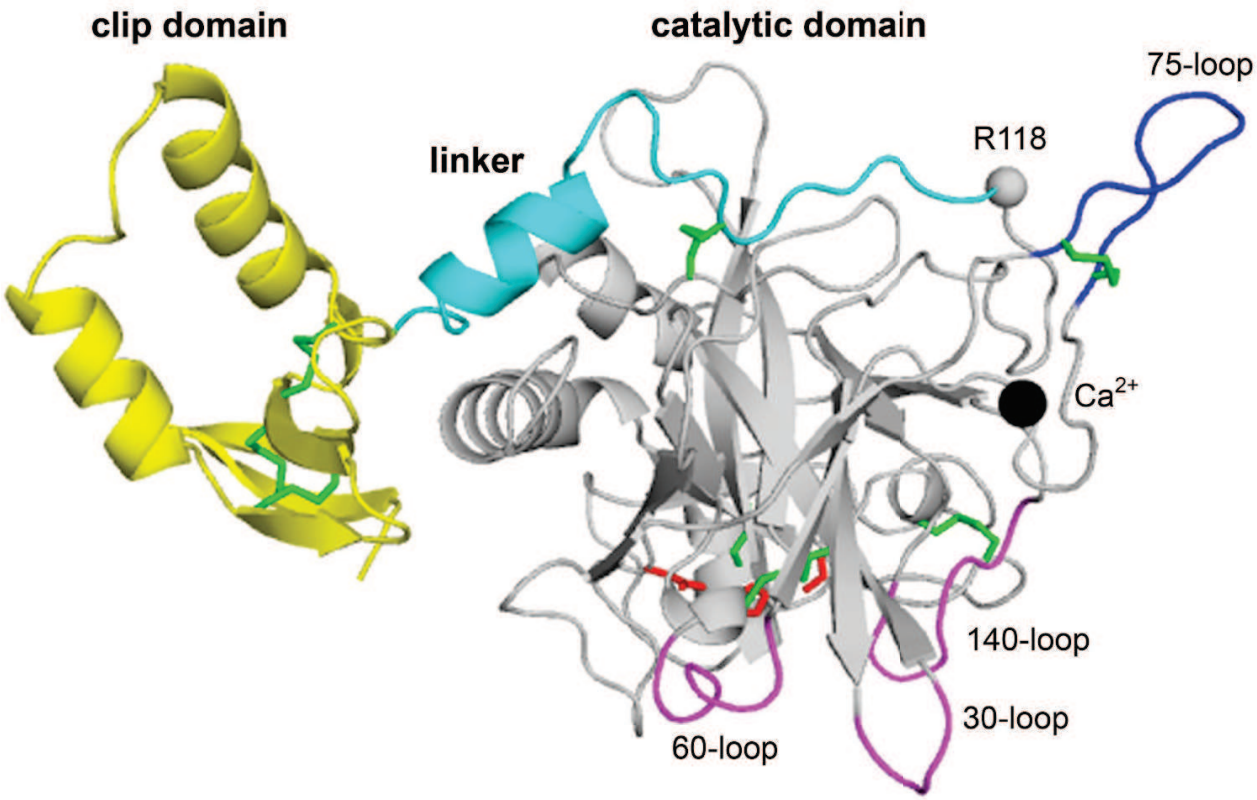


Figure 12 Structure of clip-SP Grass

Overall structure in ribbon representation, with the three domains, clip, linker and catalytic, colored in yellow, cyan and grey, respectively. The disulfide bridges and the catalytic triad are depicted as sticks, and colored in green and red, respectively. The calcium ion is represented as a dark sphere and Arg118 (P1 of the activation site) as a grey sphere. The three loops shaping the active site cleft (30, 60 and 140) and the 75-loop are colored in magenta and blue, respectively (Kellenberger et al., 2011).

domains of CLIP-SPs shows that the major difference is the existence of the 75-loop. This has prompted a new classification, based on the existence of the loop and conformation of the clip domain, into two groups, the penultimate proteases without loop (Snake, Ms-HSP6, and Tm-SAE) and terminal proteases with a loop (Easter, SPE, HP8, and Tm-SPE). Analysis of the crystal structure of Grass identifies an additional 75-loop near the activation site, which prevents spontaneous activation (Kellenberger et al., 2011) (Figure 12). According to these findings, Grass should be a terminal protease functioning at the level of SPE. However, the results of epistatic analysis, in which over expression of Grass in SPE-mutant background did not activate the Toll pathway, are consistent with the hypothesis that Grass was acting upstream of SPE. However, over expression of SPE in Grass-mutant background does not activate the pathway either. That can only position Grass at the same functional level as SPE (Vincent Leclerc, personal communication).

These prompted a search for a potential SP that could function between ModSP and Grass (or SPE), which will be described in the first part of this study.

The second part of the Results describes the characterization of the serine protease homolog Spheroide. The results show that Spheroide is specifically involved in sensing proteolytic activities produced by pathogenic bacteria. Spheroide is involved in the danger signal cascade at the level of or downstream from Persephone.

Results

1. The serine protease between ModSP and Grass

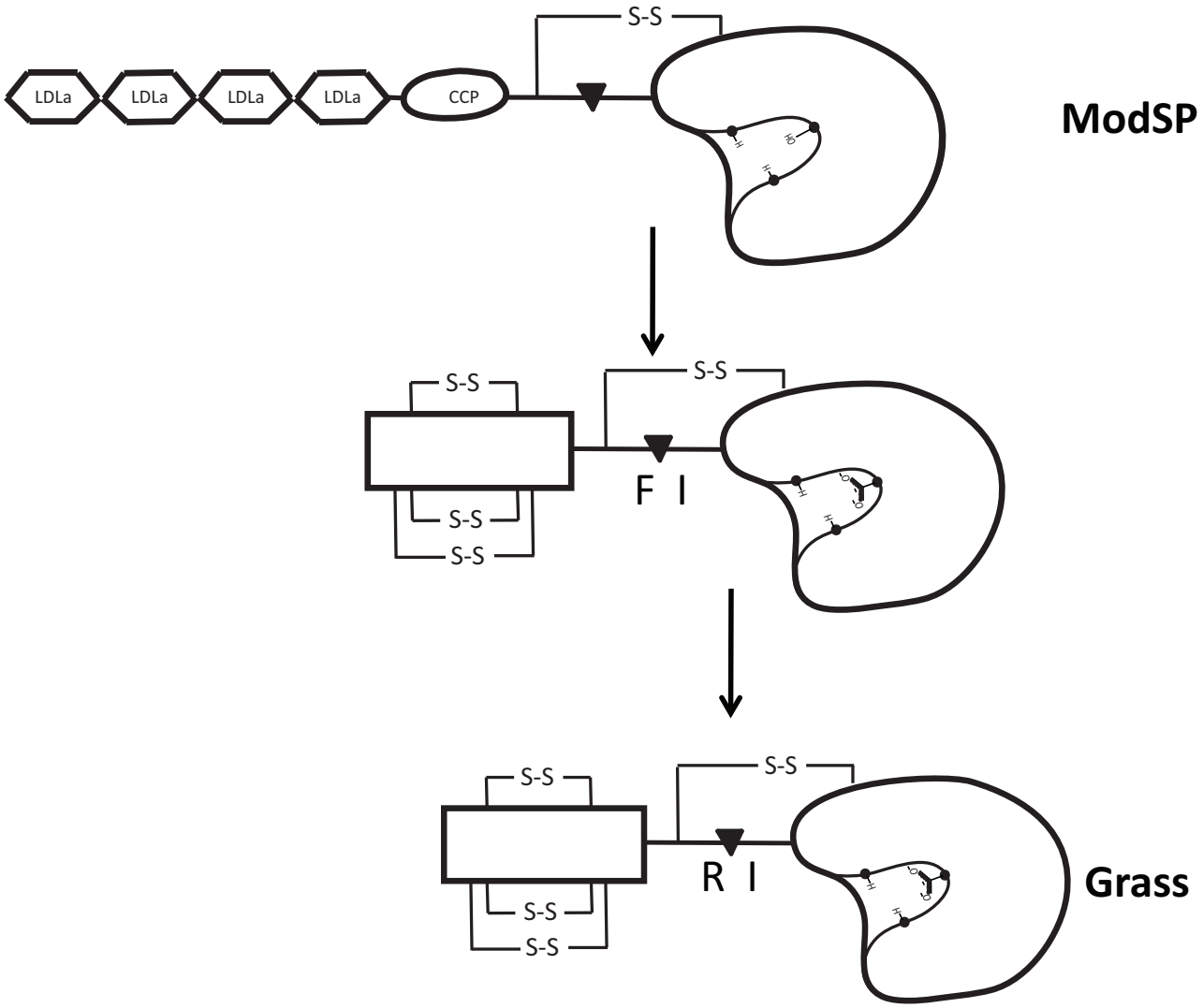


Figure 13 Schematic representation of ModSP, Grass and a putative protease with CLIP-domain, specific activation site and trypsin-like catalytic domain

A search for a potential SP that could function between ModSP and Grass was based on bioinformatic analysis, the chemical knowledge and the specificity of the SPs involved. Different criteria for the potential SP that were taken into consideration were:

- It should have a trypsin-like catalytic domain in order to activate Grass that has an arginine in the activation site (R▼VSNG).
- The activation site of the protease should fit the specificity pocket of the chymotrypsin-like catalytic domain of ModSP and therefore be composed of phenylalanine, tryptophan or tyrosine.
- It should have a CLIP domain in the N-terminal region (Figure 13).

Two independent BLAST searches were performed using the aa sequences of the catalytic domain of either Snake or Tm-SAE and three SPs were identified (in collaboration with Christine Kellenberger, Marseille). The genes encoding these SPs are *CG3700*, *CG4927* and *CG2056* (*spirit*). *CG2056* (*spirit*) was also suggested by an RNAi screen (Kambris et al., 2006).

I will describe different genetic approaches and strategies that were used to manipulate the genes in question and generate mutants and I will show the results that were obtained for *CG3700* and *CG2056* (*spirit*).

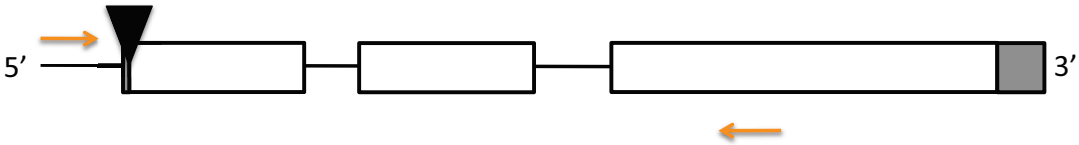
Candidate gene *CG3700*

RNA interference

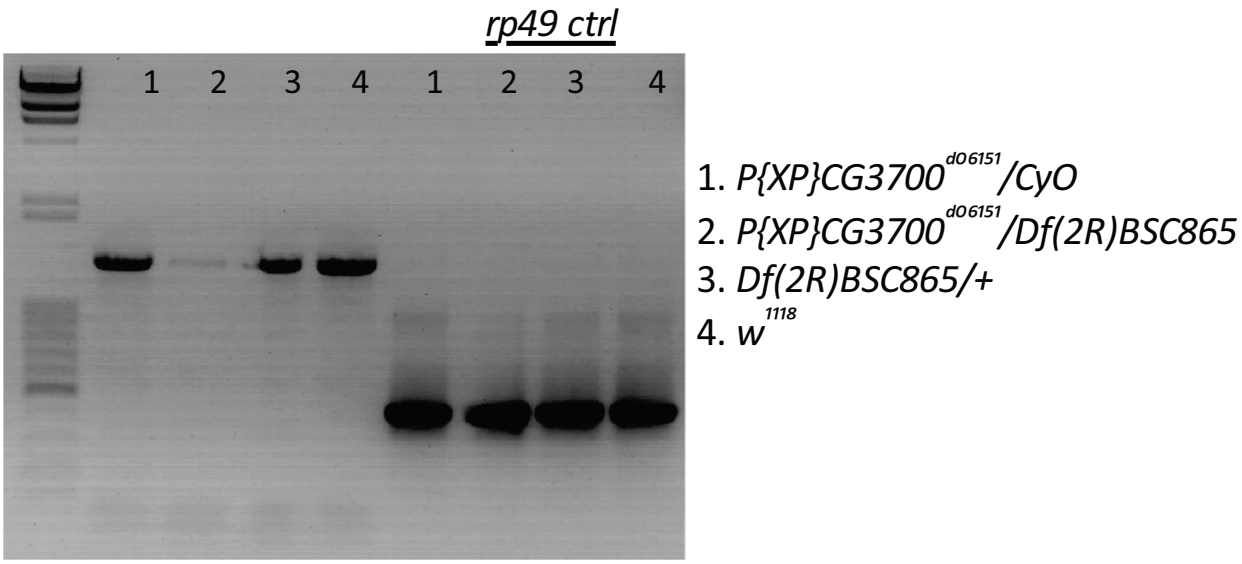
For RNAi mediated knockdown we used the UAS/Gal4 system (Brand and Perrimon, 1993). We obtained VDRC lines with inverted repeats containing inducible *UAS-RNAi* against our genes of interest. For expression of *UAS-RNAi* constructs we used different Gal4 drivers (*Act5C*, *hsp70*, *yolk*, *da*). Gal4 driven RNAi knockdown using *Act5C*, *yolk* and *da* was enhanced by incubating three days-old flies for four days at 29 °C. Gal4 driven RNAi knockdown using *hsp70* was performed on three days-old flies that were kept for three days at 29 °C, with a heat shock performed each day. The heat shock consisted of 20 minutes at 37 °C, 20 minutes at 18 °C and 20 minutes at 37 °C. To verify that expression of the gene was reduced, we performed the quantitative PCR to check the level of the gene expression.

We obtained VDRC lines with inverted repeats containing inducible *UAS-RNAi* against *CG3700* (*P{KK106926}VIE-260B*) and used it in our experiments. RNAi mediated knockdown was not efficient. Levels of gene expression were the same as in wild type flies. Before completely excluding RNAi as a strategy, Gal4 driven RNAi knockdown was enhanced during a course of time, keeping the flies for longer than three days at 29 °C, as well as performing more than three heat shocks. These experiments were repeated few times, showing no efficient reduction in gene expression of targeted genes.

A



B



C Immune challenge with *M. luteus*

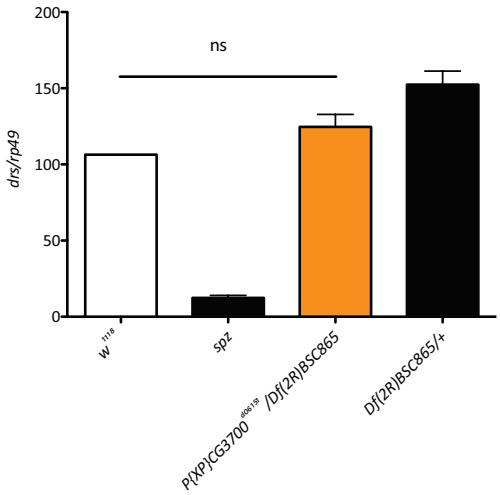


Figure 14 (to be continued on the next page)

P-element insertion

To test whether the serine protease encoded by the gene *CG3700* is involved in the activation of the Toll pathway we used flies that were heterozygous carrying the (*P{XP}CG3700*^{d06151}) P-element insertion in the 5'UTR region and the *Df(2R)BSC865* deficiency spanning this region (Figure 14A). Insertion of the P-element in *CG3700* was verified by PCR (Figure 14B).

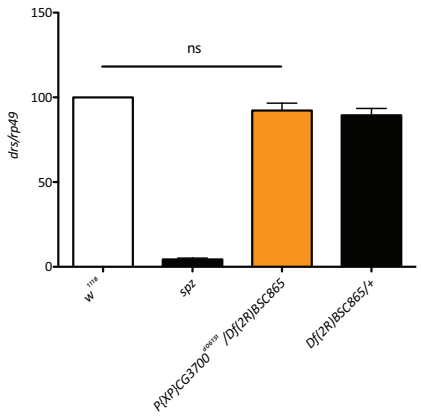
When flies were immune challenged with the non-pathogenic Gram-positive bacterium *Micrococcus luteus*, we observed *drosomycin* expression comparable to that in wild type flies (Figure 14C). This result indicated that the serine protease encoded by *CG3700* is not involved in the immune response against *Micrococcus luteus*.

When flies were immune challenge with the pathogenic Gram-positive bacterium *Enterococcus faecalis*, they showed the same susceptibility to this infection as wild type flies (Figure 14E). *drosomycin* expression was comparable to that in wild type flies (Figure 14D). This result indicated that CG3700 serine protease is not involved in the immune response against *Enterococcus faecalis*.

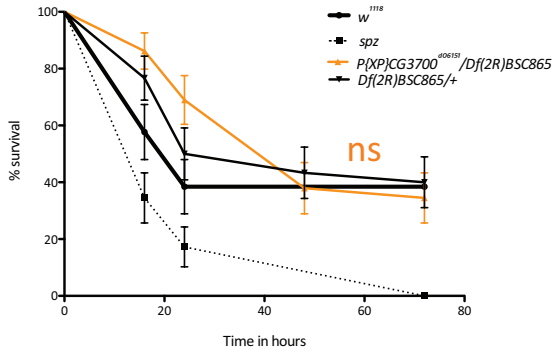
When flies were infected with the fungus *Beauveria bassiana*, they showed the same susceptibility to this infection as wild type flies (Figure 14G). *drosomycin* expression was comparable to that in wild type flies (Figure 14F). This result indicated that CG3700 serine protease is not involved in the immune response against *Beauveria bassiana*.

Immune challenge with *E. faecalis*

D

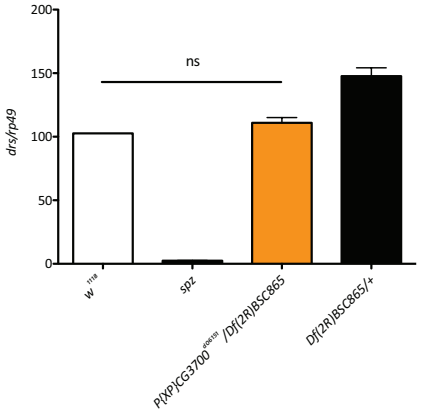


E



Natural infection with *B. bassiana*

F



G

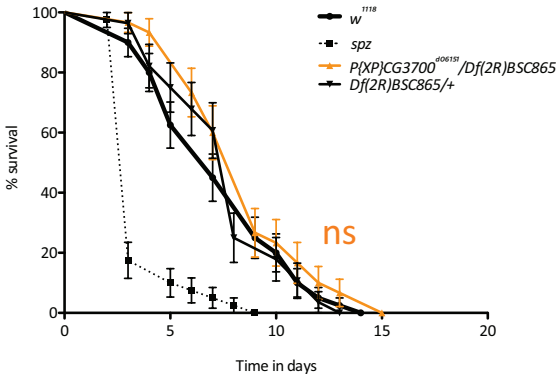


Figure 14 *CG3700* is not involved in the Toll pathway activation

Schematic representation of genomic region and P-element insertion (▼) in *CG3700* (P{XP}CG3700^{d06151}), primers used for PCR indicated with orange arrows (**A**); Insertion of the P-element in *CG3700* in flies heterozygous for the P-element insertion deficiency spanning this region (P{XP}CG3700^{d06151}/Df(2R)BSC865) shown by PCR result using the indicated primers (**B**); *drosomycin* expression 24 hours PI, normalized to *rp49* and in percentage to its expression in *w^{1118}* wild type flies after immune challenge with *M. luteus* (**C**) or *E. faecalis* (**D**) or after natural infection with *B. bassiana,* 48 hours PI (**F**). Survival rate after immune challenge with *E. faecalis* (**E**) or natural infection with *B. bassiana* (**G**). There is no statistical significance between survival of *w^{1118}* and *P{XP} CG3700^{d06151}/Df(2R)BSC865* flies (3 independent experiments)

Taken together these results demonstrate that serine protease encoded by *CG3700* is not involved in the activation of the Toll pathway after immune challenge with these specific pathogens.

My main objective was to look for the serine protease acting between ModSP and Grass and further investigation of this gene was not pursued. However, we cannot exclude involvement of this serine protease in the immune response against other pathogens. Furthermore, there are other proteolytic cascades in which this serine protease can be involved such as melanization. Further investigation and characterization of this serine protease is necessary.

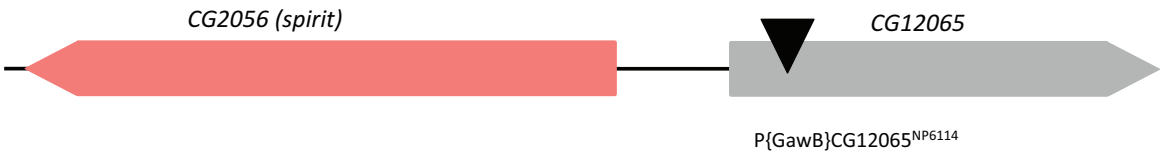


Figure 15 Schematic representation of P-element excision strategy for *CG2056 (spirit)*

Candidate gene *CG2056 (spirit)*

RNA interference

RNAi-mediated knockdown was performed as previously described. We obtained VDRC lines with inverted repeats containing inducible *UAS-RNAi* against *CG2056* (*P{GD3285}v5497*, *P{KK112114}VIE-260B*). RNAi mediated knockdown was not efficient. Levels of gene expression were the same as in wild type flies. Before completely excluding RNAi as a strategy, Gal4 driven RNAi knockdown was enhanced during a course of time, keeping the flies for longer than three days at 29 °C, as well as performing more than three heat shocks. These experiments were repeated few times, showing no efficient reduction in gene expression of targeted genes.

P-element excision

There are no transposable elements inserted in the sequence coding for *CG2056 (spirit)*. Therefore we used the availability of the P-element insertion (*P{GawB}CG12065*^{NP6114}) in the 5'UTR region of the flanking gene *CG10265* (Figure 15) to generate an imprecise P-element excision. For transposase, we used a fly line carrying *P{Δ2-3}*. Few days after performing the cross, the parents were transferred into new tubes and the tubes with the progeny were raised at 29 °C to increase the efficiency of the transposase (personal observation). After several hundreds of lines tested, only two lines with a clean excision were obtained. The gene encoding

CG12056 is suggested to be involved in lateral inhibition and we can only speculate that deletions due to P-element excision were lethal due to inactivation of *CG12056*.

Transcription Activator-Like Effector Nucleases (TALENs)

TALENs are a newly developed genomic editing tool that can be designed to create double strand (DS) breaks at specific locations within the genome. They represent an artificial fusion protein containing the high-specificity TAL DNA-binding fused to the Fok1 restriction enzyme endonuclease domain (Li et al., 2011). Each TAL domain consists of approximately 18 repeats. Each repeat is nearly identical, except for the hypervariable region consisting of two amino acids that are responsible for binding specifically to one nucleotide base. This region protrudes from the alpha helix "fingers" that are formed by the remainder of the TAL repeat being exposed to interact with nucleotide bases (Deng et al., 2012; Mak et al., 2012). It was recently discovered which hypervariable regions bind to which nucleotide bases (Boch et al., 2009). We can take advantage of this property to modularly assemble TAL repeats to bind to any nucleotide sequence (Cermak et al., 2011; Geißler et al., 2011; Morbitzer et al., 2011).

In order to cleave a specific genomic sequence, TALENs are designed to function in pairs because the Fok1 domain must dimerize in order to create a DS break. The distance between pairs is 14-16 bp. TALEN-created DS breaks can be repaired either by homologous recombination (HR) or by non-homologous end joining (NHEJ). In the first case, repair by HR will not yield any mutation. In the

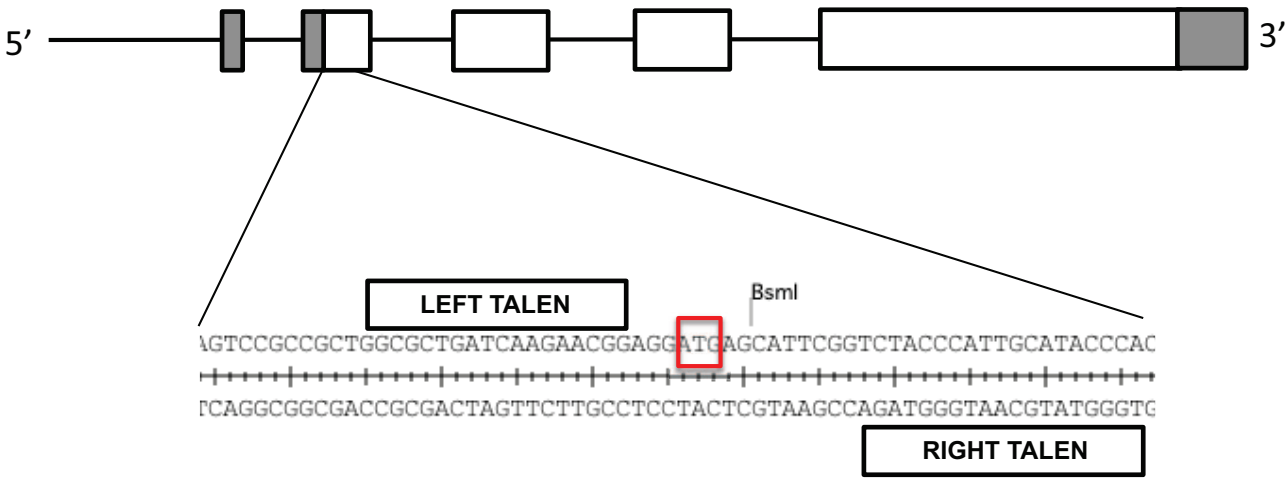


Figure 16 Schematic representation of TALEN strategy

Start codon with BsmI recognition site for *CG2056* (*spirit*)

gene	method	lines tested	mutants
spirit	transgenic lines	700	0
	TALEN-coding mRNA	100	0

Table 1

second case, repair by NHEJ should allow for small insertions and deletions, making it possible to generate targeted gene knockouts.

Two approaches were used: generation of transgenic lines (in collaboration with Eric Marois) and injection of TALEN-coding mRNA (Liu et al., 2012).

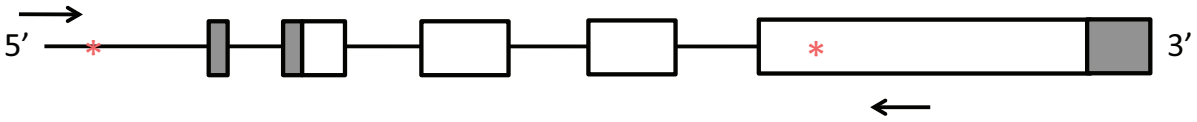
TALENs for the gene encoding Spirit were designed to target the start codon (Figure 16). The screening for mutants is based on the restriction enzyme site for BsmI found within this region. If the break occurs, this site will be disrupted. PCR amplification of this region, followed by digestion is used for screening the mutants.

The first approach using transgenic lines with TALEN constructs was performed. We used the germ line specific promoter *vasa*. In addition to this, the transgenic lines had an YFP or RFP marker under the Pax6 promoter, in the constructs for the left or right TALEN respectively. This allowed the screening of YFP/RFP progeny in which the DS break should occur. This approach did not yield any mutations in the target regions probably due to low expression of the construct under the vasa promoter or the fact that we found an 18 bp deletion in the promoter sequence.

The second approach was based on a recent publication that used TALEN-coding mRNA (Liu et al., 2012). We used the same design but with the T7 promoter. After performing *in vitro* transcription, TALEN-coding mRNA was injected into embryos. As a proof of principle, we obtained the vectors with TALENs designed to disrupt the *yellow* gene that was previously published (Liu et al., 2012). TALEN-coding mRNA for *yellow* was injected into embryos as a control for the experiment.

***CG2056* (*spirit*)**

combination 1-3



combination 2-4

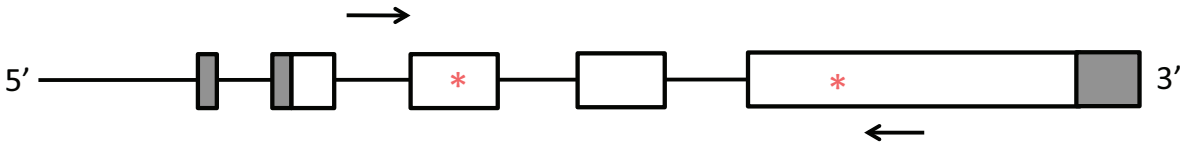


Figure 17 Schematic representation of CRISPR strategy

The asterisks represent which sequences are targeted by the sgRNA injected for *CG2056* (*spirit*). The arrows represent primers used to verify the mutations.

gene	combination	G0 flies	F1 lines tested	mutants
SPIRIT	1-3	37	121	0
	2-4	37	119	1

Table 2

Flies that were emerging showed mosaic *yellow* phenotype indicating that we were able to inactivate the *yellow* gene in somatic cells.

More than one hundred lines were tested, but unfortunately no mutations using this strategy were obtained. One possibility can be that there were no mutations in the germline. Our control for the *yellow* gene demonstrated the phenotype only in the somatic cells however we didn't check whether this was true for the germline. The other explanation is the low percentage of emerging flies upon injection of TALEN-coding mRNA, possibly due to the lethality of injections. Number of tested lines and mutants is in Table 1.

Since CRISPR/Cas9 system was developed in the lab and proved to be efficient, and due to the lack of time, we switched to this strategy.

CRISPR/Cas9 system

Another newly developed tool for genome editing is the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 system derived from bacteria where it functions as a defense system against invading viruses and plasmids (Bassett and Liu, 2014). It contains three main components: CRISPR RNA (crRNA), trans-activating CRISPR RNA (tracrRNA) and CRISPR-associated protein 9 (Cas9), a DNA nuclease. The tracrRNA will trigger Cas9 nuclease activity and crRNA will guide Cas9 to cleave the specific DNA sequence. This is achieved through base pairing between crRNA and the target DNA. This system is now being used in various different organisms and our strategy was based on using a single guide RNA (sgRNA)

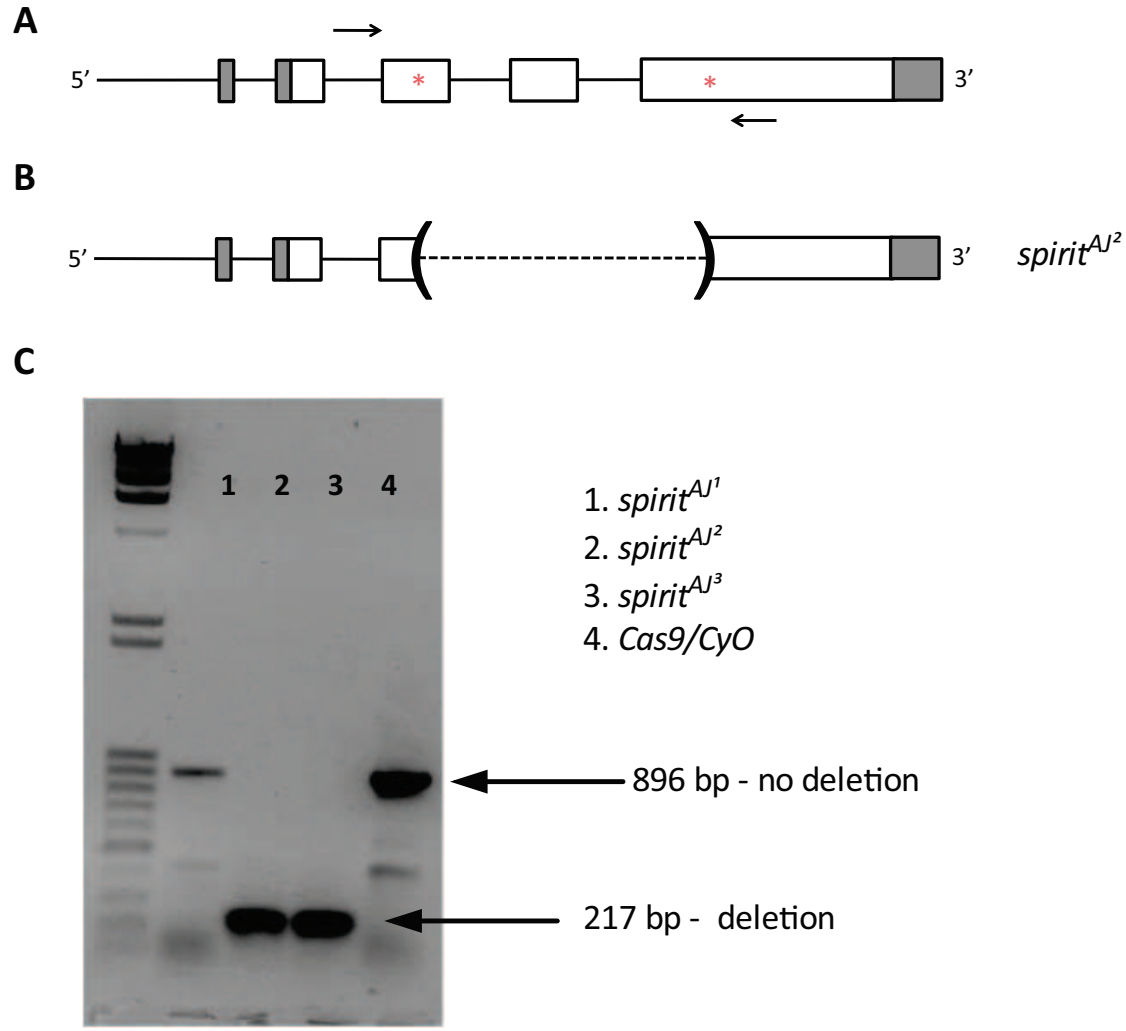


Figure 18 (to be continued on the next page)

comprising of the minimal crRNA and tracrRNA (Ren et al., 2013). The specificity is determined by a 20 bp sequence at the 5' end of the sgRNA. This sequence is designed based on our target sequence of interest. The only limitation is that it has to be followed by a specific motif, Protospacer Adjacent Motif (PAM) (NGG or NAG).

The strategy we used to generate mutants with CRISPR/Cas9 system consisted of injecting a pair of sgRNAs into transgenic line expressing Cas9 in the germ line. These sgRNAs would target specific regions of the gene, generating a deletion between them. Schematic representation of this strategy is depicted in Figure 17. We used two different combinations of sgRNA in order to generate two different mutants. Number of emerged, tested and generated mutants is in Table 2.

Using this strategy we obtained a mutant for Spirit.

To test whether the Spirit serine protease is involved in the activation of the Toll pathway we used the *spirit*^{AJ2} fly line, in which a deletion of 687 bp was generated by CRISPR system (Figure 18B and C). At the protein level this part includes 140 aa residues of the protein consisting of most part of the CLIP-domain, the activation site and His residue from the catalytic triad.

When *spirit*^{AJ2} flies were immune challenged with non-pathogenic Gram-positive bacterium *Micrococcus luteus*, we observed that *drosomycin* expression was comparable to the expression in non-injected Cas9 control line (Figure 18D). This result indicated that Spirit is not involved in the immune response against *Micrococcus luteus*.

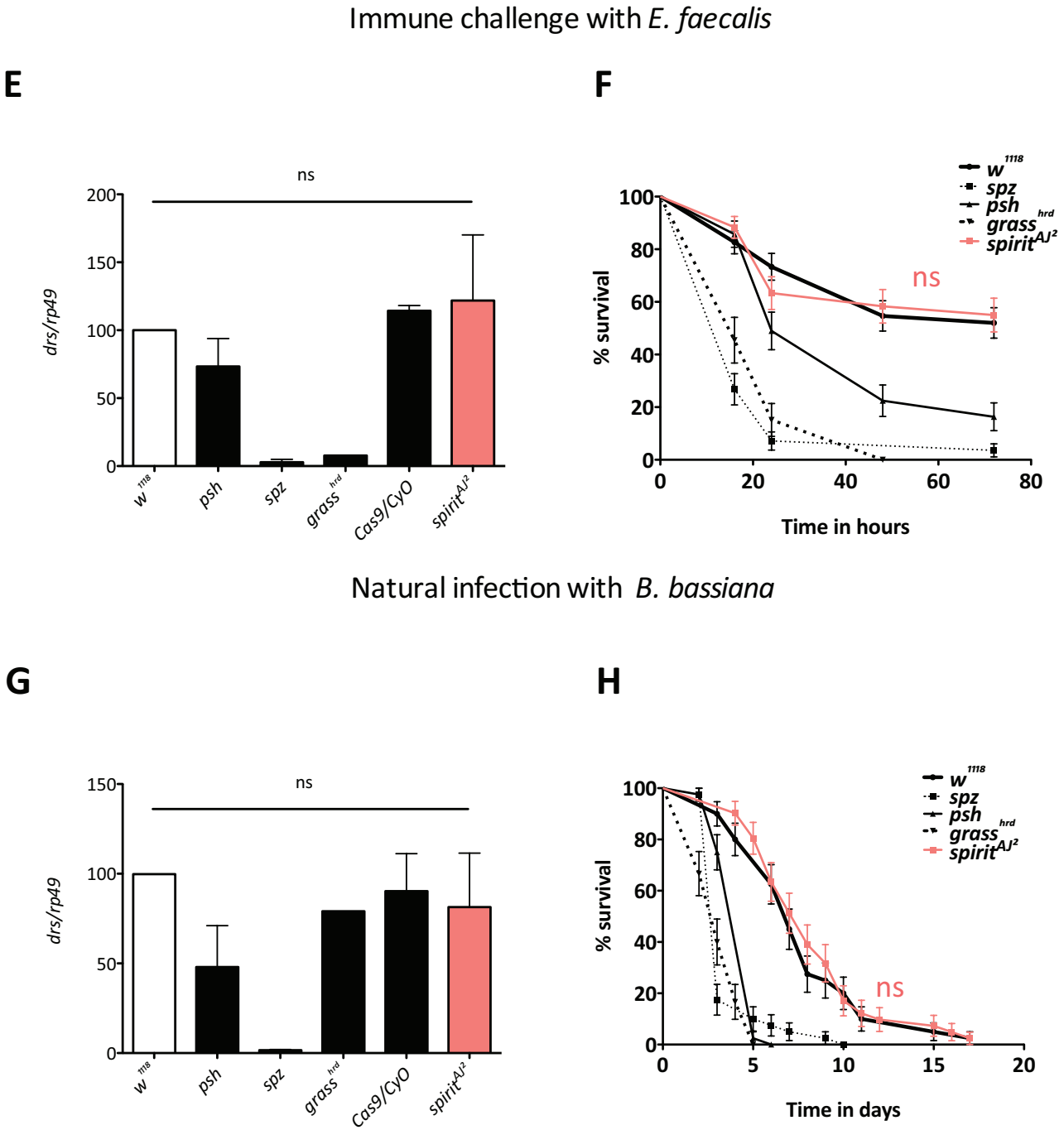


Figure 18 *spirit* is not involved in the Toll pathway activation

Schematic representation of genomic region of *CG2056* (*spirit*) (**A**) and deletion *spirit*^{AJ2} generated by CRISPR system (**B**) (the asterisks show the region targeted by the guide RNA; primers used to verify deletion are indicated with black arrows); Inactivation of *spirit* using CRISPR system shown by PCR with indicated primers (**C**); *drosomycin* expression 24 hours PI, normalized to *rp49* and in percentage to its expression in *w*^{1118} wild type flies after immune challenge with *M. luteus* (**D**) or *E. faecalis* (**E**) or after natural infection with *B. bassiana,* 48 hours PI (**G**). Survival rate after immune challenge with *E. faecalis* (**F**) or natural infection with *B. bassiana* (**H**). There is no statistical significance between survival of *w*^{1118} and *spirit*^{AJ2} flies (3 independent experiments)

When flies were immune challenge with pathogenic Gram-positive bacterium *Enterococcus faecalis*, they showed the same susceptibility to this infection as wild type flies (Figure 18F). *drosomycin* expression was comparable to that in wild type flies (Figure 18E). This result indicated that Spirit is not involved in the immune response against *Enterococcus faecalis*.

When flies were infected with fungus *Beauveria bassiana*, they showed the same susceptibility to this infection as wild type flies (Figure 18H). *drosomycin* expression was comparable to that in wild type flies (Figure 18G). This result indicated that Spirit is not involved in the immune response against *Beauveria bassiana*.

Taken together these results demonstrate that Spirit is not involved in the activation of the Toll pathway after immune challenge with these specific pathogens. However, these are preliminary results. Furthermore we cannot exclude involvement of this serine protease in the immune response against other pathogens, as well as involvement in other proteolytic cascades. Further investigation and characterization of this serine protease is necessary.

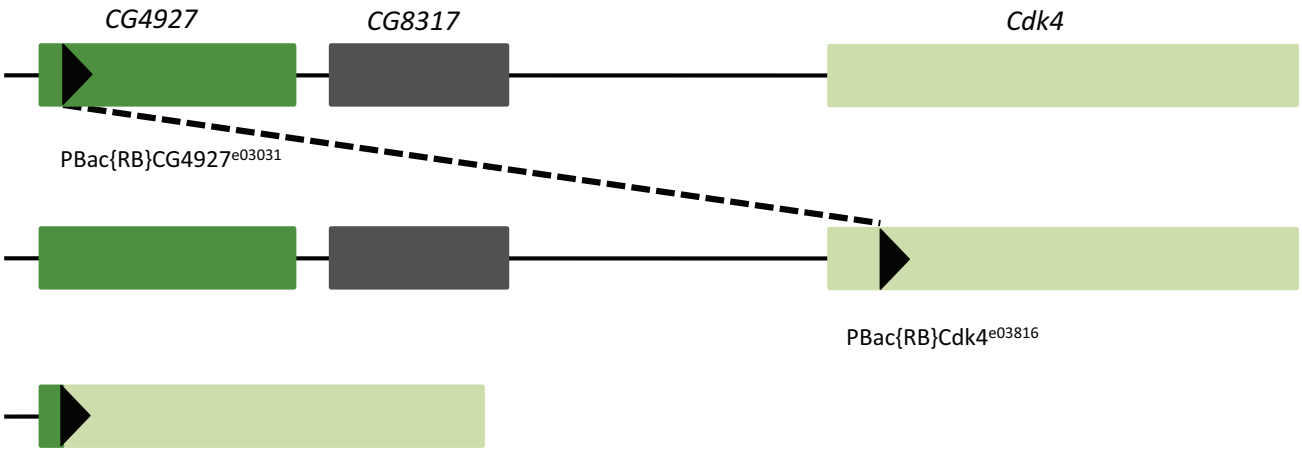


Figure 19 Schematic representation of FLP/FRT strategy for *CG4927*

Candidate gene *CG4927*

RNA interference

RNAi-mediated knockdown was performed as previously described. We obtained VDRC lines with inverted repeats containing inducible *UAS-RNAi* against *CG4927* (*P{KK110662}VIE-260B*). RNAi mediated knockdown was not efficient. Levels of gene expression were the same as in wild type flies. Before completely excluding RNAi as a strategy, Gal4 driven RNAi knockdown was enhanced during a course of time, keeping the flies for longer than three days at 29 °C, as well as performing more than three heat shocks. These experiments were repeated few times, showing no efficient reduction in gene expression of targeted genes.

FLP/FRT system

To generate a mutant in *CG4927* we used FLP/FRT system that is based on the homologous recombination between short Flippase Recognition Target (FRT) sites (Golic, 1994). We used the availability of P-bac element insertion lines carrying FRT sites in *CG4927* (*Pbac{RB}CG4927*^{e03031}) and in *Cdk4* (*Pbac{RB}Cdk4*^{e03816}) (Figure 19). Recombination between these FRT sites would generate a deletion that would include not only *CG4927*, but also *CG8317* (gene of unknown function). For the induction of recombination we used a line with a heat shock flippase (*P{70FLP}3F*) and performed heat shocks (1-2 hours at 37 °C) at different times during

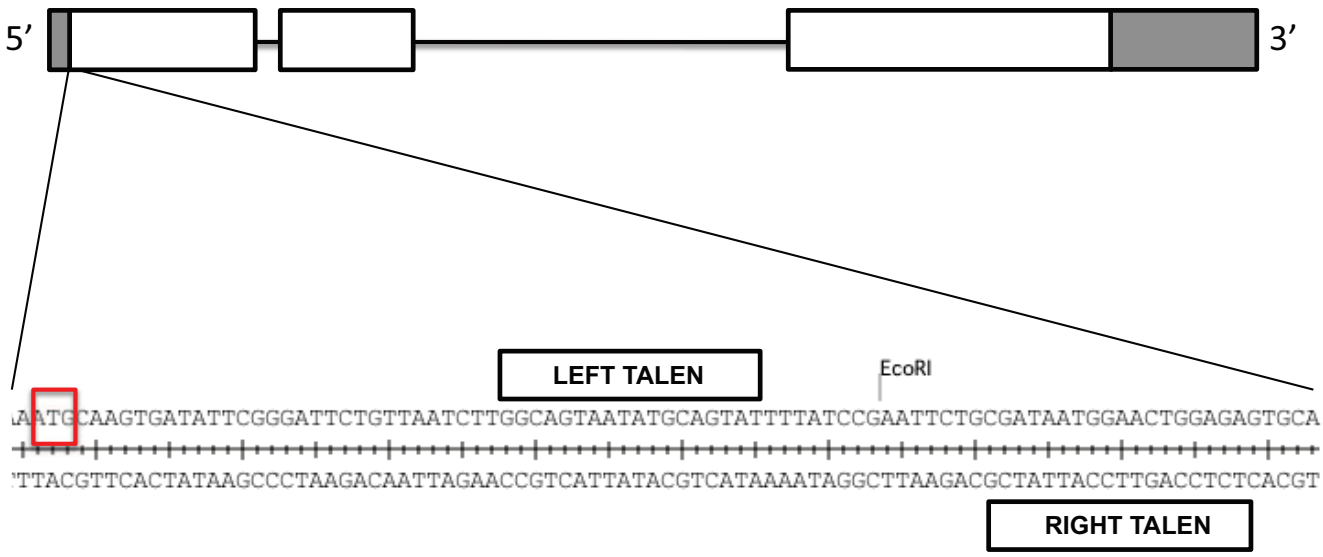


Figure 20 Schematic representation of TALEN strategy

Start codon with EcoRI recognition site for *CG4927*.

gene	method	lines tested	mutants
CG4927	TALEN-coding mRNA	250	0

Table 3

development (at 48 and 72 hours after egg laying). Unfortunately, no mutants were generated.

Transcription Activator-Like Effector Nucleases (TALENs)

We used the second approach for TALENs, injection of TALEN-coding mRNA to generate mutants in *CG4927* (Figure 20). The target site is around 50 bp downstream of the start codon because of the availability of the restriction enzyme site for EcoRI.

More than two hundred lines were tested, but unfortunately no mutations using this strategy were obtained. Number of tested lines and mutants is in Table 3.

Since CRISPR/Cas9 system was developed in the lab and proved to be efficient, and due to the lack of time, we switched to this strategy.

CRISPR/Cas9 system

We used the same strategy for generating mutants in *CG4927* as previously described. Schematic representation of this strategy is depicted in Figure 21. We used two different combinations of sgRNA in order to generate two different mutants. Number of emerged, tested and generated mutants is in Table 4. So far, no mutants were obtained and we are still pursuing the work on generating a mutant for *CG4927*.

CG4927

combination 1-3

combination 2-4

Figure 21 Schematic representation of CRISPR strategy

The asterisks represent which sequences are targeted by the sgRNA injected for *CG4927*. The arrows represent primers used to verify the mutations.

gene	combination	G0	F1 lines tested	mutants
CG4927	1-3	23	109	0
	2-4	35	131	0

Table 4

Materials and methods

(See Materials and methods in the second part of Results chapter – The serine protease homolog *spheroide* is involved in sensing of pathogenic Gram-positive bacteria.)

Fly strains

Flies carrying *UAS-RNAi* transgene against *CG3700* (*P{KK106926}VIE-260B*), *CG4927* (*P{KK110662}VIE-260B*) and *CG2056* (*P{GD3285}v5497*, *P{KK112114}VIE-260B*) were obtained from Vienna Drosophila Resource Center (VDRC). Flies carrying P-element insertion in *CG3700* (*P{XP}CG3700*^{d06151}) were obtained from Bloomington Stock Center, in *CG12065* (*P{GawB}CG12065*^{NP6114}) was obtained from Drosophila Genomics Resource Center (DGRC). Flies for the deficiency in *CG3700* (*Df(2R)BSC865*) were obtained from Bloomington Stock Center. Flies carrying P-bac element insertions in *CG4927* (*Pbac{RB}CG4927*^{e03031}) were obtained from Bloomington Stock Center, in *Cdk4* (*Pbac{RB}Cdk4*^{e03816}) was obtained from Exelixis at Harvard Medical School. Flies carrying heat shock flippase (*P{70FLP}3F*) were obtained from Bloomington Stock Center. Flies carrying transposase (*P{Δ2-3}*) were obtained from Bloomington Stock Center.

PCR analysis of *CG3700* inactivation

Primers used to verify the P-element insertion in the gene *CG3700* and its inactivation: 5Ri (5'-TGACACYYACCGCYYGACA-3'), 3Fi (5'-CCTCGATATACAGACCGATAAAAC-3'), CG3700 5' (5'-AGCCGCAGCAACCACATCA-3'),

CG3700 lower (5'-TGGACATTGAGCCGGCACAGAAC-3'), and for *rp49* control IMU85 (5'-GTGTATTCCGACCACGTTACA-3'), IMU86 (5'-ATACAGGCCCAAGATCGTGA-3'), Tm=60 ^{o}C. Amplification with CG3700 5' and CG3700 lower gives a fragment of 1500 bp in wild type flies.

TALEN design and detection of mutations

TALEN constructs for transgenic lines were made by Eric Marois using Golden Gate method as previously described (Geißler et al., 2011; Smidler et al., 2013) and injected at Genetic Services, Inc. (lines 4L and 4R). For injections of TALEN-coding mRNA the same strategy was used for cloning followed by *in vitro* transcription using mMESSAGE mMACHINE® T7 Transcription Kit (Ambion®) and Poly(A) Tailing Kit (Ambion®). TALEN sequences: for *CG2056* (*spirit*) left talen (5'-GGCGCTGATCAAGAACGG-3') and right talen (5'-GGGTATGCAATGGGTAGA-3'); for *CG4927* left talen (5'-GGCAGAATATGCAGTAT-3') and right talen (5'-CTCCAGTTCCATTATCG-3'). Primers used for verifying mutations in *CG2056* (*spirit*): SpTAL FW (5'AAGCGACCGAGCAGATAATACAGC-3') and SpTAL RV (5'GCGGGCGTGGCGTGAACT-3'), Tm=62 ^{o}C, followed by digestion with BsmI; for *CG4927:* FW2 (5'-GGTGTGTGAATTCGCTAACT-3') and RV2 (5'AGCACGAACTTGGGATGGATGATA-3'), Tm=57 ^{o}C, followed by digestion with EcoRI.

CRISPR design and detection of mutations

We used CRISPR/Cas9 strategy that was previously described (Ren et al., 2013). *spirit*^{AJ2} and *spirit*^{AJ3} mutants were generated in this study. Target CRISPR sequences used: CrisprCds1 (CC1) (5'-GAAGGGCACCTGCCGGCGCATGG-3') and

CrisprCds3 (CC3) (5’-GACAATCTAACCCTGACCGAGGG-3’). Primers used for verifying deletions in *spirit*: Spirit2FW (5’-GCCACGCCCGCCATAAGTCCGC-3’) Spirit2RV (5’-GGCTTGGCCGCCGTCTCCAGCT-3’), Tm=60 ^{o}C. Amplification in wild type flies gives a fragment of 896 bp and in flies with deletion a fragment of 217 bp. Target CRISPR sequences used for *CG4927*: CrisprIntergenic1 (CI1) (5’-TGCCTAATTTTAGGGTAGGG-3’), CrisprCds1 (CC1) (5’-GAATTCTGCGATAATGGAAC-3’), CrisprCds2 (CC2) (5’-GGGATCGCGGTCCACCAGTG-3’) and CrisprCds3 (CC3) (5’-GACACCGCAGACCAATCCAT-3’). Primers used for verifying deletions in CG4927: FW4 (5’-gtgacttttggcggcatttaaatt-3’), FW2 (5’CGACATTGCCGTGGTGGAGCTG-3’) and RV2 (5’GAGCTGCCACATGGCATGGCCA-3’), Tm=57 ^{o}C.

2. The serine protease homolog *spheroide* is involved in sensing of pathogenic Gram-positive bacteria

The serine protease homolog *spheroide* is involved in sensing of pathogenic Gram-positive bacteria

Jelena Patrnogic, Florian Veillard, Vincent Leclerc and Jean-Marc Reichhart*

UPR9022 du CNRS, Institut de Biologie Moléculaire et Cellulaire, Université de Strasbourg, 67084 Strasbourg Cedex, France

*Corresponding author: Jean-Marc Reichhart

In *Drosophila*, recognition of pathogens such as Gram-positive bacteria and fungi triggers the activation of the proteolytic cascades of the Toll pathway. This response can be achieved by either detection of pathogen associated molecular patterns (PAMPs) or by sensing microbial proteolytic activities ("danger signals"). Previous data suggested that certain serine protease homologs (serine protease folds that lack an active catalytic triad) are involved in the pathway. We generated a null mutant of the serine protease homolog *spheroide (sphe).* These mutant flies are susceptible to *Enterococcus faecalis* infection and unable to activate the Toll pathway fully. Sphe is required to activate the Toll pathway after challenge with pathogenic Gram-Positive bacteria, but not in response to fungi or non-pathogenic bacteria. Sphe functions in the danger signal pathway, downstream or at the level of Persephone.

INTRODUCTION

The fruit fly, *Drosophila melanogaster*, spends its life among decaying matter and rotten fruit, where it coexists with different microorganisms. One of the main characteristics of the immune response of *Drosophila melanogaster* is the challenge-induced synthesis and secretion of antimicrobial peptides (AMPs). This response involves the activation of two signal transduction cascades – the Toll and IMD pathways (Lemaitre and Hoffmann 2007). Gram-positive bacteria and fungi activate the Toll pathway, whereas Gram-negative bacteria activate the IMD pathway. In both cases, signaling leads to the activation of NF-κB transcription factors and expression of target genes including AMPs.

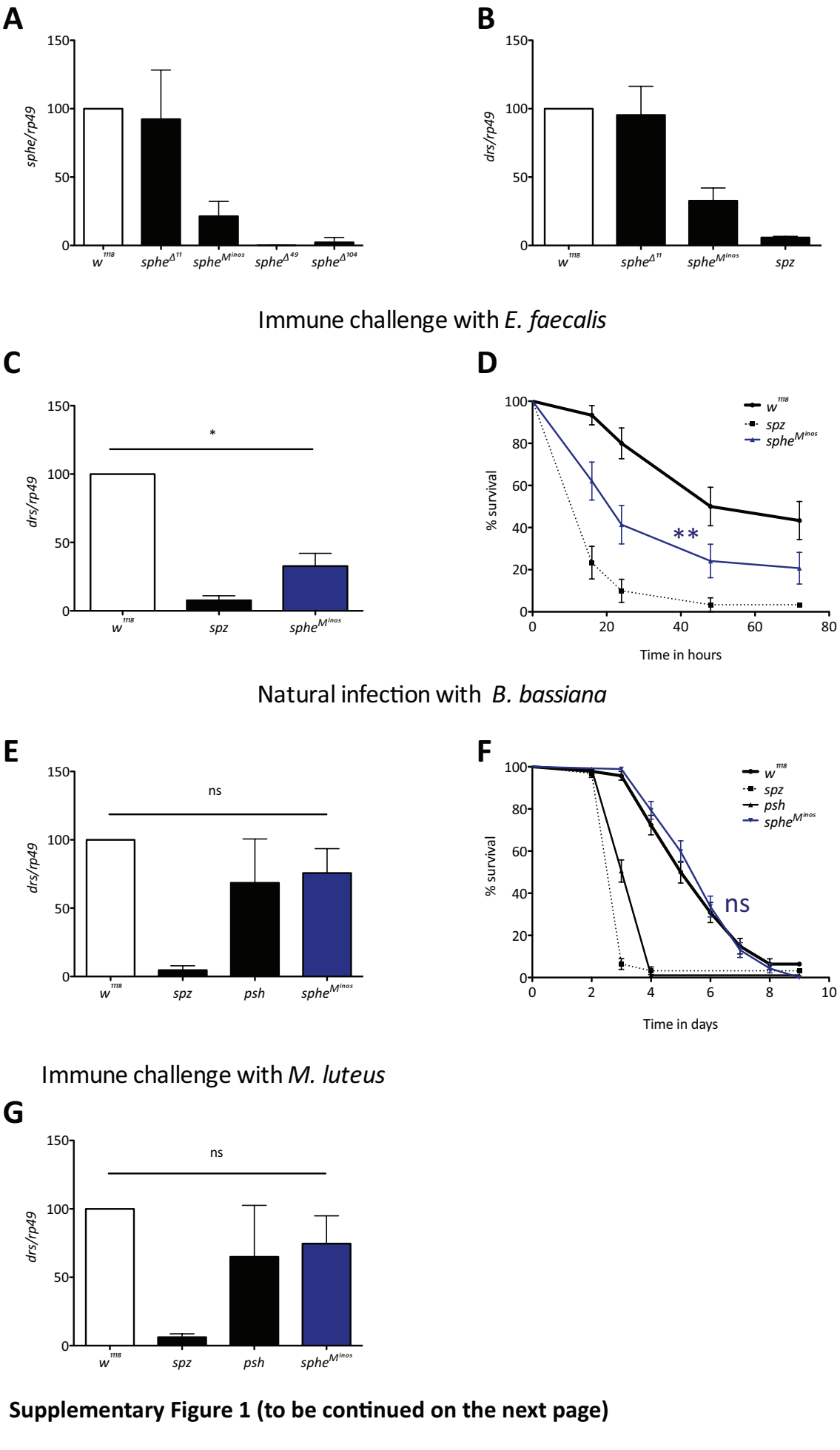
In the late 1980's, Charles Janeway proposed that the innate immune mechanisms are essential for the early detection and defense against infection. These mechanisms discriminate between self and microbial non-self. Janeway proposed the existence of germ-line encoded pathogen recognition receptors (PRRs) that recognize conserved signature molecules expressed by pathogens, referred to as Pathogen Associated Molecular Patterns (PAMPs) (Janeway, 1989).

A few years later, Polly Matzinger proposed the "danger signal" hypothesis. This hypothesis proposed that the activation of immune mechanisms is not due to discrimination between self and non-self, but rather to sensing of danger signals: either recognition of pathogens, or alarm signals produced by microbial activities or by the host's own damaged cells or tissues (Matzinger, 1994).

The Toll pathway can be activated in two ways – recognition of PAMPs by circulating Pathogen Recognition Receptors (PRRs) in the hemolymph; or by virulence factors, mostly proteases, secreted by the pathogens. This activation triggers proteolytic cascades in the hemolymph. The terminal protease in the cascade cleaves Spaetzle to its activated ligand form, which is able to bind the Toll receptor and activate the intracellular pathway. Depending on the triggering signal, two proteolytic cascades can be distinguished. First, the recognition cascade activated by PAMPs which includes 3 serine proteases, ModSP (Buchon et al., 2009), Grass (El Chamy et al., 2008; Kambris et al., 2006) and SPE (Jang et al., 2006b). Secondly, the "danger" signal cascade can be activated by pathogen-encoded, secreted proteases. Such abnormal protease activity indicates that potentially dangerous changes are happening. Danger signaling involves the serine protease Persephone (Psh) (El Chamy et al., 2008).

There are over two hundred genes coding for serine proteases (SPs) and serine protease homologs (SPHs) in the *Drosophila* genome (Ross et al., 2003). SPHs maintain the serine protease fold but lack amidase activity since one, or more of the catalytic triad residues is missing (Ross et al., 2003). The physiological functions of SPHs are poorly understood, although they have been implicated in different arthropod immune responses, in the horseshoe crab (Kawabata et al., 1996), *Manduca sexta* (Felföldi et al. 2011) and *Anopheles gambiae* (Dimopoulos et al., 1997; Povelones et al., 2013).

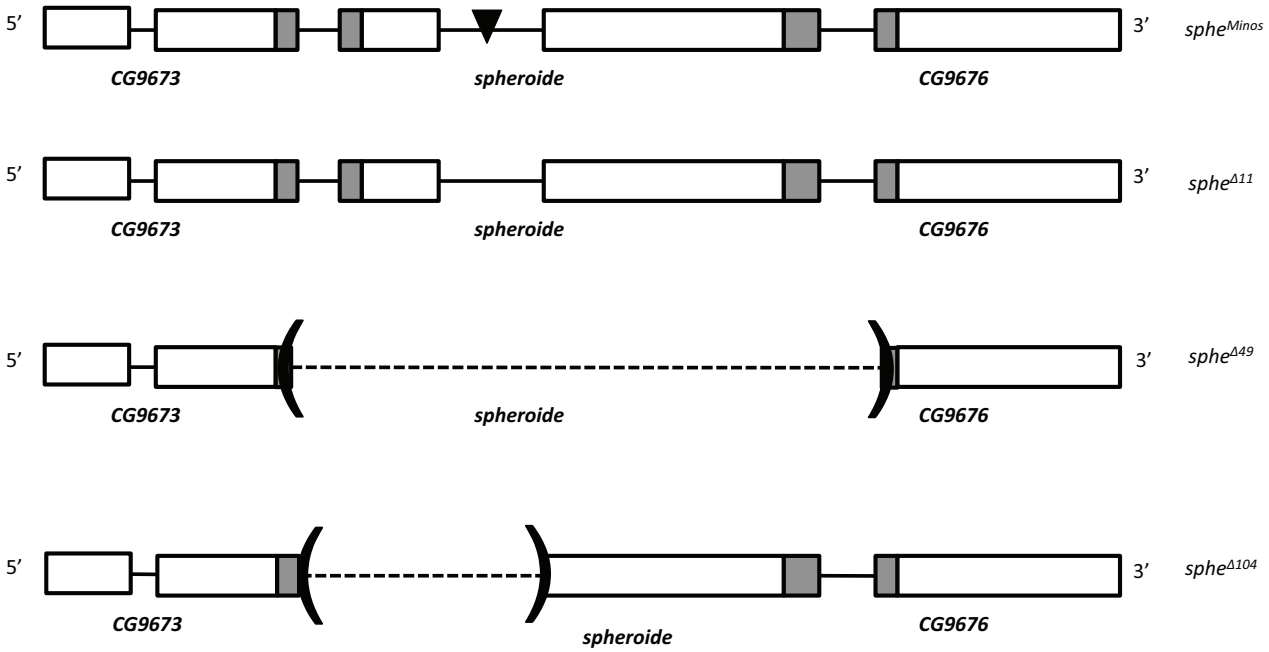
We identified the protease Grass as being required for Toll pathway activation downstream of PRRs (El Chamy et al., 2008). Grass was initially identified



Supplementary Figure 1 (to be continued on the next page)

during an RNAi based screen of serine proteases and serine protease homologs (Kambris et al., 2006), but its function was incorrectly assigned, probably due to the incomplete knockdown of the gene mediated by RNAi. We decided to verify the function of the other candidates identified in this work and focused on the serine protease homolog Spheroide (Sphe). It has been reported that Sphe is involved in the activation of Toll pathway. The knockdown of *sphe* by RNAi induced the same phenotype upon immune challenge as that of SPE, implying that Sphe might function as an adaptor or regulator of SPE (Kambris et al., 2006).

Here, we use a null mutant of *sphe* to demonstrate that Sphe is involved in the activation of Toll pathway. By using protease-deficient bacteria we conclude that Sphe is sensing the virulence factors (proteases) produced by pathogenic Gram-positive bacteria. Furthermore, using flies that are double mutants for both *sphe* and *grass*^{hrd}, we show that Sphe is involved in the "danger" signal cascade.



Supplementary Figure 1

spheroide expression 24 hours PI, normalized to *rp49* and in percentage to its expression in w^{1118} wild type flies (ctrl) (**A**). *drosomycin* expression 24 hours PI, normalized to *rp49* and in percentage to its expression in w^{1118} wild type flies after immune challenge with *E. faecalis* (w^{1118} vs. $sphe^{minos}$ $p<0.0335$) (**B** and **C**) or natural infection with *B. bassiana* 48 hours PI (**E**) or *M. luteus* (**G**). Survival rate after immune challenge with *E. faecalis* (*ctrl* vs. $sphe^{Minos}$ $p<0.0076$) (**D**) or natural infection with *B. bassiana* (3 independent experiments) (**F**). Schematic representation of deletions obtained by excision of Minos insertion element (**H**).

RESULTS

Sphe is required in the activation of the immune response after a challenge with *Enterococcus faecalis*

Due to limitations caused by RNAi-mediated knockdown, we looked for another way of inactivating *sphe* and made use of a *Minos* transposon insertion in the gene. In the fly line *Mi{ET1}spheroide*^{MB11555} a Minos element is inserted 412 bp downstream of the start codon, in an intronic sequence. This insertion reduced the expression of the *sphe* transcript compared to wild type (Supplementary Figure 1A). Flies in which *sphe* expression was reduced were more susceptible than wild type flies to infection with the Gram-positive bacterium *Enterococcus faecalis*. *Sphe* mutants showed reduced levels (30%) of *drosomycin (drs)* expression compared to wild type flies (the transcript level of the *drs* antimicrobial peptide is used here as an assay for Toll pathway activation) (Supplementary Figure 1C and D).

To confirm that the susceptibility phenotype was due to the insertion of the *Minos* transposon, we excised the insertion element and generated null alleles. We obtained a line with precise excision of the element (*sphe*$^{\Delta 11}$) that expresses wild type *sphe* mRNA levels. When *sphe*$^{\Delta 11}$ flies were challenged with *Enterococcus faecalis* they showed normal expression of *drs* (Supplementary Figure 1A and B). This excision line is used as a wild type control in subsequent experiments (*ctrl*). This demonstrates that the *Minos* insertion was indeed responsible for the susceptibility phenotype. We also obtained two imprecise excisions, *sphe*$^{\Delta 49}$ and *sphe*$^{\Delta 104}$, in which no *sphe* expression was detected. The *sphe*$^{\Delta 49}$ deletion includes the entire transcript as well as 974 bp upstream that include 46 bp of the 3'UTR of *CG9673*, and 307 bp

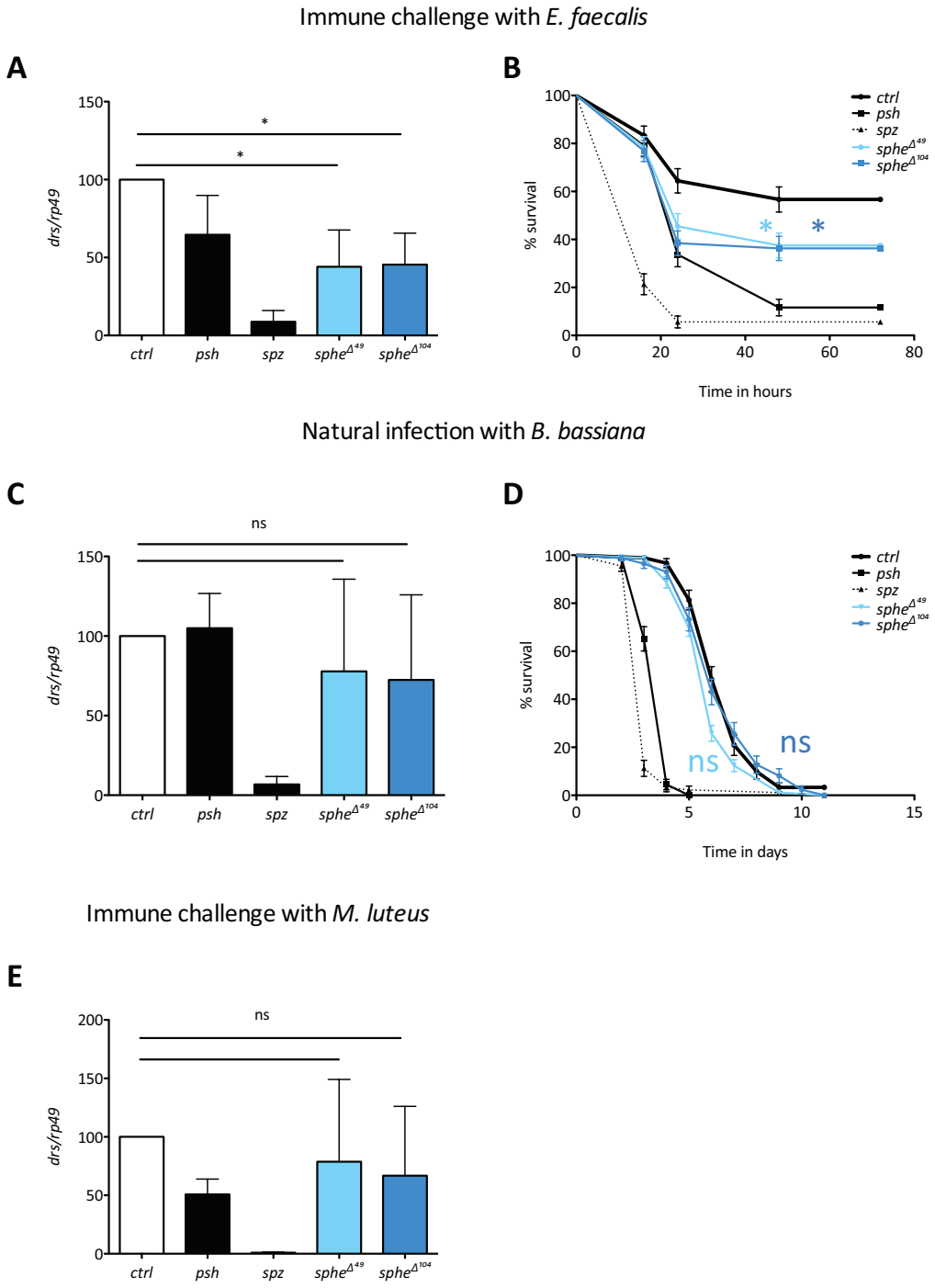


Figure 1 Sphe is involved in activation of immune response after a challenge with *E. faecalis*

drosomycin expression 24 hours PI, normalized to *rp49* and in percentage to its expression in $sphe^{\Delta 11}$ wild type flies (ctrl) after immune challenge with *E. faecalis* (*ctrl* vs. $sphe^{\Delta 49}$ p<0.0261, *ctrl* vs. $sphe^{\Delta 104}$ p<0.02) (**A**) or after natural infection with *B. bassiana* 48 hours PI (**C**) or immune challenge with *M. luteus* (**E**). Survival rate after immune challenge with *E. faecalis* (*ctrl* vs. $sphe^{\Delta 49}$ p<0.0132, *ctrl* vs. $sphe^{\Delta 104}$ p<0.0065) (**B**) or natural infection with *B. bassiana* (**D**) (3 independent experiments).

downstream that include the 5'UTR of *CG9676* (Supplementary Figure 1H). The *sphe*$^{\Delta 104}$ deletion starts at the *Minos* insertion site and includes 835 bp of upstream sequence. At the protein level, the first 74 amino acids residues are missing, which include the signal peptide and 50 amino acids residues of the catalytic domain, including the His residue from the catalytic triad. Both deletions are therefore null alleles of *sphe*.

When *sphe* null mutant flies were challenged with pathogenic Gram-positive bacterium *Enterococcus faecalis* we observed a significant decrease of *drs* levels 24 hours after infection compared to that of wild type flies (*drs* reaches 45% of wild type level) (Figure 1A). Furthermore *sphe* flies are more susceptible to this immune challenge than wild type flies (Figure 1B). Since both null alleles *sphe*$^{\Delta 49}$ and *sphe*$^{\Delta 104}$ show the same phenotype, we will describe only the results obtained with *sphe*$^{\Delta 104}$.

When *sphe* null mutant flies, were challenged with the non-pathogenic Gram-positive bacterium, *Micrococcus luteus*, or by natural infection with the entomopathogenic fungus *Beauveria bassiana, drs* expression was comparable to that in wild type flies (Figure 1C and E). Accordingly, *sphe* null mutant flies showed the same susceptibility to *Beauveria bassiana* infection as wild type flies (Figure 1D).

To confirm that the phenotype we observed is due to *sphe* inactivation, we overexpressed Sphe with the *UAS-Gal4* system using the ubiquitous *Actin5C>Gal4* driver. Sphe expressing flies are viable and show no obvious phenotype. Sphe over expression does not induce the Toll pathway as measured by levels of *drs* mRNA. We therefore expressed Sphe in *sphe* mutant background, and observed the rescue of the phenotype as assayed by the induction of *drs* expression in response to

Immune challenge with *E. faecalis*

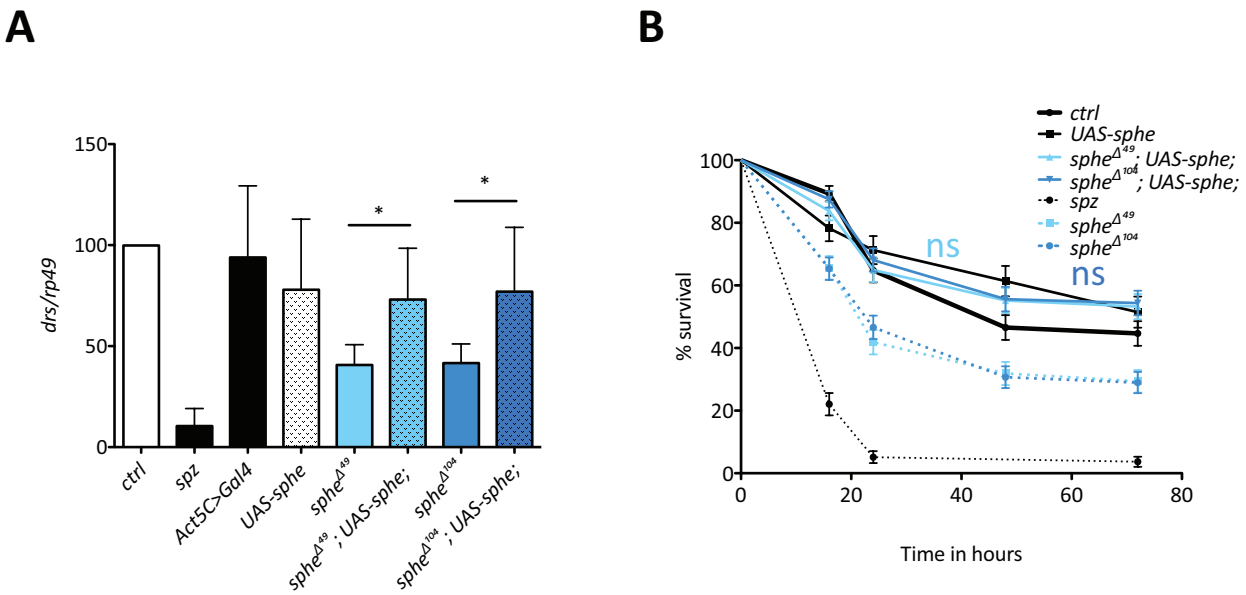


Figure 2 Rescue of *sphe* mutant phenotype using Gal4/UAS system

drosomycin expression 24 hours PI, normalized to *rp49* and in percentage to its expression in w^{1118} wild type flies (ctrl) after immune challenge with *E. faecalis* ($sphe^{\Delta 49}$ vs. $sphe^{\Delta 49}$; *UAS-sphe* p<0.0103, $sphe^{\Delta 104}$ vs. $sphe^{\Delta 104}$;*UAS-sphe* p<0.0229) (**A**). Survival rate after immune challenge with *E. faecalis* (**B**) (3 independent experiments). For over expression of *sphe* we used the ubiquitous *Act5C>Gal4* driver to activate *UAS-sphe*.

Enterococcus faecalis infection, as well as an enhanced survival to the infection (Figure 2A and B).

Taken together, these data show that Sphe is involved in Toll pathway activation and is required to activate a full and efficient response to the pathogenic Gram-positive bacterium *Enterococcus faecalis*.

Sphe is involved in the "danger" signal Toll activation cascade

Enterococcus faecalis is a pathogenic Gram-positive bacterium that activates the Toll pathway via both proteolytic branches, through recognition of Lys-type peptidoglycan (Leulier et al., 2003b), as well as through production of virulence factors that activate the "danger" signal cascade (El Chamy et al., 2008). To assess in which of these branches Sphe is functioning, we generated double mutants $sphe^{\Delta 104};grass^{hrd}$ in which the recognition cascade is blocked and we challenged these flies with *Enterococcus faecalis.* The levels of *drs* expression 24 hours after immune challenge are significantly decreased (*drs* reaches 20% of wild type level) compared to both *sphe* and $grass^{hrd}$ single mutants, to a level comparable to that of *spz* mutants flies (Figure 3A). This additive effect indicates that Sphe is acting in a parallel pathway to Grass, in the "danger" signal Toll activation cascade.

After immune challenge *Enterococcus faecalis* activation of Toll pathway is achieved via both proteolytic branches. *Enterococcus faecalis* produces several virulence factors, including cytolysin, aggregation substance, the zinc metalloprotease gelatinase GelE, and the serine protease SprE (Garsin et al., 2001;

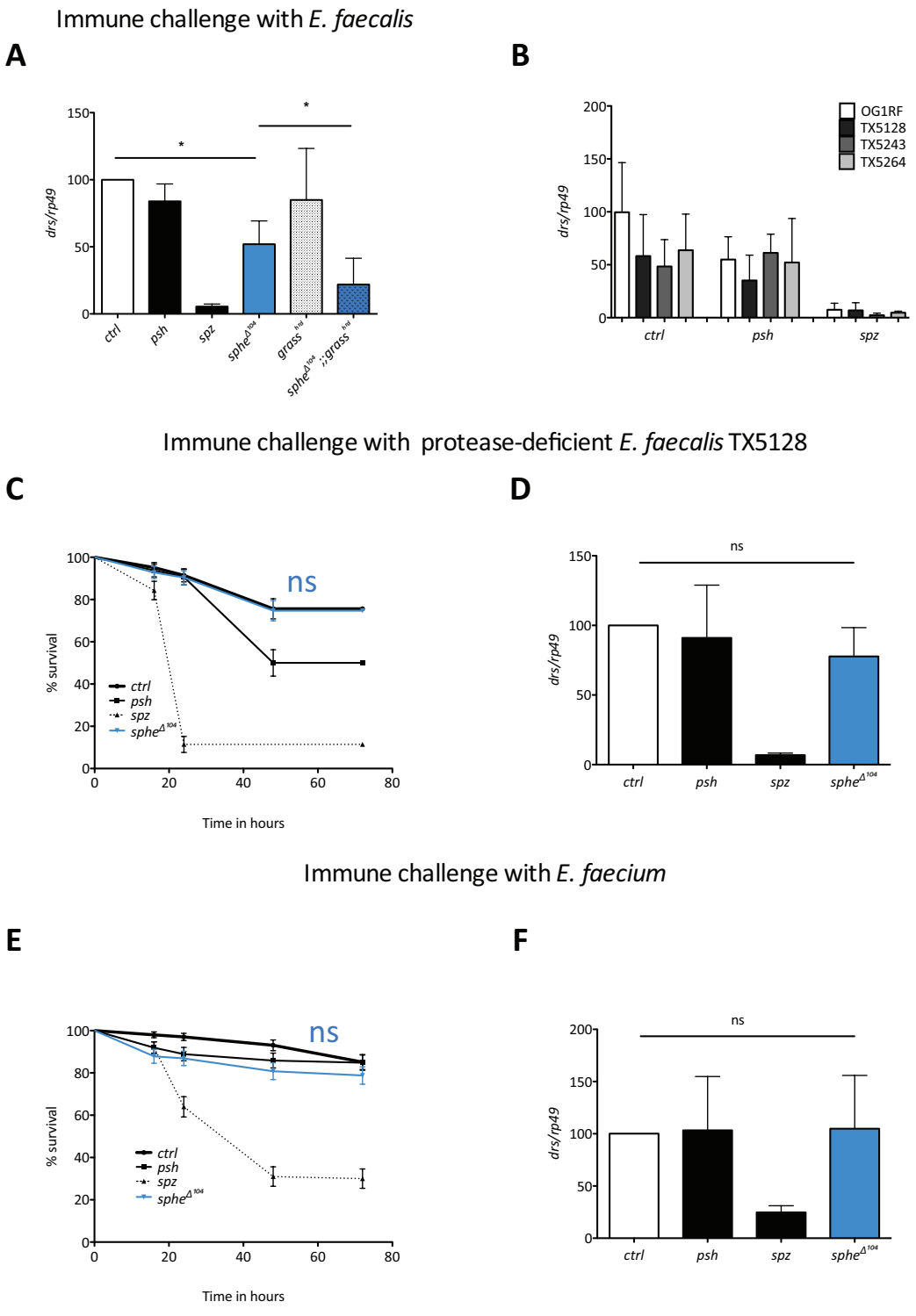
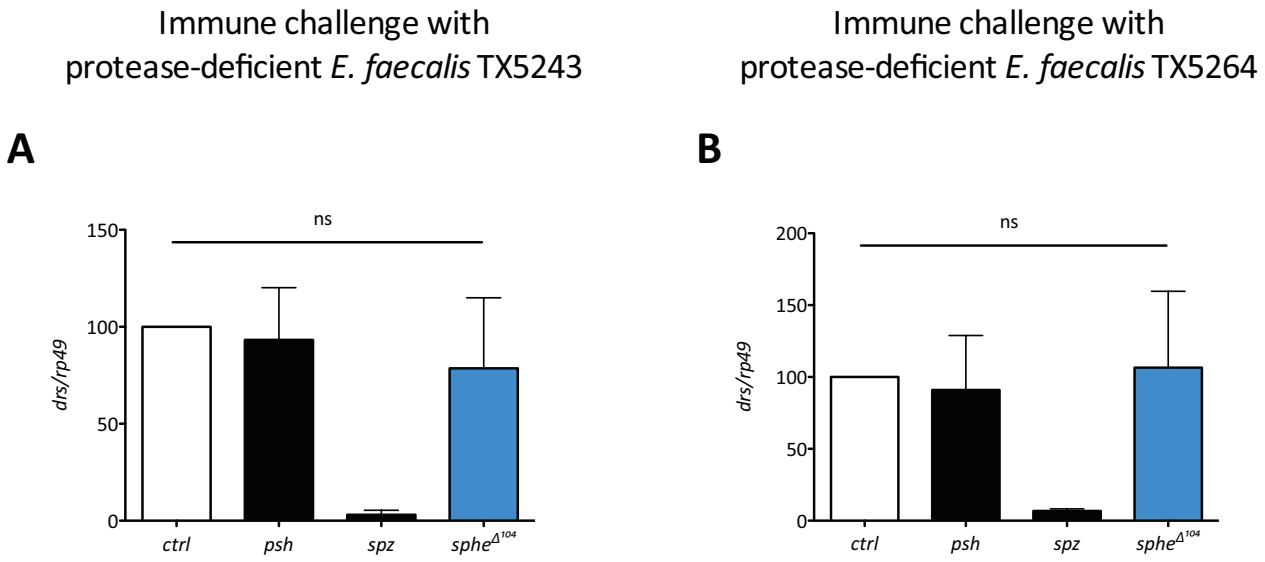


Figure 3 Sphe is involved in danger signal cascade

drosomycin expression 24 hours PI, normalized to *rp49* and in percentage to its expression in $sphe^{\Delta 11}$ wild type flies (ctrl) after immune challenge with *E. faecalis* in $sphe^{\Delta 104}$ and $grass^{hrd}$ single and double mutants (*ctrl* vs. $sphe^{\Delta 104}$ p<0.0178. *ctrl* vs. $sphe^{\Delta 104};;grass^{hrd}$ p<0.0264) (**A**). Comparison of *drs* expression after immune challenge with wild type and protease-deficient bacteria (**B**). Survival rate (**C**, **E**) and *drs* expression (**D**, **F**) in $sphe^{\Delta 104}$ flies infected with protease-deficient strain of *E. faecalis* TX5128, or with *E. faecium* (3 independent experiments).

Sifri et al., 2002). We focused on the secreted extracellular proteases GelE and SprE as potential virulence factors that might be sensed by the "danger" signal cascade. To confirm the involvement of *sphe* in this danger signal sensing, we used protease-deficient strains of *Enterococcus faecalis* that were mutant for either *gelE* (TX5264), *sprE* (TX5243), or both *gelE* and *sprE* (TX5128) (Qin et al., 2000; Sifri et al., 2002). We observe slight, but reproducible, reductions in *drs* levels when wild type flies are challenged with protease-deficient bacteria compared to those challenged with wild type bacteria. The observed decrease in *drs* levels is similar to the one observed in *psh* mutant flies challenged with wild type bacteria suggesting that these proteases are required for activating the danger signal pathway. This is confirmed with the fact that there is no additive effect when *psh* mutant flies are challenged with protease-deficient bacteria compared to the same infection in wild type flies. Furthermore, the activation of Toll pathway we observe in *psh* mutant flies is only due to the PRR cascade (Figure 3B).

After protease-deficient immune challenge, *sphe* mutants behave as wild type flies and show no susceptibility to the protease-deficient bacteria (Figure 3C); the levels of *drs* 24 hours after infection are as in wild type controls indicating normal activation of Toll pathway (Figure 3D). The same result was found using either of the single mutants for *gelE* or *sprE* (Supplementary Figure 2A and B) indicating that both of these virulence factors contribute to the activation of Toll pathway. We confirmed this observation by using the non-pathogenic bacterium *Enterococcus faecium* that lacks these virulence factors (Qin et al., 2012) and is closely related to *E. faecalis*. After immune challenge with *E. faecium, sphe* mutant



Supplementary Figure 2

drosomycin expression 24 hours PI, normalized to *rp49* and in percentage to its expression in *sphe*$^{\Delta 11}$ wild type flies (ctrl) protease-deficient *E. faecalis* TX5243 (**A**) or TX5264 (**B**).

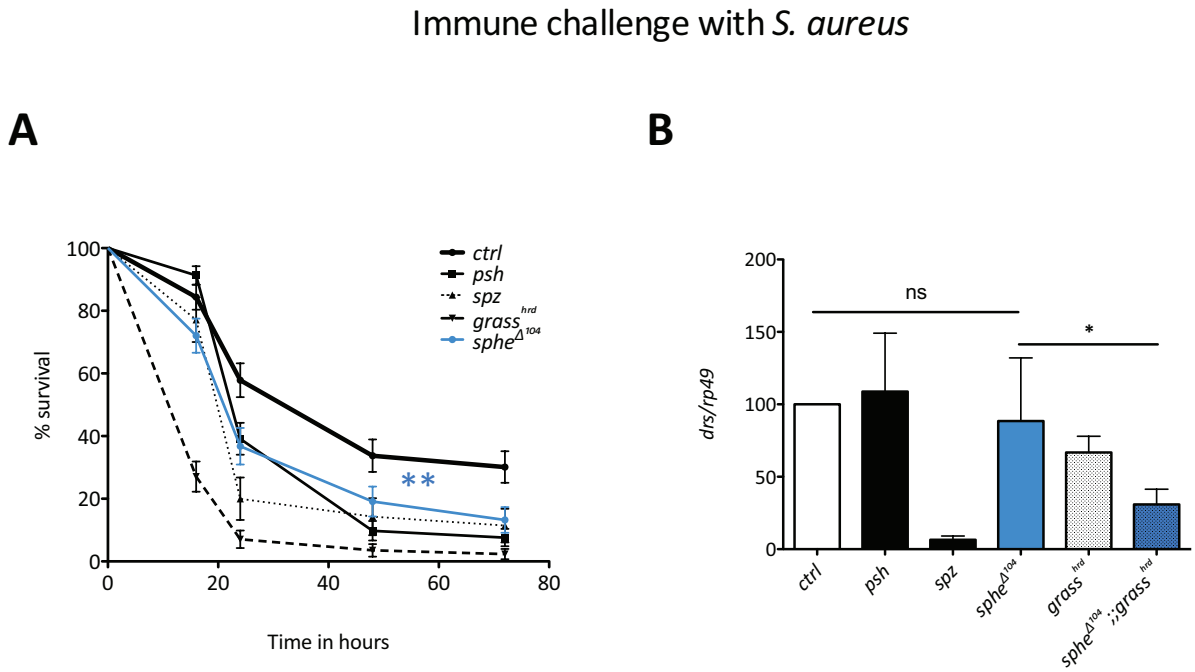


Figure 4 Sphe is involved in sensing *S. aureus*

Immune challenge with *S. aureus*: Survival rate (*ctrl* vs. *sphe*$^{\Delta 104}$ p<0.0038) (**A**). *drosomycin* expression 24 hours PI, normalized to *rp49* and in percentage to its expression in *sphe*$^{\Delta 11}$ wild type flies (ctrl) (*ctrl* vs. *sphe*$^{\Delta 104}$*;;grass*^{hrd} p<0.0144) (**B**) (3 independent experiments).

flies show no susceptibility (Figure 3E) and *drs* levels 24 hours after infection are as in wild type controls indicating normal activation of Toll pathway (Figure 3F). Taken together these data demonstrate that Sphe is involved in sensing proteases produced by *Enterococcus faecalis* for Toll pathway activation.

Sphe is involved in sensing Gram-positive pathogenic bacteria

We tested another pathogenic Gram-positive bacterium, *Staphylococcus aureus.* We observed that *sphe* null mutant flies showed the same susceptibility as *psh* mutant flies to infection compared to wild type flies (Figure 4A), but *drs* levels 24 hours upon immune challenge are as in wild type controls (Figure 4B). The wild type activation of Toll pathway could however be due to the PRR pathway. In order to confirm the involvement in the "danger" signal cascade, we used double mutant $sphe^{\Delta 104};grass^{hrd}$ flies in which the recognition cascade is blocked. *drs* levels are significantly decreased in $sphe^{\Delta 104};grass^{hrd}$ double mutants compared to the levels in wild type flies and in both *sphe* or $grass^{hrd}$ single mutant flies (it reaches 30% of wild type flies) (Figure 4B). These data demonstrate that Sphe is also involved in the sensing of virulence factors produced by *Staphylococcus aureus*.

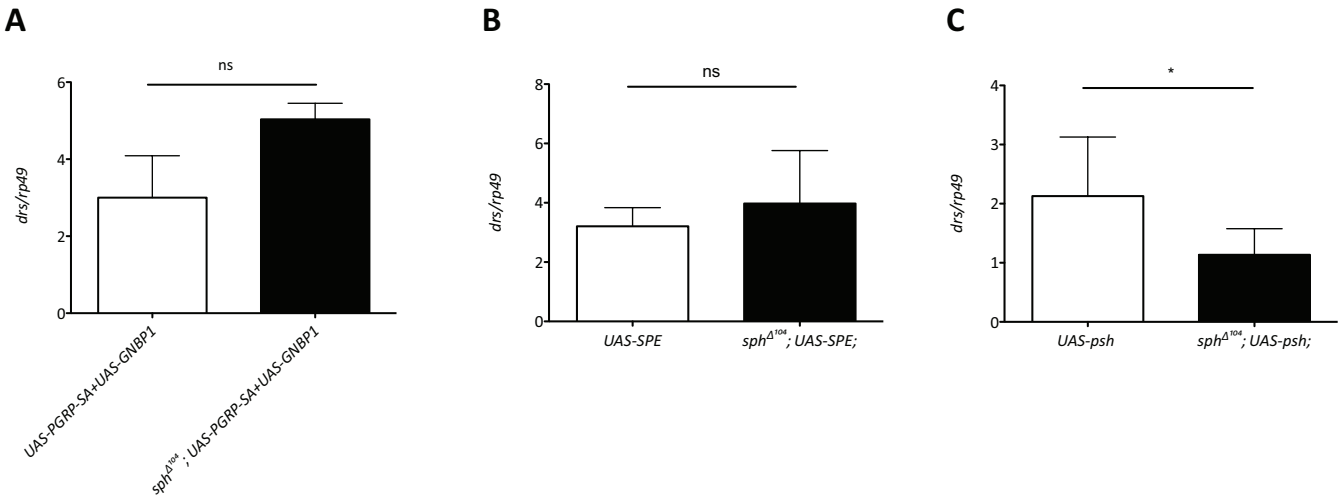


Figure 5 Spheroide is functioning in Psh pathway

drosomycin expression normalized to *rp49* after over expression, using the *yolk>Gal4* driver of both PGRP-SA and GNBP1 (**A**) or of SPE (**B**) or of Psh using the *Act5C>Gal4* driver (**C**) (*UAS-psh* vs. *sphe$^{\Delta 104}$; UAS-psh* p<0.0398).

Sphe is acting in the Persephone pathway

Over expression of serine proteases that lead to proteolytic cleavage of Spz constitutively activates the Toll pathway and induces *drs* expression in the absence of immune challenge. By over expressing the Toll pathway serine proteases in an *sphe* mutant background we assessed the position of Sphe in the cascades. As expected from the phenotype of *sphe* flies, Toll pathway activation after over expression of both PGPR-SA and GNBP1, or of SPE, is not blocked in *sphe* mutant background (Figure 5A and B). However, Toll pathway activation after Psh over expression is strongly reduced in *sphe* mutant background (Figure 5C).

These observations demonstrate that Sphe is acting downstream of (or at the same level as) Psh in the “danger” signal cascade.

DISCUSSION

A previous RNAi screen suggested that Sphe is required for Toll pathway activation. By analyzing the null mutant phenotype of this SPH, we confirm that Sphe is involved in the activation of Toll pathway. However, we show that Sphe is not required for different kind of infections as previously reported but only after the immune challenge with the pathogenic Gram-positive bacteria *Enterococcus faecalis* and *Staphylococcus aureus*. Furthermore, we demonstrate that Sphe is a component of the "danger" signal cascade acting downstream (or at the same level) of Psh, and is involved in the sensing of virulence factors (proteases) secreted by these pathogenic bacteria.

The Sphe serine protease homolog has a signal peptide and a trypsin-like protease fold. Within the catalytic triad, the active serine residue is mutated to a glycine residue, blocking the proteolytic activity. Since it has no amidase activity, Sphe cannot directly activate a downstream zymogen. Sphe might, however, itself be a target substrate for an activating protease. In this context, we note that the N-terminal of Sphe does not contain a CLIP-domain, which is conserved domain in many arthropod immune-related SPs and SPHs.

Serine protease homologs have been implicated in various physiological processes. In 1991, Hogg et al. reported that a mammalian serine protease homolog, protein Z (PZ), a vitamin K-dependent glycoprotein, binds to thrombin causing its conformational changes. When bound to PZ, thrombin is associated with phospholipid membrane vesicles. This membrane localization is important during coagulation and clotting as it partitions thrombin to the site of an injury (Hogg and

Stenflo, 1991). Later studies of protein crystal structure have demonstrated that PZ functions as a cofactor regulating proteolytic activity of Factor Xa (FXa) on phospholipid vesicles (Qureshi et al., 2014). This is achieved through interaction with FXa and PZ-dependent Protease Inhibitor (ZPI), which forms a serpin/protease complex with FXa. PZ has an NH_2-terminal domain-containing γ-carboxyglutamic acid (Gla) domain and two EGF-like domains. PZ and ZPI interact via the sites on the inactive catalytic domain, but also within the EGF2-like domain. Protein-protein interactions between PZ and FXa take place through Gla-domains. This interaction is necessary for the assembly of a protein complex on the phospholipid vesicle surface, which leads to the formation of an effective inhibitory complex containing PZ/FXa/ZPI.

One reported serine protease homolog in *Drosophila*, Masquerade, is necessary during embryonic development to promote and/or stabilize cell-matrix interactions. Masquerade has a CLIP domain in the NH_2-terminal region of the protein. One model for Masquerade function proposes that the CLIP domain mediates protease interactions by promoting cell-substrate adhesion. Another proposed model is based on the fact that Masquerade does not have amidase activity. As an inactive protease homolog, Masquerade might compete for a substrate with the active serine protease, indirectly stabilizing cell-substrate interactions (Murugasu-Oei et al., 1995).

Sphe has no conserved domains in its NH_2-terminal region. However we cannot exclude that it is interacting with other proteins, or forming a complex. The role of the CLIP-domain itself is unclear. It has been hypothesized that CLIP-domain

may be involved in mediating protein-protein interactions, but some SPHs lacking a CLIP domain also have roles in proteolytic cascades. In *Drosophila*, 16 out of the 67 SPHs have a complete, or partial CLIP-domain. Recent studies on SPs involved in the activation of Toll pathway during embryonic development reported that Gastrulation Defective (GD) forms a complex with Snake (Snk) and Easter (Ea) and that this association is required for the activation Ea by Snk. Surprisingly, this mediation of Snk activity is not dependent on the proteolytic activity of GD but still occurs in GD mutants that lack one of the active catalytic residues (Cho et al., 2012). This result establishes that a proteolytically inactive SP can function as a mediator to promote zymogen activation via another SP component of a proteolytic cascade. When GD is itself activated, its NH_2-terminal region interacts with sulfated proteins located in the ventral region of the perivitelline space. This ventral localization of GD acts to bring Ea and Snk together and promote the ventrally restricted processing of Ea.

The *Manduca sexta* SPH, SPH3, which lacks a CLIP-domain, is required in the immune response of the moth to infection with Gram-negative bacterium *Photorabdus luminescens*. SPH3 was initially identified as a target for the Repeats-in-toxin (RTX)-metalloprotease, protease A (PrtA), which is secreted by the bacterium. Upon infection, SPH3 is upregulated in both the fat body and hemocytes. RNAi-mediated knockdown of SPH3 increased the susceptibility of moths to infection by *Photorabdus luminescens* compared to untreated controls. Furthermore, the levels of antimicrobial peptides and prophenoloxidase (PPO) were reduced. Conversely, mRNA and protein levels of recognition molecules were not changed in these animals. This implied the existence of two signaling pathways, one that governs the

recognition and the other that governs the transcription of effector molecules. Based on their findings, the authors proposed the role for SPH3 in the signaling pathway that controls the effector molecules (Felföldi et al. 2011).

Sphe is involved in the sensing of proteases produced by Gram-positive pathogenic bacteria, but not proteases produced by the entomopathogenic fungus *Beauveria bassiana*. We can hypothesize that Sphe is recruited in a complex that mediates the activation of a target SP on infection with a virulent strain. Two further SPHs, *sphynx1* and *sphynx2*, were identified in the RNAi screen of Kambris et al., together with *sphe*, as putative components of the Toll pathway proteolytic cascade. Since Sphe functions upon infection with specific pathogens, it is likely that Sphynx1 and Sphynx2 might also be involved in the activation of the immune response against other pathogens.

Further investigation is necessary to elucidate the mechanisms by which Sphe is functioning and how it contributes to the sensing of these virulence factors. In addition, characterization of other *Drosophila* SPHs would give insight into their mechanisms of action in other proteolytic cascades.

MATERIALS AND METHODS

Fly strains

Stocks were raised on standard cornmeal-yeast agar medium at 25 degrees Celsius, 60% humidity. Flies carrying *UAS-RNAi* transgene against *sphe* (*P{KK112345}VIE-260B*) were obtained from Vienna Drosophila Resource Center (VDRC). Flies carrying Minos transposable element (*Mi{ET1}spheroide*^{MB11555}) were obtained from Bloomington Stock Center. Flies carrying Minos transposase (*P{hsILMiT}2.4*) were obtained from Bloomington Stock Center. Flies that were used as controls in experiments: *w*^{1118} (Bloomington Stock Center), *spz*^{rm7} were used as Toll-deficient mutant flies (Letsou et al., 1991), *psh*^{4} (Ligoxygakis et al., 2002b), *grass*^{hrd} (El Chamy et al., 2008). Flies with *UAS* constructs used in the study: *UAS-psh* (Ligoxygakis et al., 2002b), *UAS-SPE* (Jang et al., 2006b), *UAS-grass* (El Chamy et al., 2008), *UAS-PGRP-SA, UAS-GNBP1* (Gobert et al., 2003). *UAS-sphe* line was generated in this study: A Myc tag was added at the C-terminal of Spheroide protein using annealed primers IMU938 (5'-GATCCAGGGCGAGCAGAAGCTGATCTCCGAGGAGGACCTGTG-3') and IMU939 (5'-GATCCACAGGTCCTCCTCGGAGATCAGCTTCTGCTCGCCCTG-3') cloned in the BamHI site of *CG9675* (*sphe*) cDNA (clone LP05929 from DGRC). The EcoRI-XhoI fragment was inserted in pUAST (Brand and Perrimon, 1993). Flies carrying different Gal4 drivers (*Act5C*, *hsp70*, *yolk* and *da* promoters) were obtained from Bloomington Stock Center. Gal4 driven *RNAi* knockdown was enhanced by incubating three day-old flies for four days at 29 °C. Heat shock was performed as follows: 20 minutes at 37 °C, 20 minutes at 18 °C and 20 minutes at 37 °C, for three days. *Gal4* driven over

expressions were enhanced by incubating three day-old flies for two days at 29 °C (epistatic analysis) or four days at 29 °C (rescue experiment).

Microbial strains and infection

For septic injury (Reichhart et al., 2011) we used *Micrococcus luteus* (4698), *Enterococcus faecalis* strain OG1RF, protease-deficient strains of *Enterococcus faecalis* (TX5128, TX5243, TX5264) (Qin et al., 2000), *Staphylococcus aureus* (RN6390), *Enterococcus faecium* DO (TX0016) (obtained from B. Murray). Bacteria were grown in Tryptic Soy Broth (TSB) (*M. luteus*) at 30 °C or Brain-Heart infusion Broth (BHB) (*E. faecalis, S. aureus* and *E. faecium*) at 37 °C. Protease-deficient strains TX5128 and TX5243 were cultured with 2mg/ml kanamycin. Bacterial suspensions were prepared from exponential growth phase cultures and diluted to OD600 0.5 in PBS solution for immune challenge except for *M. luteus* where a pellet from overnight culture was used.

Natural infection with Beauveria bassiana was performed as described (Lemaitre et al., 1997).

At least three independent survival experiments were performed on a mix of 20-30 males and females six to eight days-old flies infected with *E. faecalis*, protease-deficient *E. faecalis* or *S. aureus*, or by natural infection with *B. bassiana* at 29 °C (Reichhart et al., 2011). . The survival data was plotted with Kaplan-Meier curves and for statistical analysis we used Log-rank (Mantel-Cox) test using the Prism software.

Q-RT-PCR analysis

Total RNA was extracted from a mix of 20-30 males and females six to eight day-old flies using TRI REAGENT® (Molecular Research Center). RNA was extracted from flies 24 hours after (or post-infection, PI) challenged by septic injury with *M. luteus*, *E. faecalis*, protease-deficient *E. faecalis*, *E. faecium* and *S. aureus*, or 48 hours (or post-infection, PI) after natural infection with *B. bassiana* at 29 °C, at least three times, independently. Reverse transcription was performed using iScript™ cDNA Synthesis Kit (BIO-RAD). iQ™ SYBR® Green Supermix (BIO-RAD) was used for Quantitative RT-PCR using CFX384™ Real-Time System and CFX Manager 3.0 (BIO-RAD) for data analysis. Student's t-test was used for statistical analysis using Prism software. *drs, sphe* and *rp49* mRNA levels were quantified using these primers: rp49FW (5'-GACGCTTCAAGGGACAGTATCTG-3'), rp49 RV (5'-AAACGCGGTTCTGCATGA-3'), Drom FW (5'-CGTGAGAACCTTTTCCAATATGAT-3'), Drom RV (5-'TCCCAGGACCACGAGCAT-3', Sph FW (5'-CATTTTGCCGCGTTTGAG-3'), Sph RV (5'-GCATCCGGACTACTATAATCTGAA-3').

PCR analysis of deletions

sphe$^{\Delta 11}$ (wild type control), *sphe*$^{\Delta 49}$ and *sphe*$^{\Delta 104}$ deletion were generated in this study. Genomic DNA was extracted from single flies using squishing buffer (Tris-HCl pH 8.2 10 mM, EDTA 1 mM, NaCl 25 mM) and proteinase K (200 µg/ml) 30 min at 37 ^{o}C and 3 min at 95 ^{o}C to inactivate PK. PCR was performed using *Taq* DNA Polymerase (Invitrogen™) with these primers: 682FW (5'-TATGTGGCTGGATGGGGTGAACTT-3'), 4012RV (5'-AATGGGCGGCGGTGACAA-3') (for *sphe*$^{\Delta 49}$), 2661RV (5'-TCACGGCCAGGTTGTTGTTCAGAT-3') (for *sphe*$^{\Delta 104}$), MinosFW

(5’-TCGAATTAATAGTGGTCACTTTTTTT-3’), MinosRV (5’-GTTCGAATTAATAGTGGTTGGGGC-3’), using Tm= 57 ^{o}C. PCR with 682FW and 4102RV showed 3331 bp fragment in wild type flies, and a 554 bp fragment in *sphe*$^{\Delta 49}$ deletion. PCR with 682FW and 2661RV showed 1980 bp fragment in wild type flies and a 725 bp fragment *sphe*$^{\Delta 104}$ deletion.

REFERENCES

Buchon, N., Poidevin, M., Kwon, H.-M., Guillou, A., Sottas, V., Lee, B.-L., and Lemaitre, B. (2009). A single modular serine protease integrates signals from pattern-recognition receptors upstream of the Drosophila Toll pathway. Proc. Natl. Acad. Sci. U. S. A. *106*, 12442–12447.

El Chamy, L., Leclerc, V., Caldelari, I., and Reichhart, J.-M. (2008). Sensing of "danger signals" and pathogen-associated molecular patterns defines binary signaling pathways "upstream" of Toll. Nat. Immunol. *9*, 1165–1170.

Cho, Y.S., Stevens, L.M., Sieverman, K.J., Nguyen, J., and Stein, D. (2012). A ventrally localized protease in the Drosophila egg controls embryo dorsoventral polarity. Curr. Biol. *22*, 1013–1018.

Dimopoulos, G., Richman, A., Muller, H.M., and Kafatos, F.C. (1997). Molecular immune responses of the mosquito Anopheles gambiae to bacteria and malariaÿparasites. Proc. Natl. Acad. Sci. *94*, 11508.

Felföldi, G., Eleftherianos, I., Ffrench-Constant, R.H., and Venekei, I. (2011a). A serine proteinase homologue, SPH-3, plays a central role in insect immunity. J. Immunol. *186*, 4828–4834.

Garsin, D.A., Sifri, C.D., Mylonakis, E., Qin, X., Singh, K. V, Murray, B.E., Calderwood, S.B., and Ausubel, F.M. (2001). A simple model host for identifying Gram-positive virulence factors. Proc. Natl. Acad. Sci. U. S. A. *98*, 10892–10897.

Gobert, V., Gottar, M., Matskevich, A. A, Rutschmann, S., Royet, J., Belvin, M., Hoffmann, J. A. , and Ferrandon, D. (2003). Dual activation of the Drosophila toll pathway by two pattern recognition receptors. Science *302*, 2126–2130.

Hogg, P.J., and Stenflo, J. (1991). Interaction of vitamin K-dependent protein Z with thrombin. Consequences for the amidolytic activity of thrombin and the interaction of thrombin with phospholipid vesicles. J. Biol. Chem. *266*, 10953–10958.

Janeway, C.A. (1989). Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harb Symp Quant Biol *54*, 1–13.

Jang, I.-H., Chosa, N., Kim, S.-H., Nam, H.-J., Lemaitre, B., Ochiai, M., Kambris, Z., Brun, S., Hashimoto, C., Ashida, M., et al. (2006). A Spätzle-processing enzyme required for toll signaling activation in Drosophila innate immunity. Dev. Cell *10*, 45–55.

Kambris, Z., Brun, S., Jang, I.-H., Nam, H.-J., Romeo, Y., Takahashi, K., Lee, W.-J., Ueda, R., and Lemaitre, B. (2006). Drosophila immunity: a large-scale in vivo RNAi

screen identifies five serine proteases required for Toll activation. Curr. Biol. *16*, 808–813.

Kawabata, S.I., Tokunaga, F., Kugi, Y., Motoyama, S., Miura, Y., Hirata, M., and Iwanaga, S. (1996). Limulus factor D, a 43-kDa protein isolated from horseshoe crab hemocytes, is a serine protease homologue with antimicrobial activity. FEBS Lett. *398*, 146–150.

Lemaitre, B., and Hoffmann, J.A. (2007). The host defense of Drosophila melanogaster. Annu. Rev. Immunol. *25*, 697–743.

Lemaitre, B., Reichhart, J.M., and Hoffmann, J.A. (1997). Drosophila host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. Proc. Natl. Acad. Sci. U. S. A. *94*, 14614–14619.

Letsou, A., Alexander, S., Orth, K., and Wasserman, S.A. (1991). Genetic and molecular characterization of tube, a Drosophila gene maternally required for embryonic dorsoventral polarity. Proc. Natl. Acad. Sci. U. S. A. *88*, 810–814.

Leulier, F., Parquet, C., Pili-Floury, S., Ryu, J.-H., Caroff, M., Lee, W.-J., Mengin-Lecreulx, D., and Lemaitre, B. (2003). The Drosophila immune system detects bacteria through specific peptidoglycan recognition. Nat. Immunol. *4*, 478–484.

Ligoxygakis, P., Pelte, N., Hoffmann, J. A. , and Reichhart, J.-M. (2002). Activation of Drosophila Toll during fungal infection by a blood serine protease. Science *297*, 114–116.

Matzinger, P. (1994). Tolerance, danger, and the extended family. Annu. Rev. Immunol. 991–1045.

Murugasu-Oei, B., Rodrigues, V., Yang, X., and Chia, W. (1995). Masquerade: a novel secreted serine protease-like molecule is required for somatic muscle attachment in the Drosophila embryo. Genes Dev. *9*, 139–154.

Povelones, M., Bhagavatula, L., Yassine, H., Tan, L.A., Upton, L.M., Osta, M. a, and Christophides, G.K. (2013). The CLIP-Domain Serine Protease Homolog SPCLIP1 Regulates Complement Recruitment to Microbial Surfaces in the Malaria Mosquito Anopheles gambiae. PLoS Pathog. *9*, e1003623.

Qin, X., Singh, K. V, Weinstock, G.M., and Murray, B.E. (2000). Effects of Enterococcus faecalis fsr Genes on Production of Gelatinase and a Serine Protease and Virulence Effects of Enterococcus faecalis fsr Genes on Production of Gelatinase and a Serine Protease and Virulence. Infect. Immun. *68*, 2579–2586.

Qin, X., Galloway-Peña, J.R., Sillanpaa, J., Roh, J.H., Nallapareddy, S.R., Chowdhury, S., Bourgogne, A., Choudhury, T., Muzny, D.M., Buhay, C.J., et al. (2012). Complete

genome sequence of Enterococcus faecium strain TX16 and comparative genomic analysis of Enterococcus faecium genomes. BMC Microbiol. *12*, 135.

Qureshi, S.H., Lu, Q., Manithody, C., Yang, L., and Rezaie, A.R. (2014). Characterization of the protein Z-dependent protease inhibitor interactive-sites of protein Z. Biochim. Biophys. Acta *1844*, 1631–1637.

Reichhart, J.-M., Gubb, D., and Leclerc, V. (2011). The Drosophila serpins: multiple functions in immunity and morphogenesis. (Elsevier Inc.).

Ross, J., Jiang, H., Kanost, M.R., and Wang, Y. (2003). Serine proteases and their homologs in the Drosophila melanogaster genome: an initial analysis of sequence conservation and phylogenetic relationships. Gene *304*, 117–131.

Sifri, C.D., Mylonakis, E., Singh, K. V, Qin, X., Garsin, D.A., Murray, B.E., Ausubel, F.M., and Calderwood, S.B. (2002). Virulence Effect of Enterococcus faecalis Protease Genes and the Quorum-Sensing Locus fsr in Caenorhabditis elegans and Mice Virulence Effect of Enterococcus faecalis Protease Genes and the Quorum-Sensing Locus fsr in Caenorhabditis elegans and Mice. Society *70*, 5647–5650.

Conclusions and perspectives

1. The serine protease between ModSP and Grass

There are 211 SP-related genes in the genome of *Drosophila melanogaster*, indicating the importance of their function. They are involved in important physiological processes, including digestion, development and the immune response (Ross et al., 2003).

The proposed model of the Toll pathway activation via PRR signaling, consisting of three previously described SPs, ModSP, Grass and SPE is fragmentary. This is based on comparison of the SPs involved in the proteolytic cascades that lead to activation of the Toll pathway during development and immune response in *Drosophila*, and immune response in *Tenebrio mollitor*, indicating that Grass is not the functional equivalent of Tm-SAE and Snake. Therefore, a search for a potential SP that could function between ModSP and Grass was based on bioinformatic analysis, the chemical knowledge and the specificity of the SPs involved. Genes encoding this potential SP are *CG3700*, *CG2056* (*spirit*) and *CG4927.*

Using a fly line with a P-element inactivating insertion in *CG3700* we demonstrated that serine protease encoded by *CG3700* is not involved in the activation of the Toll pathway after immune challenge with Gram-positive bacteria, (*Enterococcus faecalis* and *Micrococcus luteus*) and fungus (*Beauveria bassiana*). However, we cannot exclude involvement of this serine protease in the immune response against other pathogens.

So far, we have only preliminary results for serine protease encoding *CG2056* (*spirit*) demonstrating that it is not involved in the activation of the Toll pathway

after immune challenge with Gram-positive bacteria (*Enterococcus faecalis* and *Micrococcus luteus*) and fungus (*Beauveria bassiana*). As with *CG3700*, we cannot exclude involvement of this serine protease in the immune response against other pathogens.

Since the CRISPR system proved to be efficient in generating mutants, we are still pursuing the work to find a mutant for *CG4927*.

Kambris et al., performed an RNAi *in vivo* screen targeting different serine proteases and serine protease homologs in *Drosophila* genome. The results obtained suggested that five SPs and SPHs have a role in the activation of the Toll pathway. One of the SPs was Grass. However, the obtained RNAi phenotype was a hypomorphic one, which was proved later using the null allele of Grass, $grass^{herrade}$ (El Chamy et al., 2008). In addition to this, RNAi against *psh* did not give the phenotype that is observed in *psh* mutant flies (Vincent Leclerc, personal communication) indicating that RNAi-mediated knockdown of secreted proteases is note efficient enough to get a strong phenotype. It may be due to the low amount of protease required to activate the pathway.

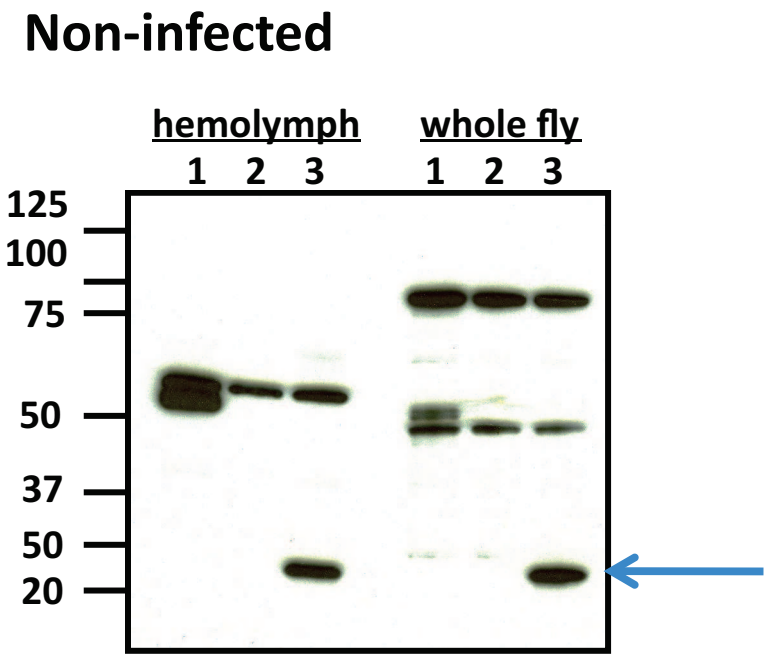
We pursued the work to verify the involvement of the other genes in the activation of the Toll pathway.

Using RNAi-mediated knockdown against *spirit*, mRNA levels were reduced to 20% of wild type levels implying the efficiency of the knockdown. These flies showed susceptibility to *E. faecalis*, *B. bassiana* and *C. albicans* infections. In addition, *drosomycin* expression was reduced when flies were challenged with *M. luteus* (as

well as only PGN from *M. luteus*), *B. bassiana* and *C. albicans* (Kambris et al., 2006). Using the same RNAi lines we didn't have an efficient knockdown of *spirit* and therefore we did not use it in our experiments. However, we used a null allele of *spirit* generated by CRISPR/Cas9 system. These mutant flies showed no reduction in *drosomycin* expression after the immune challenge with *E. faecalis* and *M. luteus*, and natural infection with *B. bassiana*. In addition, these mutant flies showed no susceptibility to *E. faecalis* and *B. bassiana* infections. Our results differ from those obtained in the RNAi screen which could imply an off target effect that the authors of Kambris paper obtained. However, we can still not exclude *spirit* to be involved against other pathogens, such as *C. albicans* as they demonstrated. Furthermore, we cannot exclude the involvement of these serine proteases in other proteolytic cascades in *Drosophila*. Further investigation and characterization of these serine proteases is necessary.

The work I have done aimed to characterize the various molecules that are involved in proteolytic cascades upstream of Toll. This will help to reconstitute these cascades *in vitro* and understand their organization and potential complex formation between SPs. In addition, characterization of other CLIP-SPs found in *Drosophila* is important to understand the mechanism of their function and their potential roles in different physiological processes.

A



1 – *UAS-nec myc tag*
2 – *CTRL*
3 – *UAS-sphe myc tag*

Act5cGal4
2 days @29°C
@myc

Figure 22 (to be continued on the next page)

2. The serine protease homolog Spheroide

The importance of SP-related genes in *Drosophila* has already been described. In the second part of my work I focused on the components of the danger signal cascade.

Based on a previous RNAi screen, the serine protease homolog *spheroide* was suggested to be required for the activation of Toll pathway. By analyzing the null mutant phenotype, we have demonstrated a new role for Spheroide in the activation of Toll pathway specifically after the immune challenge with pathogenic Gram-positive bacteria (*Enterococcus faecalis* and *Staphylococcus aureus*). Furthermore, we demonstrate that Spheroide is functioning as a component of the danger signal cascade, sensing virulence factors (proteases) produced by these pathogenic bacteria.

Spheroide is an SPH that has a signal peptide and trypsin-like catalytic domain with mutated serine residue rendering the protein inactive. Since it has no amidase activity it cannot actively function in the proteolytic cascade. Furthermore, there are no reported conserved cystein residues in the NH_2-terminal region of the protein that could form a CLIP-domain. Based on the knowledge of previously described SPHs we can propose a model in which Spheroide is functioning as a cofactor, forming a complex with other components of the cascade in order to mediate a proper immune response against these specific invading pathogens.

We undertook a biochemical approach where we over expressed Spheroide with the use of the *UAS-Gal4* system with the ubiquitous *Actin5C>Gal4* driver. This

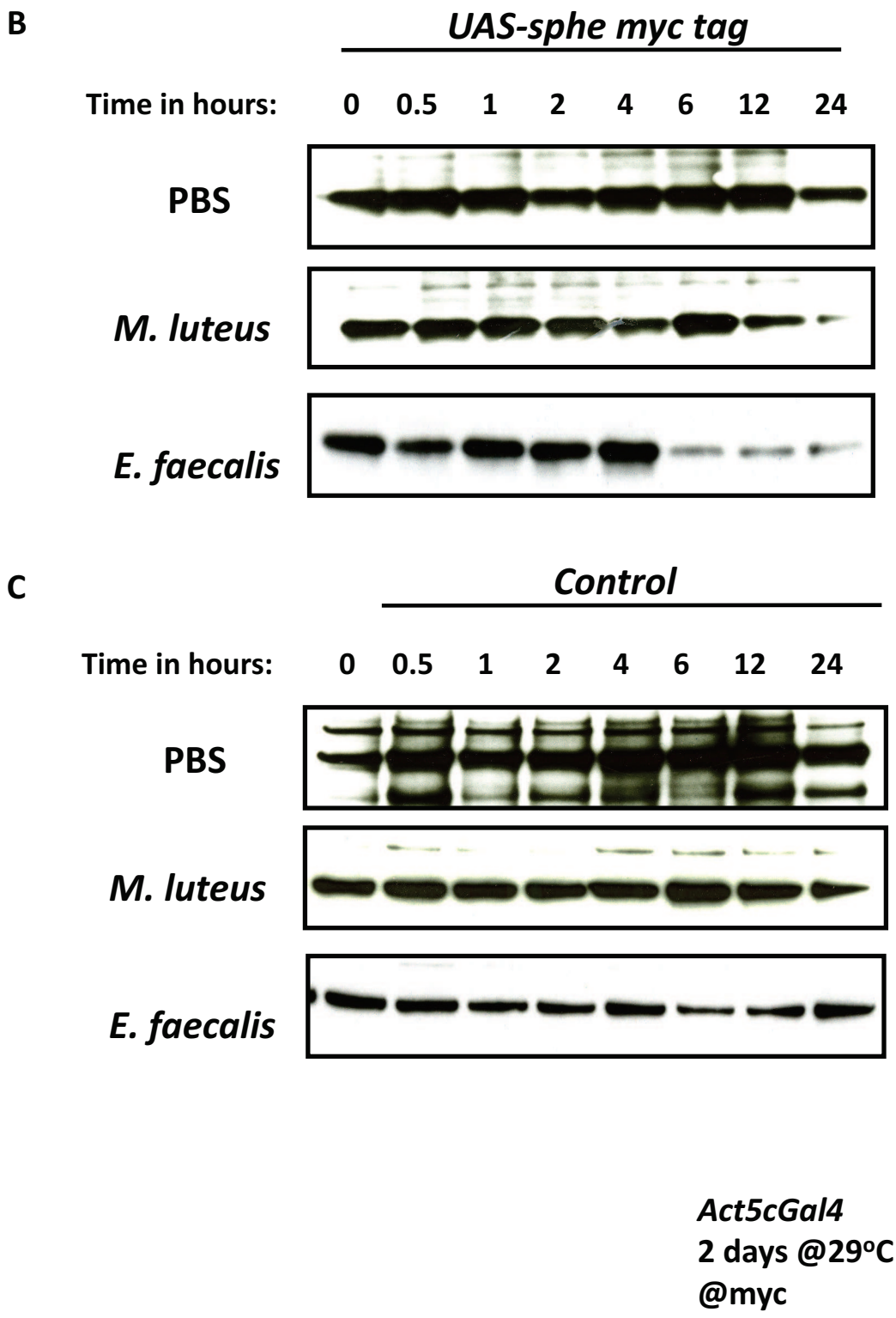


Figure 22

Over expression of Spheroide using *Act5C>Gal4* (for 2 days at 29 °C) with anti-myc antibody: localization in the hemolymph – the blu line indicates the signal corresponding to molecular weight of Spheroide (**A**); Time course experiment after clean injury with PBS, and immune challenge with *M.luteus* and *E. faecalis* (**B**). Unspecific bands observed during the time course experiment (used as a loading control) (**C**). Whole fly extract prepared with lysis buffer 30 mM Hepes pH 7.4, 150 mM NaCl, 2 mM MgAc, 1 % NP-40.

construct also has a Myc tag that we can use to localize the protein. Indeed, using an anti-Myc antibody we observed the localization of Spheroide in the hemolymph (Figure 22A). In addition, we performed time course experiments in order to see what is happening with the protein during a course of infection, whether we can observe a cleavage or the formation of a complex. Our preliminary data seems to indicate that the concentration of Spheroide in the hemolymph is decreased 6 hours after the immune challenge with *E. faecalis*, whereas it remains the same after the immune challenge with *M. luteus* (Figure 22B). Instead of a loading control, we used the observation that some unspecific bands detected by this antibody showed that the protein amount is not changed during this course of (Figure 22C).

Further investigation is necessary to elucidate the mechanisms by which Spheroide is functioning and how it is sensing these virulence factors. In addition, complementary approaches at biochemical level would provide additional answers.

Bibliography

Adams, M.D. (2000). The Genome Sequence of Drosophila melanogaster. Science (80). *287*, 2185–2195.

Aderem, A., and Ulevitch, R. (2000). Toll-like receptors in the induction of the innate immune response. Nature *406*.

Akira, S., Takeda, K., and Kaisho, T. (2001). Toll-like receptors: critical proteins linking innate and acquired immunity. Nat. Immunol. *2*, 675–680.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. Cell *124*, 783–801.

An, C., Zhang, M., Chu, Y., and Zhao, Z. (2013). Serine Protease MP2 Activates Prophenoloxidase in the Melanization Immune Response of Drosophila melanogaster. PLoS One *8*, e79533.

Anderson, K. V (2000). Toll signaling pathways in the innate immune response. Curr. Opin. Immunol. *12*, 13–19.

Avadhanula, V., Weasner, B.P., Hardy, G.G., Kumar, J.P., and Hardy, R.W. (2009). A Novel System for the Launch of Alphavirus RNA Synthesis Reveals a Role for the Imd Pathway in Arthropod Antiviral Response. PLoS Pathog *5*, e1000582.

Bangham, J., Jiggins, F., and Lemaitre, B. (2006). Insect Immunity: The Post-Genomic Era. Immunity *25*, 1–5.

Bassett, A.R., and Liu, J.-L. (2014). CRISPR/Cas9 and genome editing in Drosophila. J. Genet. Genomics *41*, 7–19.

Belvin, M.P., and Anderson, K. V (1996). A conserved signaling pathway: the Drosophila toll-dorsal pathway. Annu. Rev. Cell Dev. Biol. *12*, 393–416.

Bidla, G., Lindgren, M., Theopold, U., and Dushay, M.S. (2005). Hemolymph coagulation and phenoloxidase in Drosophila larvae. Dev. Comp. Immunol. *29*, 669–679.

Bidla, G., Dushay, M.S., and Theopold, U. (2007). Crystal cell rupture after injury in Drosophila requires the JNK pathway, small GTPases and the TNF homolog Eiger. J. Cell Sci. *120*, 1209–1215.

Binggeli, O., Neyen, C., Poidevin, M., and Lemaitre, B. (2014). Prophenoloxidase Activation Is Required for Survival to Microbial Infections in Drosophila. PLoS Pathog. *10*, e1004067.

Bischoff, V., Vignal, C., Boneca, I.G., Michel, T., Hoffmann, J. A., and Royet, J. (2004). Function of the drosophila pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. Nat. Immunol. *5*, 1175–1180.

Bischoff, V., Vignal, C., Duvic, B., Boneca, I.G., Hoffmann, J. A., and Royet, J. (2006). Downregulation of the Drosophila immune response by peptidoglycan- recognition proteins SC1 and SC2. PLoS Pathog. *2*, 0139–0147.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. Science *326*, 1509–1512.

Boman, H.G., Nilsson, I., and Rasmunson, B. (1972). Inducible Antibacterial Defence System in Drosophila. Nature *237*, 232–235.

Boutros, M., Agaisse, H., and Perrimon, N. (2002). Sequential activation of signaling pathways during innate immune responses in Drosophila. Dev. Cell *3*, 711–722.

Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development *118*, 401–415.

Buchon, N., Poidevin, M., Kwon, H.-M., Guillou, A., Sottas, V., Lee, B.-L., and Lemaitre, B. (2009). A single modular serine protease integrates signals from pattern-recognition receptors upstream of the Drosophila Toll pathway. Proc. Natl. Acad. Sci. U. S. A. *106*, 12442–12447.

Castillejo-López, C., and Häcker, U. (2005). The serine protease Sp7 is expressed in blood cells and regulates the melanization reaction in Drosophila. Biochem. Biophys. Res. Commun. *338*, 1075–1082.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. a, Somia, N. V, Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. *39*, e82.

El Chamy, L., Leclerc, V., Caldelari, I., and Reichhart, J.-M. (2008). Sensing of “danger signals” and pathogen-associated molecular patterns defines binary signaling pathways “upstream” of Toll. Nat. Immunol. *9*, 1165–1170.

Chasan, R., and Anderson, K. V (1989). The Role of easter, an Apparent Serine Protease , in Organizing the Dorsal-Ventral Pattern of the Drosophila Embryo. *56*, 391–400.

Chasan, R., Jin, Y., and Anderson, K. V (1992). Activation of the easter zymogen is regulated by five other genes to define dorsal-ventral polarity in the Drosophila embryo. Development *115*, 607–616.

Cho, Y.S., Stevens, L.M., Sieverman, K.J., Nguyen, J., and Stein, D. (2012). A ventrally localized protease in the Drosophila egg controls embryo dorsoventral polarity. Curr. Biol. *22*, 1013–1018.

Choe, K.-M., Werner, T., Stöven, S., Hultmark, D., and Anderson, K. V (2002). Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in Drosophila. Science *296*, 359–362.

Chrostek, E., Marialva, M.S.P., Esteves, S.S., Weinert, L. a, Martinez, J., Jiggins, F.M., and Teixeira, L. (2013). Wolbachia variants induce differential protection to viruses in Drosophila melanogaster: a phenotypic and phylogenomic analysis. PLoS Genet. *9*, e1003896.

Corbo, J.C., and Levine, M. (1996). Characterization of an immunodeficiency mutant in Drosophila. Mech. Dev. *55*, 211–220.

Costa, A., Jan, E., Sarnow, P., and Schneider, D. (2009). The Imd Pathway Is Involved in Antiviral Immune Responses in Drosophila. PLoS One *4*, e7436.

DeLotto, Y., and DeLotto, R. (1998). Proteolytic processing of the Drosophila Spatzle protein by Easter generates a dimeric NGF-like molecule with ventralising activity. Mech. Dev. *72*, 141–148.

Deng, D., Yan, C., Pan, X., Mahfouz, M., and Wang, J. (2012). Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. Science (80-.). *335*, 720–723.

Dimopoulos, G., Richman, A., Muller, H.M., and Kafatos, F.C. (1997). Molecular immune responses of the mosquito Anopheles gambiae to bacteria and malaria parasites. Proc. Natl. Acad. Sci. *94*, 11508.

Dissing, M., Giordano, H., and DeLotto, R. (2001). Autoproteolysis and feedback in a protease cascade directing Drosophila dorsal-ventral cell fate. EMBO J. *20*, 2387–2393.

Dushay, M.S., and Eldon, E.D. (1998). Drosophila immune responses as models for human immunity. Am. J. Hum. Genet. *62*, 10–14.

Dushay, M.S., Asling, B., and Hultmark, D. (1996). Origins of immunity: Relish, a compound Rel-like gene in the antibacterial defense of Drosophila. Proc. Natl. Acad. Sci. U. S. A. *93*, 10343–10347.

Engström, Y., Kadalayil, L., Sun, S.C., Samakovlis, C., Hultmark, D., and Faye, I. (1993). kappa B-like motifs regulate the induction of immune genes in Drosophila. J. Mol. Biol. *232*, 327–333.

Felföldi, G., Eleftherianos, I., Ffrench-Constant, R.H., and Venekei, I. (2011). A serine proteinase homologue, SPH-3, plays a central role in insect immunity. J. Immunol. *186*, 4828–4834.

Ferrandon, D., Jung, a C., Criqui, M., Lemaitre, B., Uttenweiler-Joseph, S., Michaut, L., Reichhart, J., and Hoffmann, J. A. (1998). A drosomycin-GFP reporter transgene reveals a local immune response in Drosophila that is not dependent on the Toll pathway. EMBO J. *17*, 1217–1227.

Ferrandon, D., Imler, J.-L., and Hoffmann, J. A. (2004). Sensing infection in Drosophila: Toll and beyond. Semin. Immunol. *16*, 43–53.

Filipe, S.R., Tomasz, A., and Ligoxygakis, P. (2005). Requirements of peptidoglycan structure that allow detection by the Drosophila Toll pathway. EMBO Rep. *6*, 327–333.

Franc, N.C., Dimarcq, J.L., Lagueux, M., Hoffmann, J., and Ezekowitz, R. a (1996). Croquemort, a novel Drosophila hemocyte/macrophage receptor that recognizes apoptotic cells. Immunity *4*, 431–443.

Geißler, R., Scholze, H., Hahn, S., Streubel, J., Bonas, U., Behrens, S.-E., and Boch, J. (2011). Transcriptional activators of human genes with programmable DNA-specificity. PLoS One *6*, e19509.

Georgel, P., Naitza, S., Kappler, C., Ferrandon, D., Zachary, D., Swimmer, C., Kopczynski, C., Duyk, G., Reichhart, J.-M., and Hoffmann, J. A. (2001). Drosophila Immune Deficiency (IMD) Is a Death Domain Protein that Activates Antibacterial Defense and Can Promote Apoptosis. Dev. Cell *1*, 503–514.

Gobert, V., Gottar, M., Matskevich, A. a, Rutschmann, S., Royet, J., Belvin, M., Hoffmann, J. A., and Ferrandon, D. (2003). Dual activation of the Drosophila toll pathway by two pattern recognition receptors. Science *302*, 2126–2130.

Golic, K.G. (1994). Local Transposition of P Elements in Drosophila melanogaster and Recombination Between Duplicated Elements Using a Site-Specific Recombinase. *563*, 551–563.

Gottar, M., Gobert, V., Michel, T., Belvin, M., Duyk, G., Hoffmann, J. A., Ferrandon, D., and Royet, J. (2002). The Drosophila immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein. Nature *416*, 640–644.

Gottar, M., Gobert, V., Matskevich, A. a, Reichhart, J.-M., Wang, C., Butt, T.M., Belvin, M., Hoffmann, J. A., and Ferrandon, D. (2006). Dual detection of fungal infections in Drosophila via recognition of glucans and sensing of virulence factors. Cell *127*, 1425–1437.

Gubb, D., Sanz-Parra, A., Barcena, L., Troxler, L., and Fullaondo, A. (2010). Protease inhibitors and proteolytic signalling cascades in insects. Biochimie *92*, 1749–1759.

De Gregorio, E., Spellman, P.T., Rubin, G.M., and Lemaitre, B. (2001). Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc. Natl. Acad. Sci. U. S. A. *98*, 12590–12595.

De Gregorio, E., Han, S.-J., Lee, W.-J., Baek, M.-J., Osaki, T., Kawabata, S.-I., Lee, B.-L., Iwanaga, S., Lemaitre, B., and Brey, P.T. (2002). An immune-responsive Serpin regulates the melanization cascade in Drosophila. Dev. Cell *3*, 581–592.

Ha, E.M., Oh, C., Bae, Y., and Lee, W. (2005a). A direct role for dual oxidase in Drosophila gut immunity. Science (80-.). *847*.

Ha, E.M., Oh, C.T., Ryu, J.H., Bae, Y.S., Kang, S.W., Jang, I. hwan, Brey, P.T., and Lee, W.J. (2005b). An antioxidant system required for host protection against gut infection in Drosophila. Dev. Cell *8*, 125–132.

Hashimoto, C., Hudson, K.L., and Anderson, K. V (1988). The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell *52*, 269–279.

Hashimoto, C., Kim, D.R., Weiss, L. a, Miller, J.W., and Morisato, D. (2003). Spatial regulation of developmental signaling by a serpin. Dev. Cell *5*, 945–950.

Haskel-Ittah, M., Ben-Zvi, D., Branski-Arieli, M., Schejter, E.D., Shilo, B.-Z., and Barkai, N. (2012). Self-organized shuttling: generating sharp dorsoventral polarity in the early Drosophila embryo. Cell *150*, 1016–1028.

Hedstrom, L. (2002). Serine protease mechanism and specificity. Chem. Rev. *102*, 4501–4524.

Hoffmann, J. A., and Reichhart, J.-M. (2002). Drosophila innate immunity: an evolutionary perspective. Nat. Immunol. *3*, 121–126.

Hogg, P.J., and Stenflo, J. (1991). Interaction of vitamin K-dependent protein Z with thrombin. Consequences for the amidolytic activity of thrombin and the interaction of thrombin with phospholipid vesicles. J. Biol. Chem. *266*, 10953–10958.

Hu, X., Yagi, Y., Tanji, T., Zhou, S., and Ip, Y.T. (2004). Multimerization and interaction of Toll and Spätzle in Drosophila. Proc. Natl. Acad. Sci. U. S. A. *101*, 9369–9374.

Huang, T.S., Wang, H., Lee, S.Y., Johansson, M.W., Söderhäll, K., and Cerenius, L. (2000). A cell adhesion protein from the crayfish Pacifastacus leniusculus, a serine proteinase homologue similar to Drosophila masquerade. J. Biol. Chem. *275*, 9996–10001.

Huntington, J.A., Read, R.J., and Carrell, R.W. (2000). Structure of a serpin-protease complex shows inhibition by deformation. Nature *407*, 923–926.

Imler, J.-L., and Bulet, P. (2005). Antimicrobial peptides in Drosophila: structures, activities and gene regulation. Chem. Immunol. Allergy *86*, 1–21.

Ip, Y.T., Reach, M., Engstrom, Y., Kadalayil, L., Cai, H., Gonzalez-Crespo, S., Tatei, K., and Levine, M. (1993). Dif, a dorsal-related gene that mediates an immune response in Drosophila. Cell *75*, 753–763.

Irving, P., Troxler, L., Heuer, T.S., Belvin, M., Kopczynski, C., Reichhart, J.-M., Hoffmann, J. A., and Hetru, C. (2001). A genome-wide analysis of immune responses in Drosophila. Proc. Natl. Acad. Sci. U. S. A. *98*, 15119–15124.

Irving, P., Ubeda, J.M., Doucet, D., Troxler, L., Lagueux, M., Zachary, D., Hoffmann, J. A., Hetru, C., and Meister, M. (2005). New insights into Drosophila larval haemocyte functions through genome-wide analysis. Cell. Microbiol. *7*, 335–350.

Janeway, C.A. (1989). Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harb Symp Quant Biol *54*, 1–13.

Jang, I.-H., Chosa, N., Kim, S.-H., Nam, H.-J., Lemaitre, B., Ochiai, M., Kambris, Z., Brun, S., Hashimoto, C., Ashida, M., et al. (2006). A Spätzle-processing enzyme required for toll signaling activation in Drosophila innate immunity. Dev. Cell *10*, 45–55.

Jang, I.-H., Nam, H.-J., and Lee, W.-J. (2008). CLIP-domain serine proteases in Drosophila innate immunity. BMB Rep. *41*, 102–107.

Jiang, H., and Kanost, M.R. (2000). The clip-domain family of serine proteinases in arthropods. Insect Biochem. Mol. Biol. *30*, 95–105.

Jung, S.-H., Evans, C.J., Uemura, C., and Banerjee, U. (2005). The Drosophila lymph gland as a developmental model of hematopoiesis. Development *132*, 2521–2533.

Kambris, Z., Brun, S., Jang, I.-H., Nam, H.-J., Romeo, Y., Takahashi, K., Lee, W.-J., Ueda, R., and Lemaitre, B. (2006). Drosophila immunity: a large-scale in vivo RNAi screen identifies five serine proteases required for Toll activation. Curr. Biol. *16*, 808–813.

Kaneko, T., Goldman, W.E., Mellroth, P., Steiner, H., Fukase, K., Kusumoto, S., Harley, W., Fox, A., Golenbock, D., and Silverman, N. (2004). Monomeric and polymeric gram-negative peptidoglycan but not purified LPS stimulate the Drosophila IMD pathway. Immunity *20*, 637–649.

Kaneko, T., Yano, T., Aggarwal, K., Lim, J.-H., Ueda, K., Oshima, Y., Peach, C., Erturk-Hasdemir, D., Goldman, W.E., Oh, B.-H., et al. (2006). PGRP-LC and PGRP-LE have essential yet distinct functions in the drosophila immune response to monomeric DAP-type peptidoglycan. Nat. Immunol. *7*, 715–723.

Kang, D., Liu, G., Lundström, A., Gelius, E., and Steiner, H. (1998). A peptidoglycan recognition protein in innate immunity conserved from insects to humans. Proc. Natl. Acad. Sci. U. S. A. *95*, 10078–10082.

Kappler, C., Meister, M., Lagueux, M., Gateff, E., Hoffmann, J. A., and Reichhart, J.-M. (1993). Insect immunity. Two 17 bp repeats nesting a kappa B-related sequence confer inducibility to the diptericin gene and bind a polypeptide in bacteria-challenged Drosophila. EMBO J. *12*, 1561–1568.

Kawabata, S.I., Tokunaga, F., Kugi, Y., Motoyama, S., Miura, Y., Hirata, M., and Iwanaga, S. (1996). Limulus factor D, a 43-kDa protein isolated from horseshoe crab hemocytes, is a serine protease homologue with antimicrobial activity. FEBS Lett. *398*, 146–150.

Kellenberger, C., Leone, P., Coquet, L., Jouenne, T., Reichhart, J.-M., and Roussel, A. (2011). Structure-function analysis of grass clip serine protease involved in Drosophila Toll pathway activation. J. Biol. Chem. *286*, 12300–12307.

Kemp, C., Mueller, S., Goto, A., Barbier, V., Paro, S., Bonnay, F., Dostert, C., Troxler, L., Hetru, C., Meignin, C., et al. (2013). Broad RNA interference-mediated antiviral immunity and virus-specific inducible responses in Drosophila. J. Immunol. *190*, 650–658.

Kim, C.-H., Kim, S.-J., Kan, H., Kwon, H.-M., Roh, K.-B., Jiang, R., Yang, Y., Park, J.-W., Lee, H.-H., Ha, N.-C., et al. (2008). A three-step proteolytic cascade mediates the activation of the peptidoglycan-induced toll pathway in an insect. J. Biol. Chem. *283*, 7599–7607.

Kim, M.S., Baek, M.J., Lee, M.H., Park, J.W., Lee, S.Y., Soderhall, K., and Lee, B.L. (2002). A new easter-type serine protease cleaves a masquerade-like protein during prophenoloxidase activation in Holotrichia diomphalia larvae. J. Biol. Chem. *277*, 39999–40004.

Kocks, C., Cho, J.H., Nehme, N., Ulvila, J., Pearson, A.M., Meister, M., Strom, C., Conto, S.L., Hetru, C., Stuart, L.M., et al. (2005). Eater, a transmembrane protein mediating phagocytosis of bacterial pathogens in Drosophila. Cell *123*, 335–346.

Lamiable, O., and Imler, J.-L. (2014). Induced antiviral innate immunity in Drosophila. Curr. Opin. Microbiol. *20C*, 62–68.

Leclerc, V., Pelte, N., El Chamy, L., Martinelli, C., Ligoxygakis, P., Hoffmann, J. A., and Reichhart, J.-M. (2006). Prophenoloxidase activation is not required for survival to microbial infections in Drosophila. EMBO Rep. *7*, 231–235.

Lee, S.Y., and Soderhall, K. (2001). Characterization of a Pattern Recognition Protein, a Masquerade-Like Protein, in the Freshwater Crayfish Pacifastacus leniusculus. J. Immunol. *166*, 7319–7326.

Lee, S.Y., Cho, M.Y., Hyun, J.H., Lee, K.M., Homma, K.I., Natori, S., Kawabata, S.I., Iwanaga, S., and Lee, B.L. (1998). Molecular cloning of cDNA for pro-phenol-oxidase-activating factor i, a serine protease is induced by lipopolysaccharide or 1,3-β-glucan in coleopteran insect, Holotrichia diomphalia larvae. Eur. J. Biochem. *257*, 615–621.

Lee, W.J., Lee, J.D., Kravchenko, V. V, Ulevitch, R.J., and Brey, P.T. (1996). Purification and molecular cloning of an inducible gram-negative bacteria-binding protein from the silkworm, Bombyx mori. Proc. Natl. Acad. Sci. U. S. A. *93*, 7888–7893.

Lemaitre, B., Meister, M., Govind, S., Georgel, P., Steward, R., Reichhart, J.-M., and Hoffmann, J. A. (1995a). Functional analysis and regulation of nuclear import of dorsal during the immune response in Drosophila. EMBO J. *14*, 536–545.

Lemaitre, B., Kromer-Metzger, E., Michaut, L., Nicolas, E., Meister, M., Georgel, P., Reichhart, J.-M., and Hoffmann, J. A. (1995b). A recessive mutation, immune deficiency (imd), defines two distinct control pathways in the Drosophila host defense. Proc. Natl. Acad. Sci. U. S. A. *92*, 9465–9469.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.-M., and Hoffmann, J. A. (1996). The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell *86*, 973–983.

LeMosy, E.K., Kemler, D., and Hashimoto, C. (1998). Role of Nudel protease activation in triggering dorsoventral polarization of the Drosophila embryo. Development *125*, 4045–4053.

LeMosy, E.K., Tan, Y.Q., and Hashimoto, C. (2001). Activation of a protease cascade involved in patterning the Drosophila embryo. Proc. Natl. Acad. Sci. U. S. A. *98*, 5055–5060.

Lesch, C., Goto, A., and Lindgren, M. (2007). A role for Hemolectin in coagulation and immunity in Drosophila melanogaster. Dev. *31*, 1255–1263.

Leulier, F., Parquet, C., Pili-Floury, S., Ryu, J.-H., Caroff, M., Lee, W.-J., Mengin-Lecreulx, D., and Lemaitre, B. (2003). The Drosophila immune system detects bacteria through specific peptidoglycan recognition. Nat. Immunol. *4*, 478–484.

Levashina, E. A, Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J. A., and Reichhart, J.-M. (1999). Constitutive activation of toll-mediated antifungal defense in serpin-deficient Drosophila. Science *285*, 1917–1919.

Levashina, E.A., Ohresser, S., Lemaitre, B., and Imler, J. -L. (1998). Two distinct pathways can control expression of the gene encoding the Drosophila antimicrobial peptide metchnikowin. J. Mol. Biol. *278*, 515–527.

Levy, F., Bulet, P., and Ehret-Sabatier, L. (2004a). Proteomic analysis of the systemic immune response of Drosophila. Mol. Cell. Proteomics *3*, 156–166.

Levy, F., Rabel, D., Charlet, M., Bulet, P., Hoffmann, J. A., and Ehret-Sabatier, L. (2004b). Peptidomic and proteomic analyses of the systemic immune response of Drosophila. Biochimie *86*, 607–616.

Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P., and Yang, B. (2011). TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res. *39*, 359–372.

Li, Q., and Verma, I.M. (2002). NF-kappaB regulation in the immune system. Nat. Rev. Immunol. *2*, 725–734.

Ligoxygakis, P., Pelte, N., Ji, C., Leclerc, V., Duvic, B., Belvin, M., Jiang, H., Hoffmann, J. A., and Reichhart, J.-M. (2002a). A serpin mutant links Toll activation to melanization in the host defence of Drosophila. EMBO J. *21*, 6330–6337.

Ligoxygakis, P., Pelte, N., Hoffmann, J. A., and Reichhart, J.-M. (2002b). Activation of Drosophila Toll during fungal infection by a blood serine protease. Science *297*, 114–116.

Ligoxygakis, P., Roth, S., and Reichhart, J.-M. (2003). A Serpin Regulates Dorsal-Ventral Axis Formation in the Drosophila Embryo. Curr. Biol. *13*, 2097–2102.

Liu, B., Behura, S.K., Clem, R.J., Schneemann, A., Becnel, J., Severson, D.W., and Zhou, L. (2013). P53-Mediated Rapid Induction of Apoptosis Conveys Resistance to Viral Infection in Drosophila melanogaster. PLoS Pathog *9*, e1003137.

Liu, J., Li, C., Yu, Z., Huang, P., Wu, H., Wei, C., Zhu, N., Shen, Y., Chen, Y., Zhang, B., et al. (2012). Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy. J. Genet. Genomics *39*, 209–215.

Mak, A.N., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. *335*, 716–719.
Matzinger, P. (1994). Tolerance, danger, and the extended family. Annu. Rev. Immunol. 991–1045.

Matzinger, P. (2002). An innate sense of danger. Ann. N. Y. Acad. Sci. *961*, 341–342.
Medzhitov, R., and Janeway, C. (2000). Innate immune recognition: mechanisms and pathways. Immunol. Rev. *173*, 89–97.

Medzhitov, R., Preston-Hurlburt, P., and Janeway, C.A. (1997). A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. Nature *388*, 394–397.

Meister, M. (2004). Blood cells of Drosophila: Cell lineages and role in host defence. Curr. Opin. Immunol. *16*, 10–15.

Meister, M., Braun, A., Kappler, C., Reichhart, J.-M., and Hoffmann, J. A. (1994). Insect immunity. A transgenic analysis in Drosophila defines several functional domains in the diptericin promoter. EMBO J. *13*, 5958–5966.

Merkling, S.H., and van Rij, R.P. (2013). Beyond RNAi: antiviral defense strategies in Drosophila and mosquito. J. Insect Physiol. *59*, 159–170.

Michel, T., Reichhart, J.-M., Hoffmann, J. A., and Royet, J. (2001). Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature *414*, 756–759.

Ming, M., Obata, F., Kuranaga, E., and Miura, M. (2014). Persephone/Spatzle Pathogen Sensors Mediate the Activation of Toll Receptor Signaling in Response to Endogenous Danger Signals in Apoptosis-deficient Drosophila. J. Biol. Chem. *289*, 7558–7568.

Mizuguchi, K., Parker, J.S., Blundell, T.L., and Gay, N.J. (1998). Getting knotted: A model for the structure and activation of Spatzle. Trends Biochem. Sci. *23*, 239–242.

Morbitzer, R., Elsaesser, J., Hausner, J., and Lahaye, T. (2011). Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res. *39*, 5790–5799.

Moy, R.H., Gold, B., Molleston, J.M., Schad, V., Yanger, K., Salzano, M.-V., Yagi, Y., Fitzgerald, K.A., Stanger, B.Z., Soldan, S.S., et al. (2014). Antiviral Autophagy Restricts Rift Valley Fever Virus Infection and Is Conserved from Flies to Mammals. Immunity *40*, 51–65.

Murugasu-Oei, B., Rodrigues, V., Yang, X., and Chia, W. (1995). Masquerade: a novel secreted serine protease-like molecule is required for somatic muscle attachment in the Drosophila embryo. Genes Dev. *9*, 139–154.

Mushegian, A., and Medzhitov, R. (2001). Evolutionary perspective on innate immune recognition. J. Cell Biol. *155*, 705–710.

Muta, T., and Iwanaga, S. (1996). The role of hemolymph coagulation in innate immunity. Curr. Opin. Immunol.

Muta, T., Hashimoto, R., Miyata, T., Nishimura, H., Toh, Y., and Iwanaga, S. (1990). Proclotting enzyme from horseshoe crab hemocytes. cDNA cloning, disulfide locations, and subcellular localization. J. Biol. Chem. *265*, 22426–22433.

Nappi, A.J., Vass, E., Frey, F., and Carton, Y. (1995). Superoxide anion generation in Drosophila during melanotic encapsulation of parasites. Eur. J. Cell Biol. *68*, 450–456.

Nappi, A.J., Vass, E., Frey, F., and Carton, Y. (2000). Nitric oxide involvement in Drosophila immunity. Nitric Oxide *4*, 423–430.

Nayak, A., Tassetto, M., Kunitomi, M., and Andino, R. (2013). RNA Interference-Mediated Intrinsic Antiviral Immunity in Invertebrates. In Intrinsic Immunity SE - 7, B.R. Cullen, ed. (Springer Berlin Heidelberg), pp. 183–200.

Neyen, C., Bretscher, A.J., Binggeli, O., and Lemaitre, B. (2014). Methods to study Drosophila immunity. Methods.

Page, M.J., and Di Cera, E. (2008). Serine peptidases: classification, structure and function. Cell. Mol. Life Sci. *65*, 1220–1236.

Pearson, A.M., Baksa, K., Rämet, M., Protas, M., Mckee, M., Brown, D., and Ezekowitz, R.A.B. (2003). Identification of cytoskeletal regulatory proteins required for efficient phagocytosis in Drosophila. *5*, 815–824.

Perona, J.J., and Craik, C.S. (1995). Structural basis of substrate specificity in the serine proteases. Protein Sci. *4*, 337–360.

Poltorak, A., He, X., Smirnova, I., Liu, M.-Y., Huffel, C. Van, Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C., et al. (1998). Defective LPS Signaling in C3H/HeJ and C57BL/10ScCr Mice: Mutations in Tlr4 Gene. Sci. *282* , 2085–2088.

Povelones, M., Bhagavatula, L., Yassine, H., Tan, L.A., Upton, L.M., Osta, M. a, and Christophides, G.K. (2013). The CLIP-Domain Serine Protease Homolog SPCLIP1 Regulates Complement Recruitment to Microbial Surfaces in the Malaria Mosquito Anopheles gambiae. PLoS Pathog. *9*, e1003623.

Qureshi, S.H., Lu, Q., Manithody, C., Yang, L., and Rezaie, A.R. (2014). Characterization of the protein Z-dependent protease inhibitor interactive-sites of protein Z. Biochim. Biophys. Acta *1844*, 1631–1637.

Rämet, M., Manfruelli, P., Pearson, A., Mathey-Prevot, B., and Ezekowitz, R.A.B. (2002). Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for E. coli. Nature *416*, 644–648.

Rancès, E., Johnson, T.K., Popovici, J., Iturbe-Ormaetxe, I., Zakir, T., Warr, C.G., and O'Neill, S.L. (2013). The toll and Imd pathways are not required for wolbachia-mediated dengue virus interference. J. Virol. *87*, 11945–11949.

Rawlings, N.D., and Barrett, A.J. (1999). MEROPS: The peptidase database. Nucleic Acids Res. *27*, 325–331.

Rawlings, N.D., Tolle, D.P., and Barrett, A.J. (2004). Evolutionary families of peptidase inhibitors. Biochem. J. *378*, 705–716.

Reichhart, J.-M., Georgel, P., Meister, M., Lemaitre, B., Kappler, C., and Hoffmann, J. A. (1993). Expression and nuclear translocation of the rel/NF-kappa B-related morphogen dorsal during the immune response of Drosophila. C. R. Acad. Sci. III. *316*, 1218–1224.

Ren, X., Sun, J., Housden, B.E., Hu, Y., Roesel, C., Lin, S., Liu, L.-P., Yang, Z., Mao, D., Sun, L., et al. (2013b). Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. Proc. Natl. Acad. Sci. U. S. A. *10*, 19012–19017.

Ross, J., Jiang, H., Kanost, M.R., and Wang, Y. (2003). Serine proteases and their homologs in the Drosophila melanogaster genome: an initial analysis of sequence conservation and phylogenetic relationships. Gene *304*, 117–131.

Rubin, G.M. (2000). Comparative Genomics of the Eukaryotes. Science (80-.). *287*, 2204–2215.

Russo, J., Dupas, S., Frey, F., Carton, Y., and Brehelin, M. (1996). Insect immunity: early events in the encapsulation process of parasitoid (Leptopilina boulardi) eggs in resistant and susceptible strains of Drosophila. Parasitology *112 (Pt 1*, 135–142.

Scherfer, C., Karlsson, C., Loseva, O., Bidla, G., Goto, A., Havemann, J., Dushay, M.S., and Theopold, U. (2004). Isolation and Characterization of Hemolymph Clotting Factors in Drosophila melanogaster by a Pullout Method. *14*, 625–629.

Scherfer, C., Qazi, M.R., Takahashi, K., Ueda, R., Dushay, M.S., Theopold, U., and Lemaitre, B. (2006). The Toll immune-regulated Drosophila protein Fondue is involved in hemolymph clotting and puparium formation. Dev. Biol. *295*, 156–163.

Scherfer, C., Tang, H., Kambris, Z., Lhocine, N., Hashimoto, C., and Lemaitre, B. (2008). Drosophila Serpin-28D regulates hemolymph phenoloxidase activity and adult pigmentation. Dev. Biol. *323*, 189–196.

Shelly, S., Lukinova, N., Bambina, S., Berman, A., and Cherry, S. (2014). Autophagy Is an Essential Component of Drosophila Immunity against Vesicular Stomatitis Virus. Immunity *30*, 588–598.

Smidler, A.L., Terenzi, O., Soichot, J., Levashina, E. a, and Marois, E. (2013). Targeted mutagenesis in the malaria mosquito using TALE nucleases. PLoS One *8*, e74511.

Soderhall, K., and Cerenius, L. (1998). Role of the prophenoloxidase-activating system in invertebrate immunity. Curr. Opin. Immunol. *10*, 23–28.

Stein, D., and Nusslein-Volhard, C. (1992). Multiple extracellular activities in Drosophila egg perivitelline fluid are required for establishment of embryonic dorsal-ventral polarity. Cell *68*, 429–440.

Stein, D., Cho, Y., Zhang, Z., and Stevens, L. (2008). No requirement for localized Nudel protein expression in Drosophila embryonic axis determination. Fly (Austin). *2*, 220–228.

Stenbak, C.R., Ryu, J.-H., Leulier, F., Pili-Floury, S., Parquet, C., Hervé, M., Chaput, C., Boneca, I.G., Lee, W.-J., Lemaitre, B., et al. (2004). Peptidoglycan molecular requirements allowing detection by the Drosophila immune deficiency pathway. J. Immunol. *173*, 7339–7348.

Steward, R. (1987). Dorsal, an embryonic polarity gene in Drosophila, is homologous to the vertebrate proto-oncogene, c-rel. Science *238*, 692–694.

Takehana, A., Katsuyama, T., Yano, T., Oshima, Y., Takada, H., Aigaki, T., and Kurata, S. (2002). Overexpression of a pattern-recognition receptor, peptidoglycan-recognition protein-LE, activates imd/relish-mediated antibacterial defense and the prophenoloxidase cascade in Drosophila larvae. Proc. Natl. Acad. Sci. U. S. A. *99*, 13705–13710.

Takehana, A., Yano, T., Mita, S., Kotani, A., Oshima, Y., and Kurata, S. (2004). Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in Drosophila immunity. EMBO J. *23*, 4690–4700.

Tang, H., Kambris, Z., Lemaitre, B., and Hashimoto, C. (2006). Two proteases defining a melanization cascade in the immune system of Drosophila. J. Biol. Chem. *281*, 28097–28104.

Tang, H., Kambris, Z., Lemaitre, B., and Hashimoto, C. (2008). A serpin that regulates immune melanization in the respiratory system of Drosophila. Dev. Cell *15*, 617–626.

Tauszig-Delamasure, S., Bilak, H., Capovilla, M., Hoffmann, J. A., and Imler, J.-L. (2002). Drosophila MyD88 is required for the response to fungal and Gram-positive bacterial infections. Nat. Immunol. *3*, 91–97.

Theopold, U., Schmidt, O., Söderhäll, K., and Dushay, M.S. (2004). Coagulation in arthropods: defence, wound closure and healing. Trends Immunol. *25*, 289–294.

Tzou, P., Ohresser, S., Ferrandon, D., Capovilla, M., Reichhart, J.-M., Lemaitre, B., Hoffmann, J. A., and Imler, J. -L. (2000). Tissue-specific inducible expression of antimicrobial peptide genes in Drosophila surface epithelia. Immunity *13*, 737–748.

Valanne, S., Wang, J.-H., and Rämet, M. (2011). The Drosophila Toll signaling pathway. J. Immunol. *186*, 649–656.

Wang, L., Weber, A.N.R., Atilano, M.L., Filipe, S.R., Gay, N.J., and Ligoxygakis, P. (2006). Sensing of Gram-positive bacteria in Drosophila: GNBP1 is needed to process and present peptidoglycan to PGRP-SA. EMBO J. *25*, 5005–5014.

Wang, L., Gilbert, R.J.C., Atilano, M.L., Filipe, S.R., Gay, N.J., and Ligoxygakis, P. (2008). Peptidoglycan recognition protein-SD provides versatility of receptor formation in Drosophila immunity. Proc. Natl. Acad. Sci. U. S. A. *105*, 11881–11886.

Wang, Z., Wilhelmsson, C., Hyrsl, P., Loof, T.G., Dobes, P., Klupp, M., Loseva, O., Mörgelin, M., Iklé, J., Cripps, R.M., et al. (2010). Pathogen entrapment by transglutaminase--a conserved early innate immune mechanism. PLoS Pathog. *6*, e1000763.

Weber, A.N.R., Tauszig-delamasure, S., Hoffmann, J. A., Lelièvre, E., Gascan, H., Ray, K.P., Morse, M.A., Imler, J., and Gay, N.J. (2003). Binding of the Drosophila cytokine Spätzle to Toll is direct and establishes signaling. *4*, 794–800.

Weber, A.N.R., Gangloff, M., Moncrieffe, M.C., Hyvert, Y., Imler, J.-L., and Gay, N.J. (2007). Role of the Spatzle Pro-domain in the generation of an active toll receptor ligand. J. Biol. Chem. *282*, 13522–13531.

Werner, T., Liu, G., Kang, D., Ekengren, S., Steiner, H., and Hultmark, D. (2000). A family of peptidoglycan recognition proteins in the fruit fly Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. *97*, 13772–13777.

Xie, J., Butler, S., Sanchez, G., and Mateos, M. (2013). Male killing Spiroplasma protects Drosophila melanogaster against two parasitoid wasps. Heredity (Edinb). *112*, 399–408.

Xu, J., and Cherry, S. (2014). Viruses and antiviral immunity in Drosophila. Dev. Comp. Immunol. *42*, 67–84.

Yoshida, H., Kinoshita, K., and Ashida, M. (1996). Purification of a Peptidoglycan Recognition Protein from Hemolymph of the Silkworm , Bombyx mori. J. Biol. Chem.

Zaidman-Rémy, A., Hervé, M., Poidevin, M., Pili-Floury, S., Kim, M.-S., Blanot, D., Oh, B.-H., Ueda, R., Mengin-Lecreulx, D., and Lemaitre, B. (2006). The Drosophila amidase PGRP-LB modulates the immune response to bacterial infection. Immunity *24*, 463–473.

Zambon, R.A., Nandakumar, M., Vakharia, V.N., and Wu, L.P. (2005). The Toll pathway is important for an antiviral response in Drosophila. Proc. Natl. Acad. Sci. United States Am. *102* , 7257–7262.

Résumé de thèse

Pour se défendre contre les infections, la drosophile dispose d'une réponse immunitaire innée, c'est à dire une réponse qui ne nécessite pas de recombinaison pour former des récepteurs spécifiques comme les anticorps. La réponse immunitaire innée est rapide et généralement très efficace. Chez les vertébrés elle est la première ligne de défense avant qu'une réponse adaptative ne puisse se mettre en place. Sa principale caractéristique, en parallèle avec la phagocytose, est la sécrétion dans l'hémolymphe de peptides antimicrobiens (PAM) par les cellules du corps gras, équivalents fonctionnels du foie de mammifères. Les PAM sont toxiques pour les bactéries ou champignons en provoquant des perforations de leurs parois. La synthèse des PAM dépend de deux voies de signalisation, la voie Toll activée par la plupart des bactéries à Gram positif (possédant un peptidoglycan de type Lys) et les champignons et la voie IMD activée par la plupart des bactéries à Gram négatif (possédant un peptidoglycan de type DAP).

A la fin des années 1980, Charles Janeway proposa que la réponse immunitaire innée repose principalement sur la discrimination entre le soi et le non-soi. Il postula l'existence de récepteurs reconnaissant les pathogènes (*Pattern/Pathogen Recognition Receptors : PRR)* encodés par des gènes ne nécessitant pas de recombinaison somatique. Ils sont capables de reconnaître des molécules conservées chez les pathogènes, en particulier des composants de leurs parois, qu'il a appelé PAMP pour *Pathogen Associated Molecular Patterns*.

Quelques années plus tard, Polly Matzinger proposa que les mécanismes principaux d'activation de la réponse immunitaire innée ne reposent pas sur la discrimination entre le soi et le non-soi mais sur la détection de signaux de danger produits par les activités microbiennes: ce peuvent être soit des facteurs de virulence microbiens, soit des molécules libérées par les cellules endommagées de l'hôte.

Notre laboratoire a pu montrer que chez la drosophile les deux mécanismes sont à l'œuvre pour l'activation de la voie Toll. Toll est un récepteur membranaire activé par un ligand, Spaetzle, qui doit subir préalablement une maturation par coupure protéolytique de sa partie N-terminale. Ceci est le résultat de l'activation de plusieurs cascades protéolytiques où chaque protéase est synthétisée sous forme d'un zymogène qui doit être activé par coupure de son domaine N-terminal. Les PAMP sont reconnus par des récepteurs circulants et induisent l'activation de la cascade Modular Serine Protease (ModSP), Grass et Spaetzle Processing Enzyme (SPE). Les protéases microbiennes et des signaux endogènes non identifiés activent la protéase Persephone (Psh) qui active également SPE. Ces cascades, telles qu'elles sont connues actuellement, ne sont pas complètes et l'objet de mon travail de thèse a été d'identifier des protéases manquantes.

Dans une première partie, je me suis intéressée à la cascade en aval des récepteurs circulants. Nous savons que ModSP n'a pas la spécificité requise pour pouvoir activer Grass : c'est une chymotrypsine et Grass doit être activée par une trypsine. Il y a donc une protéase qui doit être activée par ModSP et capable d'activer Grass. Parmi les quelque 200 protéases à sérine identifiées dans le génome de la drosophile trois gènes répondaient aux différents critères issus d'une analyse bioinformatique: *CG4927, CG2056* et *CG3700*. Mon travail a consisté à générer des mutants pour ces gènes et à tester leur implication dans l'activation de la voie Toll.

Une lignée comportant une insertion inactivatrice dans le gène *CG3700* m'a permis de montrer que ce gène n'est pas impliqué dans la voie Toll. Pour les deux autres gènes, aucun des outils de génétique classique (insertion et excision d'éléments transposables, ARN interférence…) ne m'ont permis d'obtenir leur inactivation fonctionnelle. J'ai alors tenté une approche dirigée et utilisé les Transcription Activator-Like Effector Nucleases (TALEN). Cette méthode a été validée en 2012 chez la drosophile, mais je n'ai pas pu obtenir non plus de mutant. La dernière approche, l'utilisation du système Clustered, Regularly Interspaced, Short Palindromic Repeats (CRISPR)/Cas9 m'a enfin permis d'obtenir un mutant du gène *CG2056* qui est en cours de caractérisation.

Dans une deuxième partie, je me suis intéressée à la cascade en aval des signaux de danger. J'ai en effet caractérisé la fonction du gène *spheroide*. Ce gène code pour un homologue de protéase à sérine, ce qui veut dire que la protéine est très similaire à une protéase mais que certains des 3 acides aminés essentiels du site catalytique sont mutés. La protéine est donc inactive. Certaines de ces protéines ont été identifiées dans des cascades de protéases et jouent donc un rôle qui n'a pour l'instant pas pu être compris. Une approche par ARN interférence avait montré que Spheroide pouvait être impliquée dans l'activation de la réponse immunitaire.

J'ai utilisé une insertion inactivatrice dans le gène *spheroide* ainsi qu'une délétion provoquée par l'excision de cet élément transposable pour obtenir des allèles nuls de *spheroide*. Les mouches portant ces allèles sont susceptibles à des infections par les bactéries pathogéniques *Enterococcus faecalis* et *Staphylococcus aureus*. Elles présentent une réponse immunitaire réduite en réponse à ces infections. En revanche, leur réponse immunitaire à une infection par des bactéries non pathogéniques, *Micrococcus luteus,* est parfaitement normale. Ceci est confirmé par l'infection des mouches par des bactéries *E. faecalis* mutantes qui n'expriment pas les deux protéases principales : *spheroide* n'est alors plus nécessaire pour activer la réponse immunitaire. Sachant que les bactéries non pathogéniques n'activent que la cascade en aval des récepteurs circulants aux PAMP et pas la cascade activée par les signaux de danger, ceci suggère que Spheroide est impliquée spécifiquement dans cette deuxième voie. Nous savons que l'activation de la voie Toll en réponse à une infection par *E. faecalis* n'est complètement abolie que dans des mouches mutantes pour à la fois Psh, impliquée dans la voie des signaux de danger, et Grass, impliquée dans la voie des PAMP. J'ai pu confirmer de la même manière que dans les mouches mutantes à la fois pour Spheroide et pour Grass, la voie Toll n'est plus activée. Spheroide fonctionne donc dans la même cascade que Psh, en aval des signaux de danger. De plus les analyses épistatiques nous montrent que Spheroide est requise en aval de Psh pour l'activation de la voie Toll.

J'ai donc pu montrer, pour la première fois chez la drosophile, qu'une protéase inactive est nécessaire au bon fonctionnement d'une cascade protéolytique. Spheroide ne subit pas de modification lors d'une infection. A ce jour, il ne m'est pas

possible de préciser sa fonction. Il est probable que Spheroide fonctionne comme un cofacteur nécessaire soit à l'activation de Psh soit à la formation d'un complexe permettant l'activation de la cascade protéolytique. Cependant, j'ai aussi pu montrer que Spheroide n'est pas requise pour activer la voie Toll en réponse à une infection par un champignon pathogénique, *Beauveria bassiana*. Nous savons pourtant que ce champignon exprime également des protéases qui activent Psh. Ceci suggère donc soit que la présence d'un co-facteur comme Spheroide n'est nécessaire que dans certaines circonstances, soit, et c'est plus probable, qu'un autre co-facteur est requis.

Le travail que j'ai mené vise à caractériser les différentes molécules impliquées dans les cascades protéolytiques en amont de Toll. Ceci permettra de reconstituer ces cascades *in vitro* et de comprendre comment elles s'organisent, comment des complexes peuvent se former, à quel endroit (circulant ou à la surface des cellules). Une question fondamentale que nous souhaitons approcher est de comprendre l'intérêt de ce système complexe d'activation du récepteur Toll suite à des cascades protéolytiques, sachant que ce système n'est pas conservé pour l'activation de la voie IMD par des bactéries à Gram-négatif ni pour l'activation des Toll-like receptors (TLR) chez les vertébrés. Une autre question est de savoir comment les signaux de danger sont exactement perçus, sachant que l'activation de la protéase Psh doit se faire spécifiquement par coupure après une histidine. Nous savons en revanche que cette détection des signaux de danger est essentielle. J'ai en effet pu montrer que c'est le premier événement qui active la voie Toll : la mouche détecte d'abord une activité suspecte avant de détecter le microorganisme. Cette précocité est vitale, car les mouches mutantes pour *psh* ou *spheroide*, qui peuvent monter une réponse immunitaire normale mais avec un léger retard, meurent beaucoup plus vite d'une infection que les mouches sauvages.

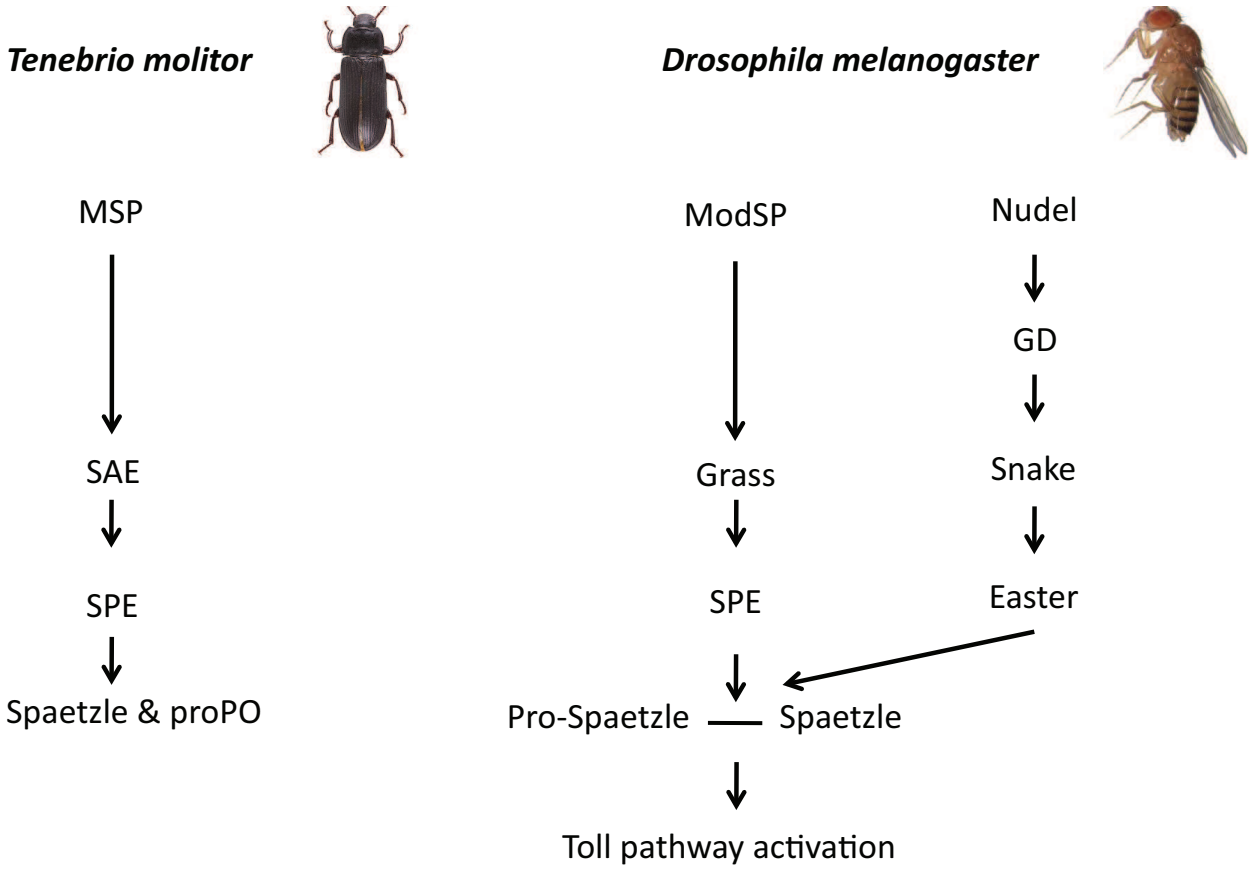


Figure 1 Comparaison des cascades protéolytiques contenant des protéases à sérine à domaine CLIP

A gauche, la cascade impliquée dans les réponses immunitaires chez *Tenebrio molitor*. A droite, les cascades impliquées dans l'activation de la voie Toll chez *Drosophila melanogaster* (réponse immunitaire et développement)

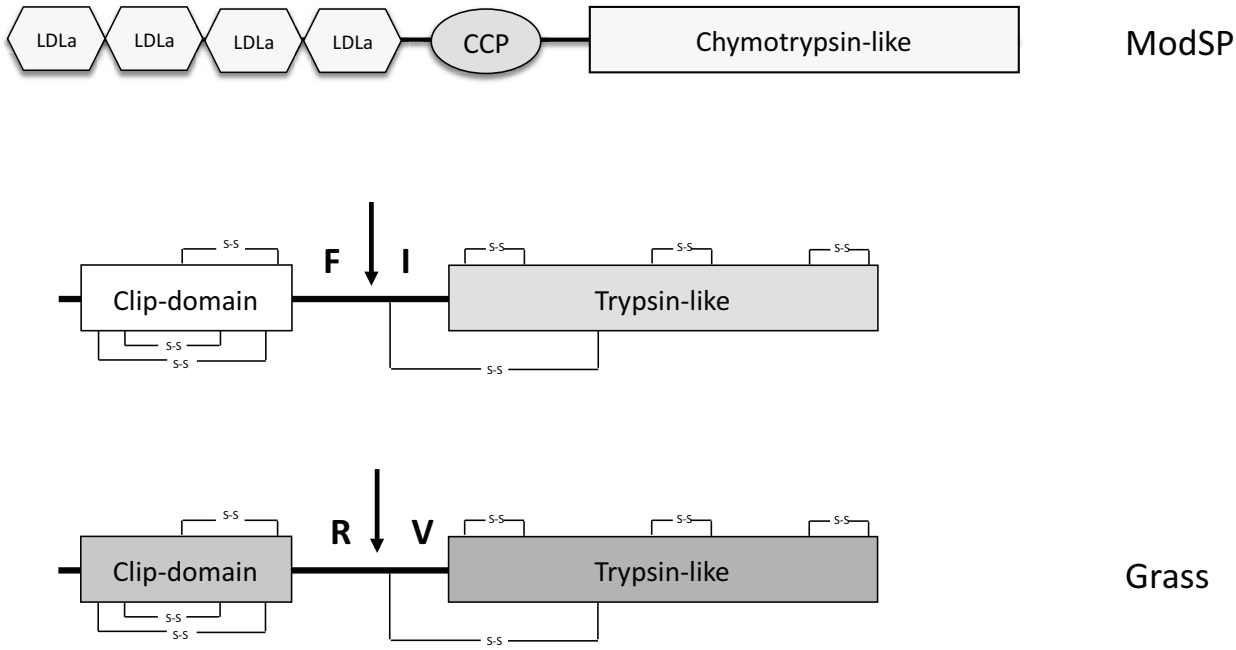
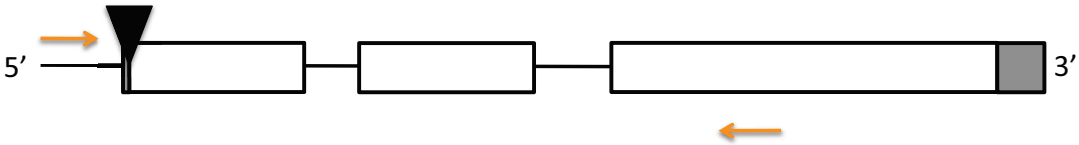
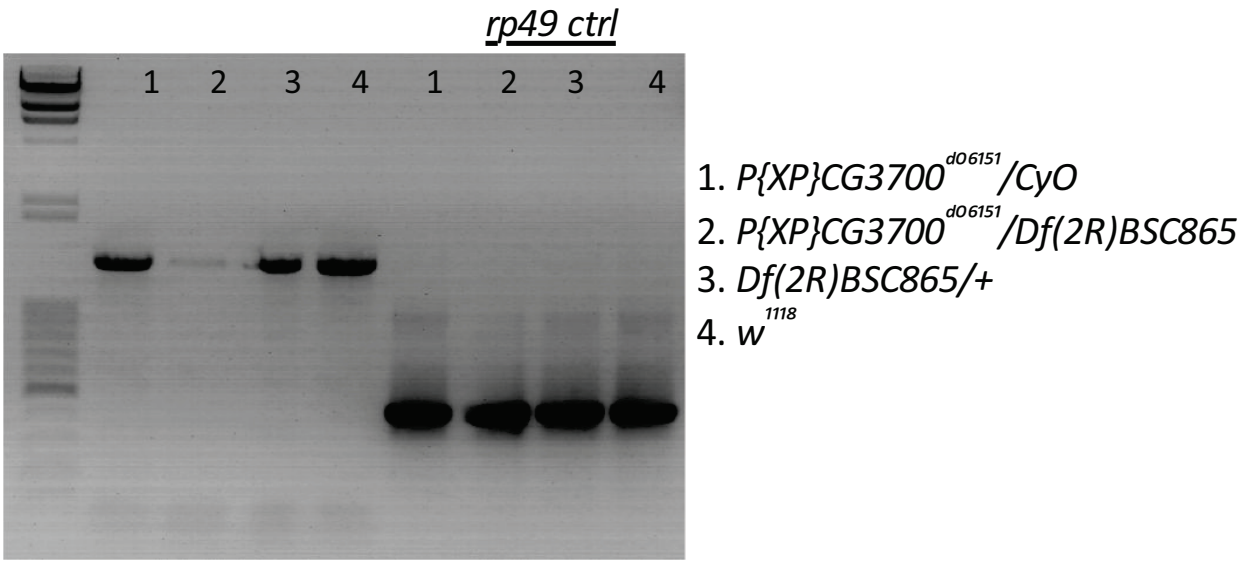


Figure 2 : représentation schématique de ModSP, Grass et la protéase potentielle comprenant un domain CLIP, le site d'activation spécifique et un domain catalytique de type trypsine.

A



B



C Immune challenge with *M. luteus*

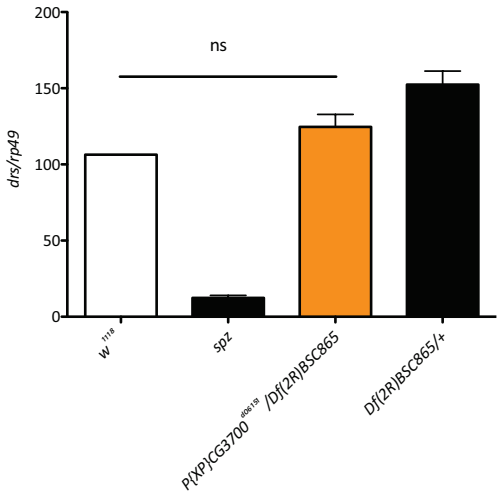
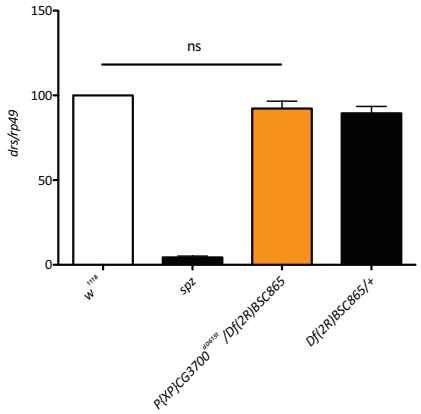


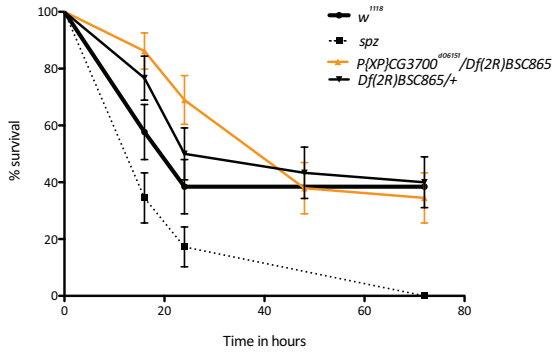
Figure 3 (suite page suivante)

Immune challenge with *E. faecalis*

D

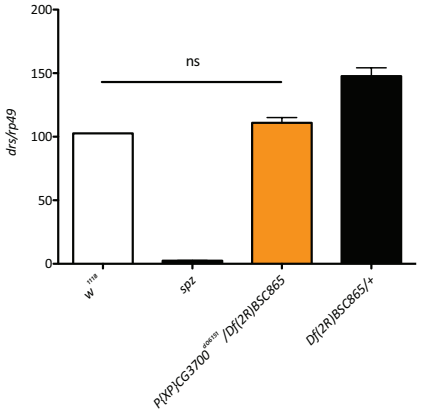


E



Natural infection with *B. bassiana*

F



G

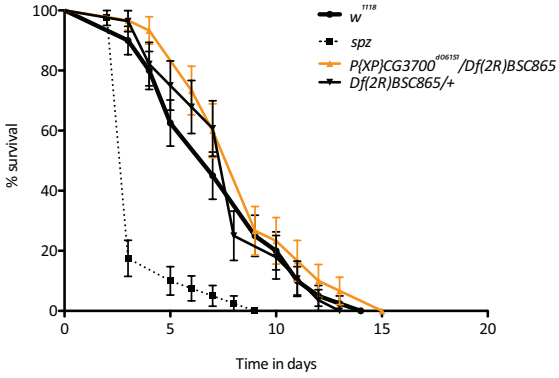


Figure 3 : *CG3700* n'est pas impliqué dans l'activation de la voie Toll

Representation schématique de la région génomique et de l'insertion de l'élément P (▼) dans *CG3700* (P{XP}CG3700^{d06151}). Les amorces utilisées pour la PCR sont indiquées par des flèches orange (**A**); Insertion de l'élément P dans *CG3700* dans de smouches hétérozygotes pour l'insertion et une délétion de cette région (P{XP}CG3700^{d06151}/ Df(2R)BSC865) visualisée par PCR (**B**); Expression de la *drosomycin* 24 heures après infection (PI), normalisée par rapport à *rp49* et en pourcentage de son expression dans des mouches sauvages *w^{1118}* après infection par piqure septique avec *M. luteus* (**C**) ou *E. faecalis* (**F**) ou infection naturelle avec *B. bassiana,* 48 hours PI (**G**). Courbe de survie après infection par piqure septique avec *E. faecalis* (**E**) ou infection naturelle avec *B. bassiana* (**G**).

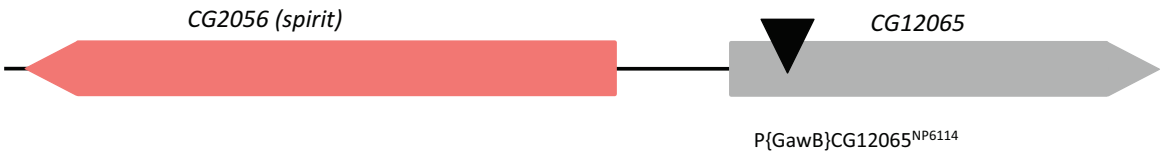


Figure 4 : représentation schématique de la stratégie d'excision de l'élément P pour *CG2056 (spirit)*

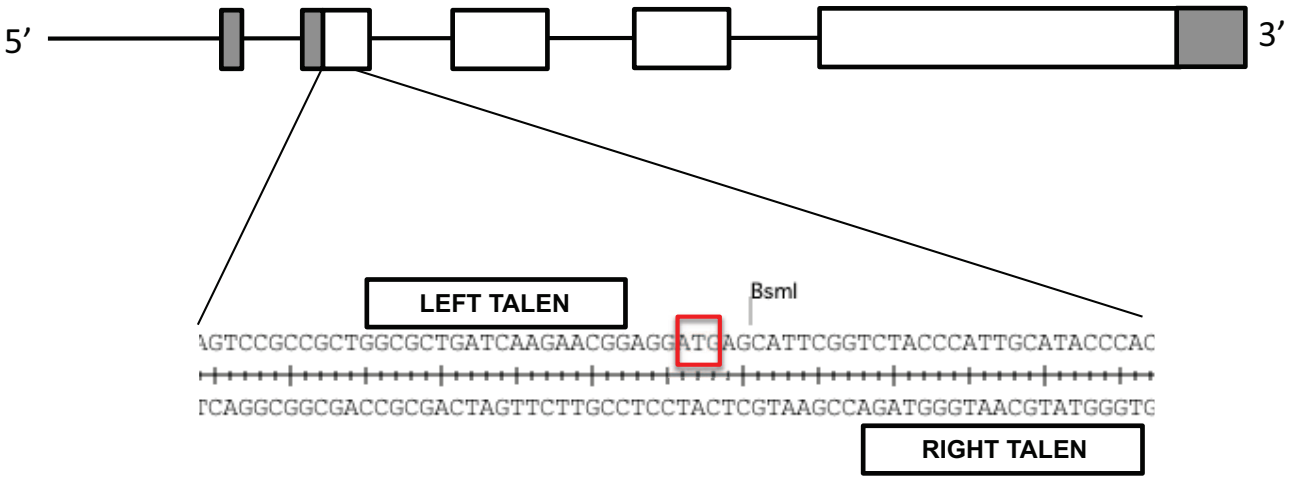
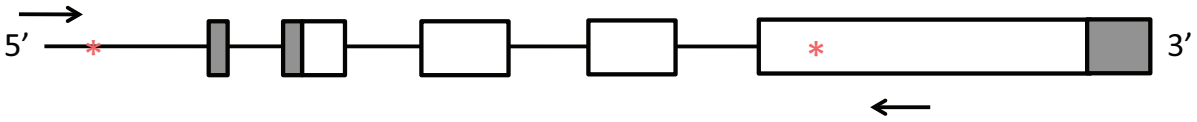


Figure 5 : représentation schématique de la stratégie utilisant les TALEN

Codon Start avec site de restriction pour BsmI dans *CG2056* (*spirit*)

***CG2056* (*spirit*)**

combination 1-3



combination 2-4

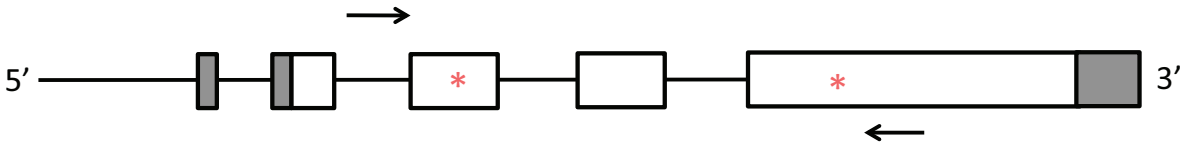


Figure 6 : représentation schématique de la stratégie CRISPR

Les asterisques représentent les séquences de *CG2056* (*spirit*) ciblées par les sgRNA injectés. Les flèches représentent les amorces utilisées pour vérifier les mutations.

gene	combination	G0 flies	F1 lines tested	mutants
SPIRIT	1-3	37	121	0
	2-4	37	119	1

Table 1 : résultats de la mutagénèse de *spirit* par CRIPSR

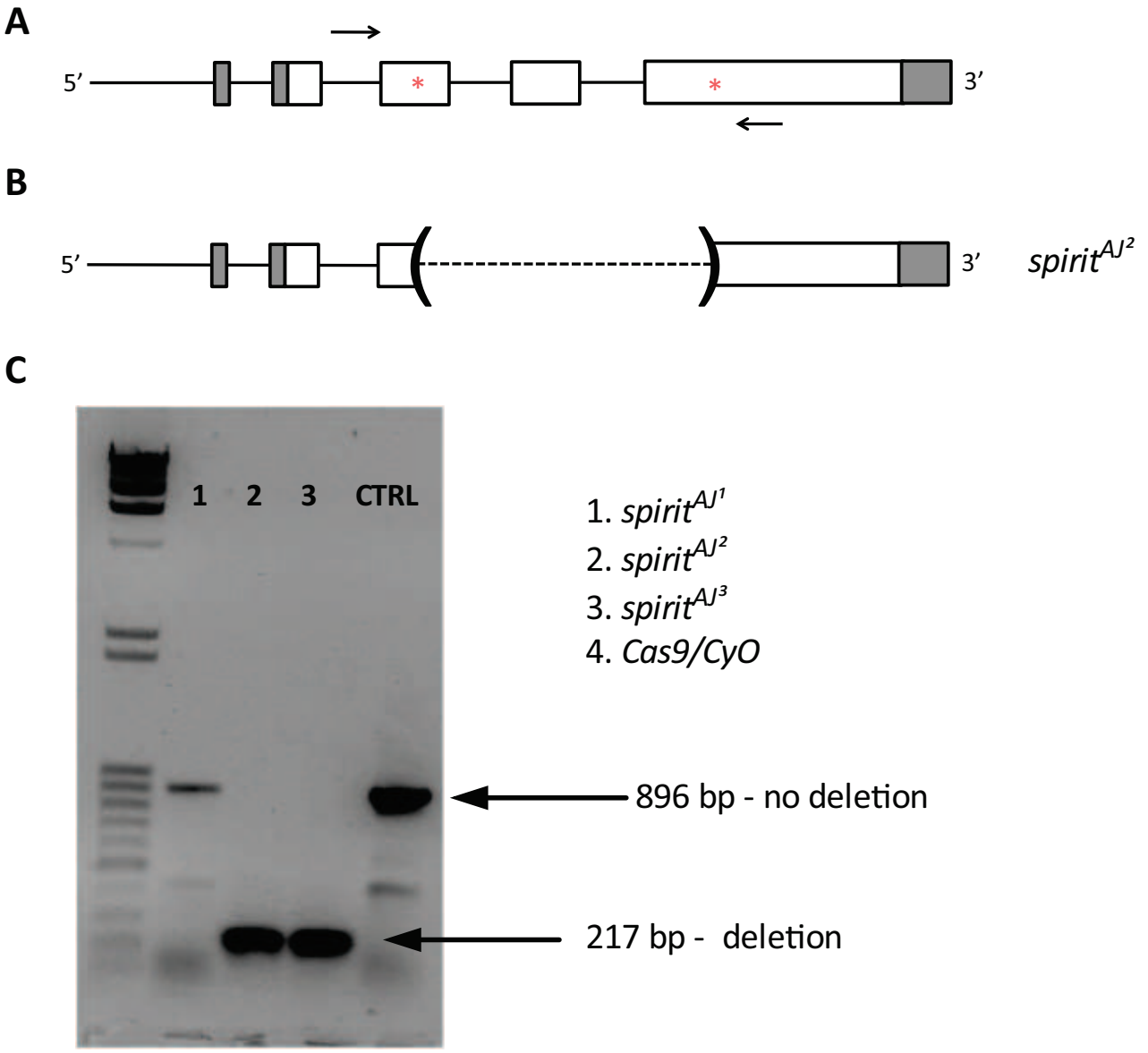


Figure 7 (suite page suivante)

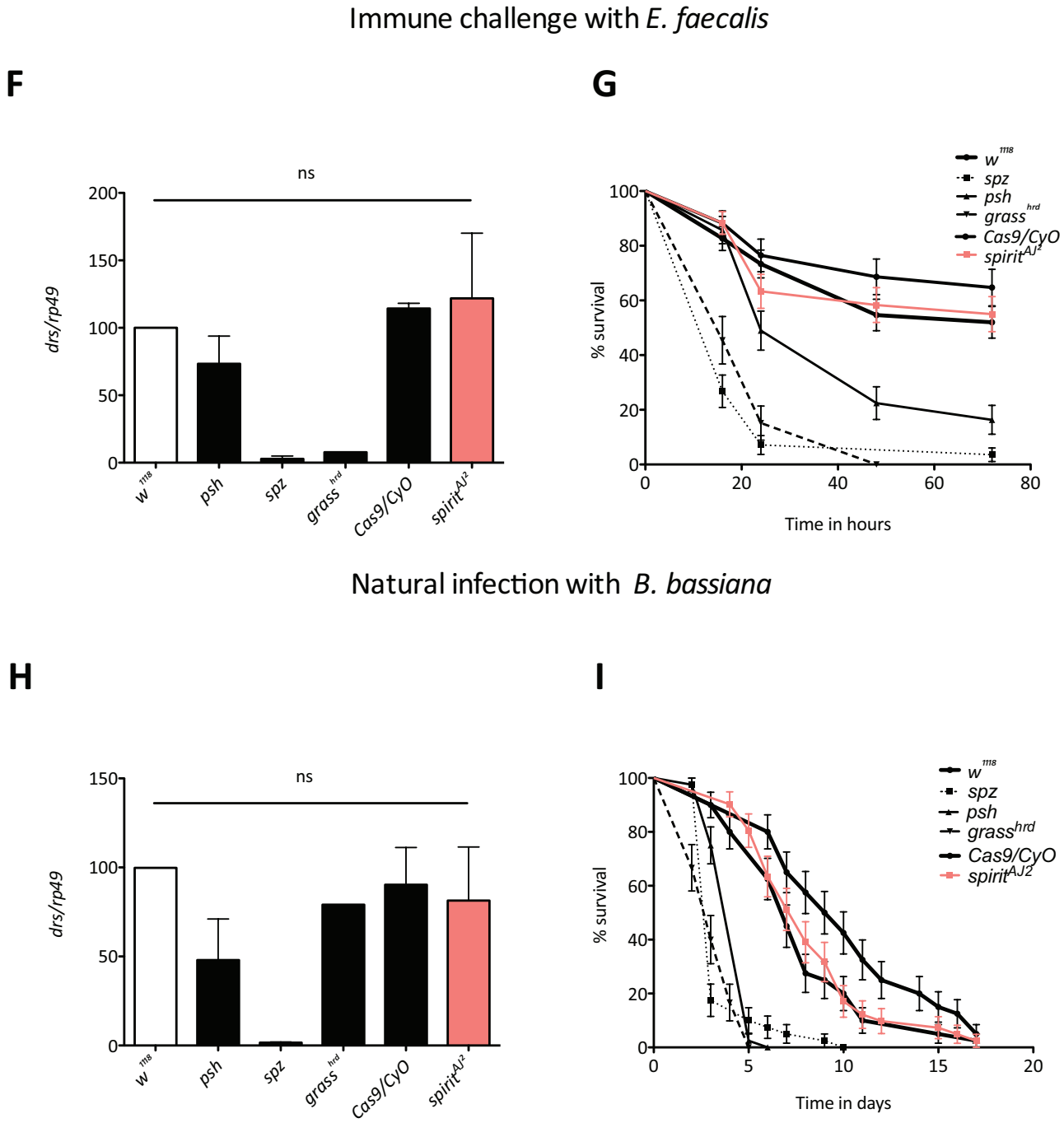


Figure 7: *Spirit* n'est pas impliqué dans l'activation de la voie Toll

Representation schématique de la région génomique de *CG2056 (spirit*) (**A**) et de la délétion *spirit^{AJ}* générée par le system CRISPR (**B**). Les asterisques représentent les séquences de *CG2056* (*spirit*) ciblées par les sgRNA injectés. Les flèches représentent les amorces utilisées pour vérifier les mutations (**C**). Expression de *spirit* normalisée par rapport à *rp49* (**D**). Expression de la *drosomycin* 24 heures après infection (PI), normalisée par rapport à *rp49* et en pourcentage de son expression dans des mouches sauvages *w^{1118}* après infection par piqure septique avec *M. luteus* (**E**) ou *E. faecalis* (**H**) ou infection naturelle avec *B. bassiana,* 48 hours PI (**I**). Courbe de survie après infection par piqure septique avec *E. faecalis* (**F**) ou infection naturelle avec *B. bassiana* (**G**).

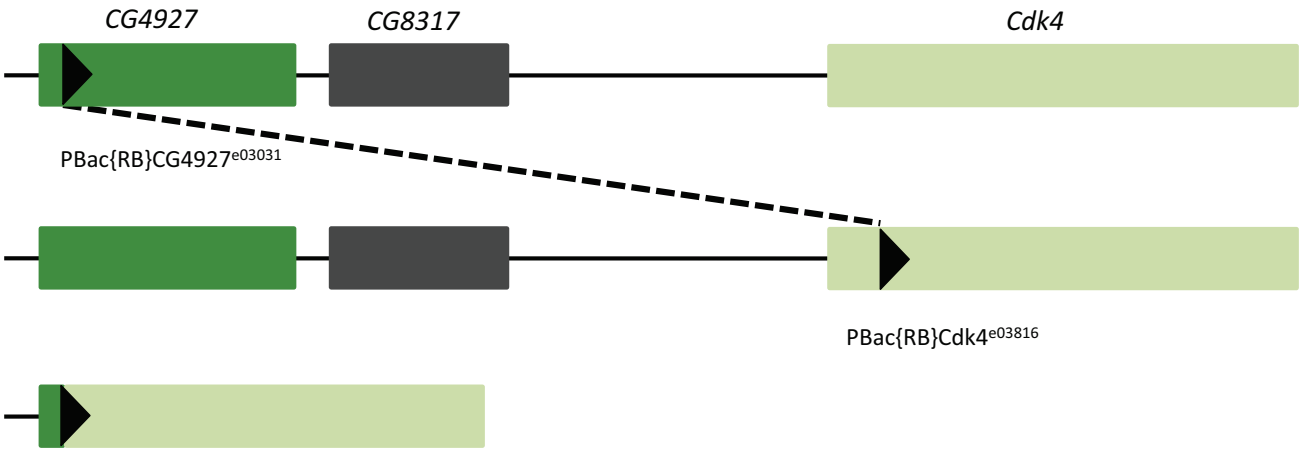


Figure 8 : représentation schématique de la stratégie FLP/FRT pour *CG4927*

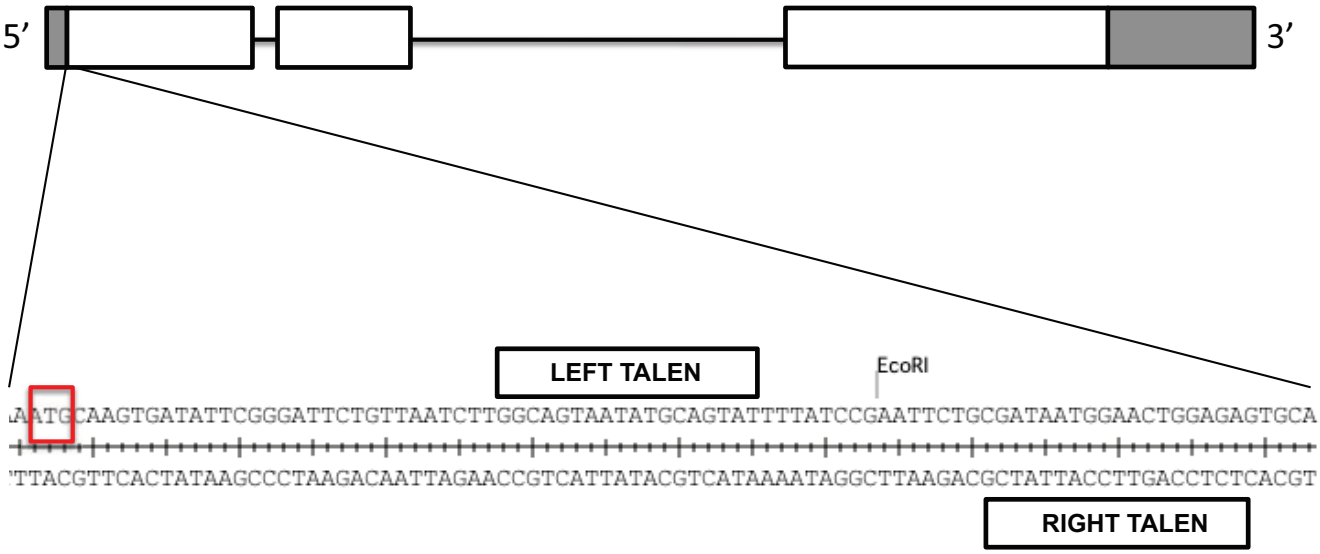


Figure 9: Représentation schématique de la stratégie utilisant les TALEN

Codon Start avec site de restriction pour EcoR1 dans *CG4927*

CG4927

combination 1-3

combination 2-4

Figure 10 : représentation schématique de la stratégie CRISPR

Les asterisques représentent les séquences de *CG4927* ciblées par les sgRNA injectés. Les flèches représentent les amorces utilisées pour vérifier les mutations.

gene	combination	G0	F1 lines tested	mutants
CG4927	1-3	23	109	0
	2-4	35	131	0

Table 2 : résultats de la mutagénèse de *CG4927* par CRISPR

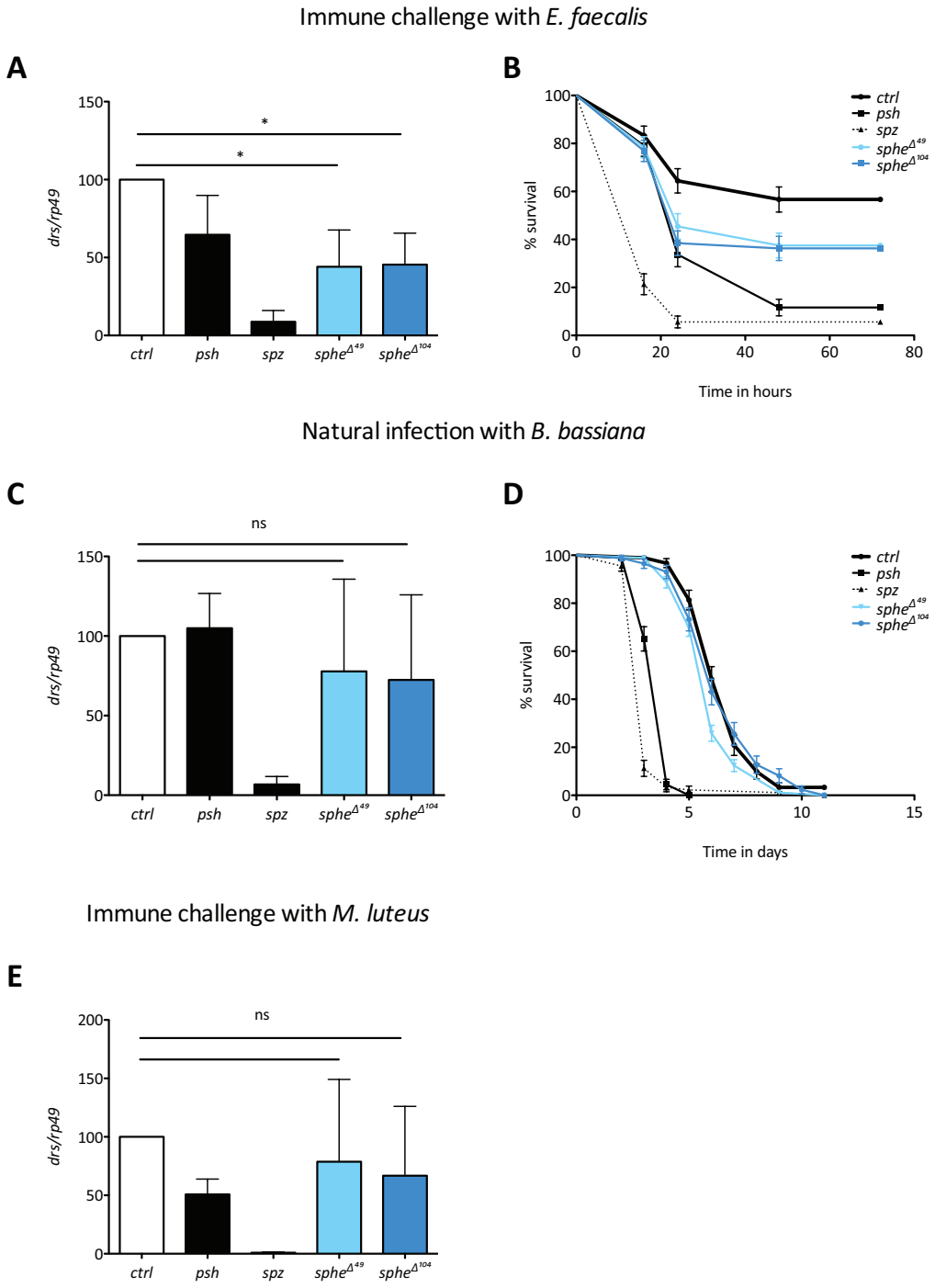


Figure 11 : Spheroide est impliquée dans l'activation de la réponse immunitaire après infection par *E. faecalis*

Expression de la *drosomycin* 24 heures PI, normalisée par rapport à *rp49* et en pourcentage de son expression dans les mouches sauvages $sphe^{\Delta 11}$ (ctrl) après infection par piqure septique avec *E. faecalis* (*ctrl* vs. $sphe^{\Delta 49}$ p<0.0261, *ctrl* vs. $sphe^{\Delta 104}$ p<0.02) (**A**) ou infection naturelle avec *B. bassiana* 48 hours PI (**C**) ou infection par piqure septique avec *M. luteus* (**E**). Courbe de survie infection par *E. faecalis* (**B**) ou infection naturelle par *B. bassiana* (**D**).

Immune challenge with *E. faecalis*

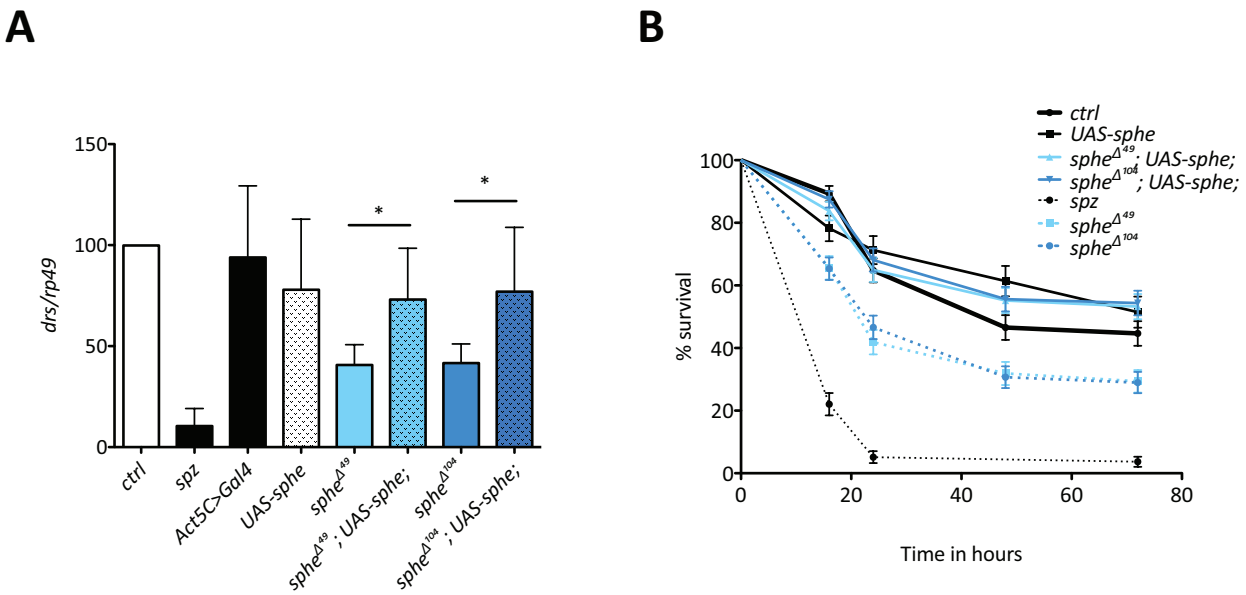


Figure 12 : sauvetage du phénotype mutant *sphe* à l'aide du système

Expression de la *drosomycin* 24 heures PI, normalisée par rapport à *rp49* et en pourcentage de son expression dans les mouches sauvages $sphe^{\Delta 11}$ (ctrl) après infection par piqure septique avec *E. faecalis* ($sphe^{\Delta 49}$ vs. $sphe^{\Delta 49}$; *UAS-sphe* p<0.0103, $sphe^{\Delta 104}$ vs. $sphe^{\Delta 104}$;*UAS-sphe* p<0.0229) (**A**). Courbe de survie infection par *E. faecalis* (**B**). Le système de surexpression ubiquitaire *Act5C>Gal4* a été utilisé pour induire l'expression de *UAS-sphe*.

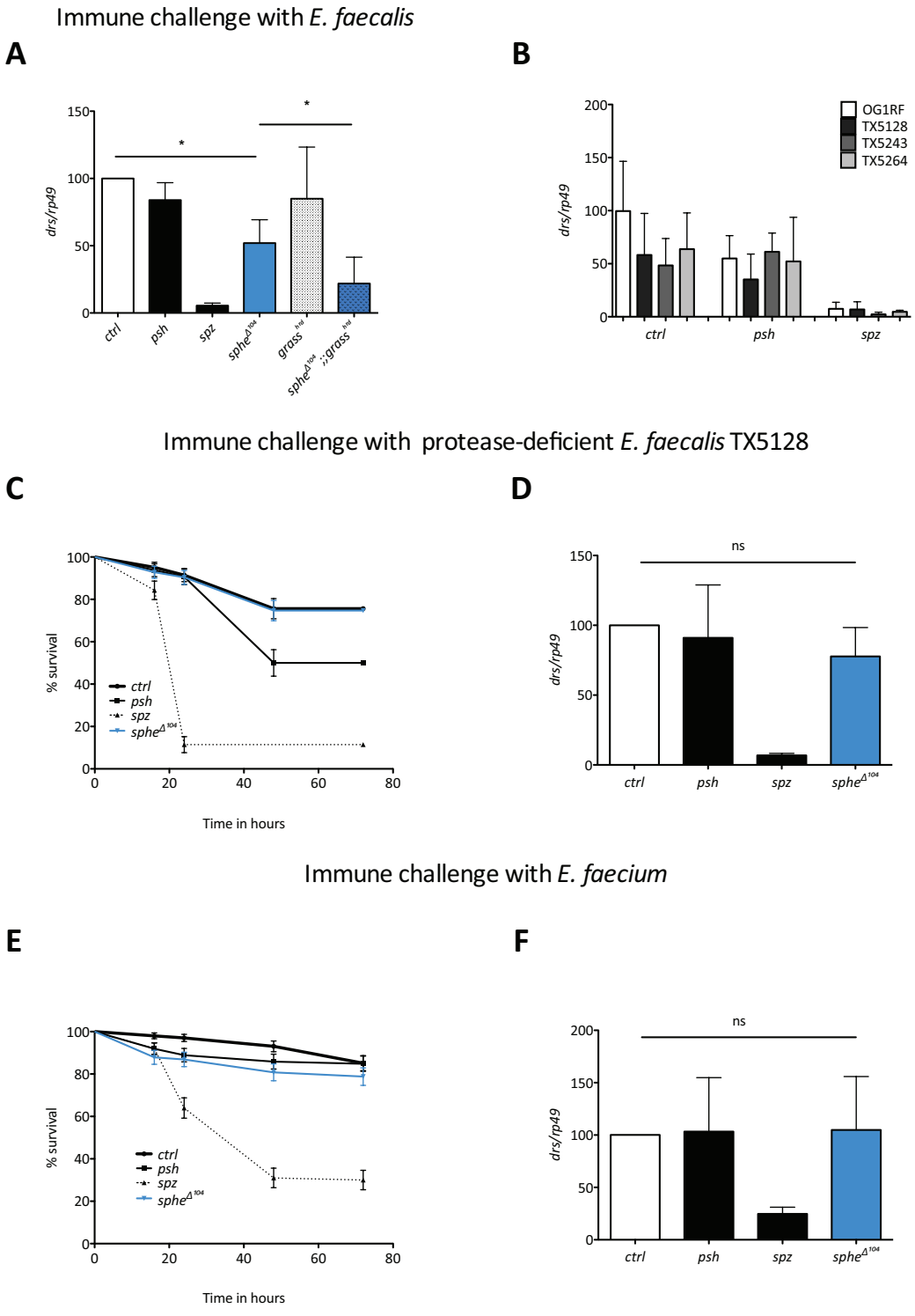


Figure 13 : Sphe est impliquée dans la cascade du signal de danger.
Expression de la *drosomycin* 24 heures PI, normalisée par rapport à *rp49* et en pourcentage de son expression dans les mouches sauvages $sphe^{\Delta 11}$ (ctrl) après infection par piqure septique avec *E. faecalis* de mouches simple ou double mutantes $sphe^{\Delta 104}$ ou/ et $grass^{hrd}$ (*ctrl* vs. $sphe^{\Delta 104}$ p<0.0178. *ctrl* vs. $sphe^{\Delta 104};;grass^{hrd}$ p<0.0264) (**A**) Comparaison de l'expression de *drs* après infection avec des bactéries sauvages ou n'exprimant pas de protéases (**B**). Courbe de survie (**C,E**) ou expression de drs (**D,F**) des mouches $sphe^{\Delta 104}$ après infection avec des bactéries *E. faecalis* TX5128 n'exprimant pas de protéases ou infection par piqure septique (**D**), ou *E. faecium* (**F**).

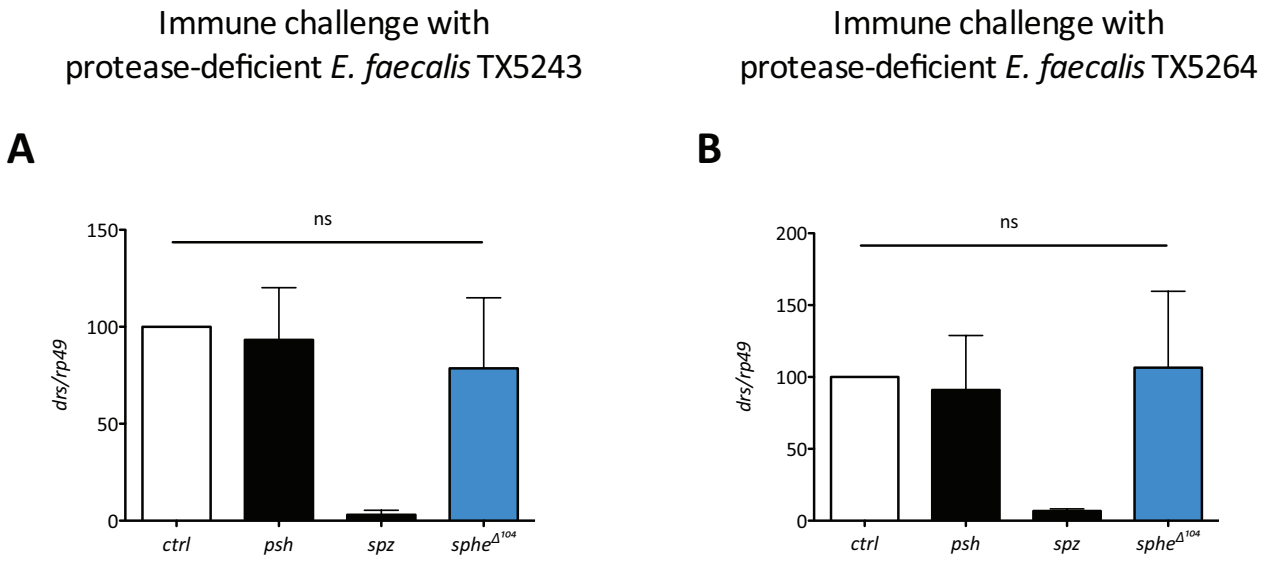


Figure 14 : Sphe est impliquée dans la détection des protéases bactériennes

Expression de la *drosomycin* 24 heures PI, normalisée par rapport à *rp49* et en pourcentage de son expression dans les mouches sauvages *sphe*$^{\Delta 11}$ wild type flies (ctrl) après infection par les bactéries *E. faecalis n'exprimant pas de protéases* TX5243 (**A**) or TX5264 (**B**).

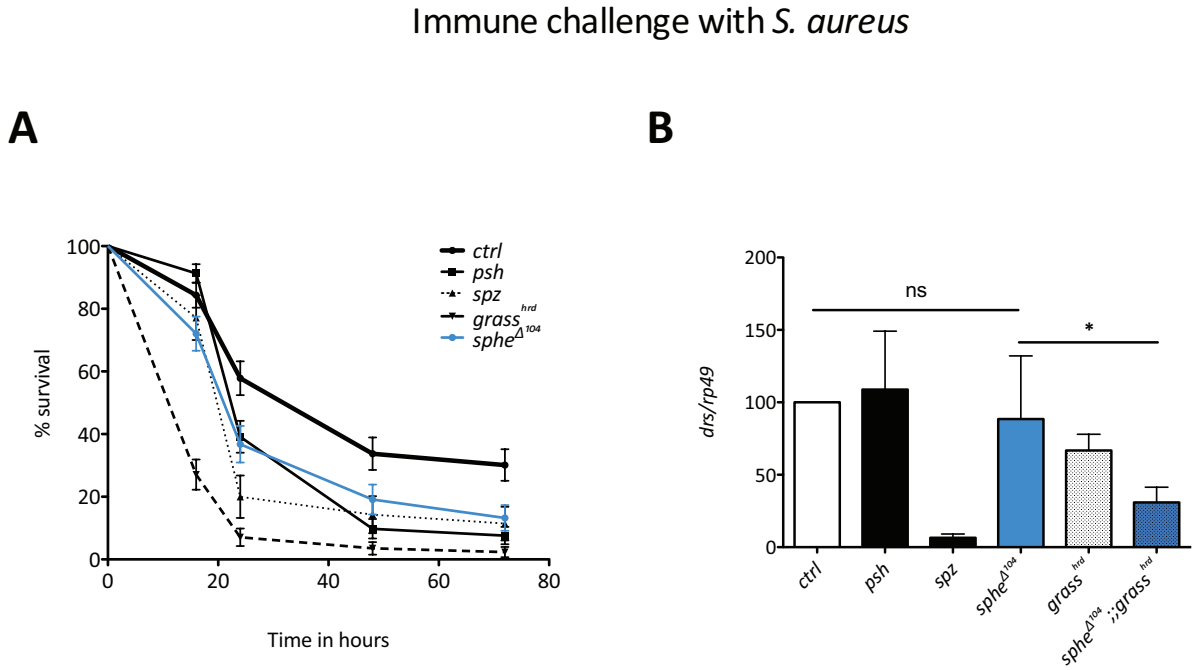


Figure 15 : Sphe est impliquée dans la détection de *S. aureus*

Infection par *S. aureus* : Courbe de survie(**A**). Expression de la *drosomycin* 24 heures PI, normalisée par rapport à *rp49* et en pourcentage de son expression dans les mouches sauvages *sphe*$^{\Delta 11}$ wild type flies (ctrl)) (*ctrl* vs. *sphe*$^{\Delta 104}$*;;grass*^{hrd} p<0.0144) (**B**).

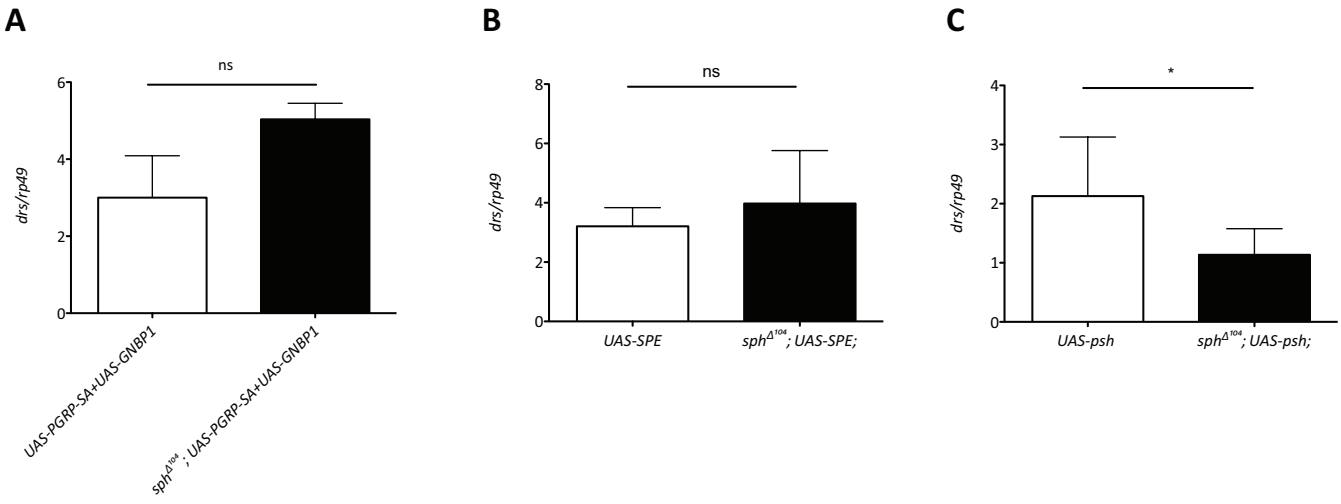
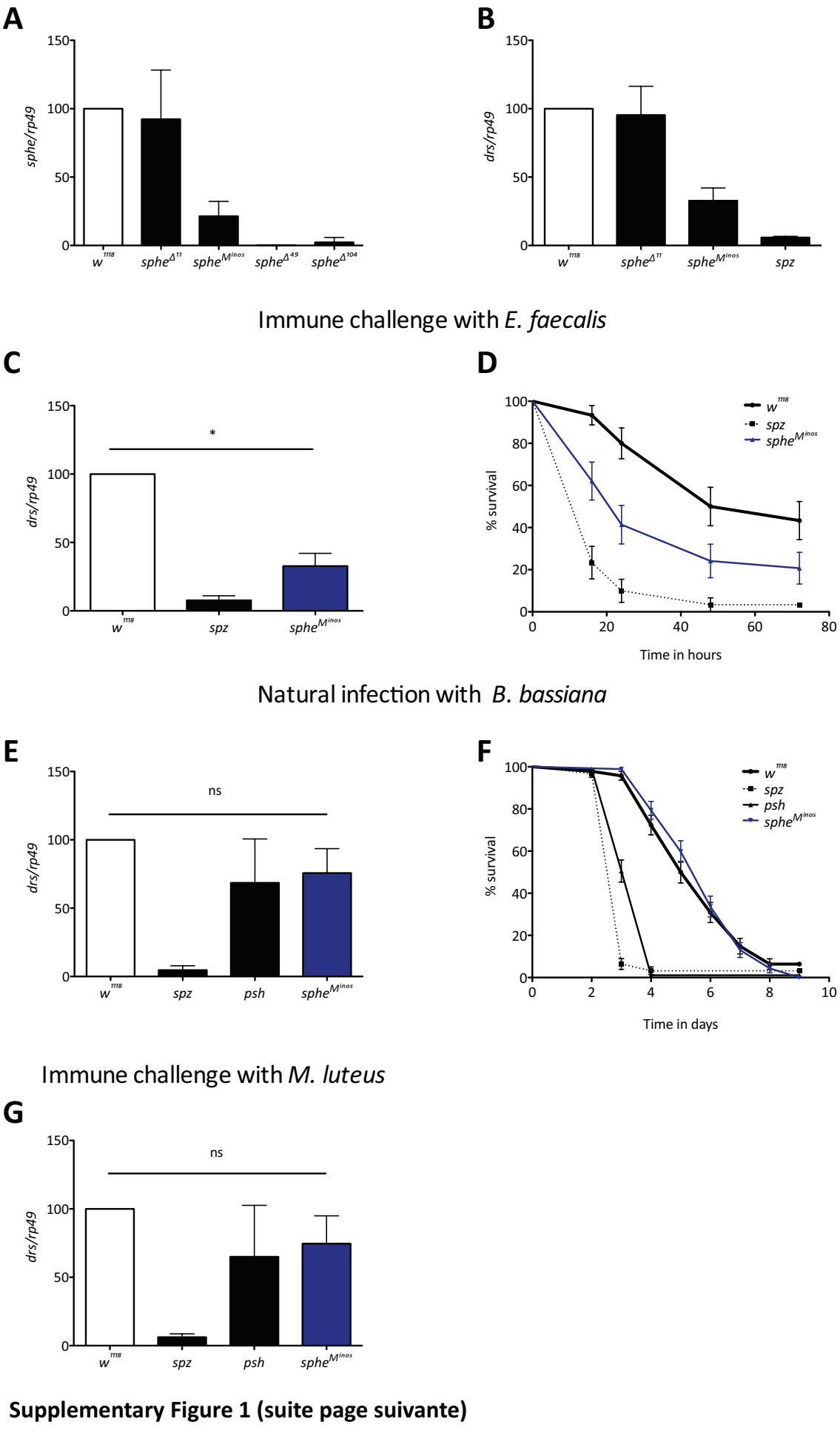


Figure 16 : Spheroide fonctionne dans la voie Psh

Expression de la *drosomycin* 24 heures PI, normalisée par rapport à *rp49* après surexpression à l'aide du *yolk>Gal4* driver de PGRP-SA et GNBP1 (**A**) ou de SPE (**B**) ou de Psh à l'aide du driver *Act5C>Gal4* (**C**) (*UAS-psh* vs. *sphe$^{\Delta 104}$; UAS-psh* p<0.0398).



Supplementary Figure 1 (suite page suivante)

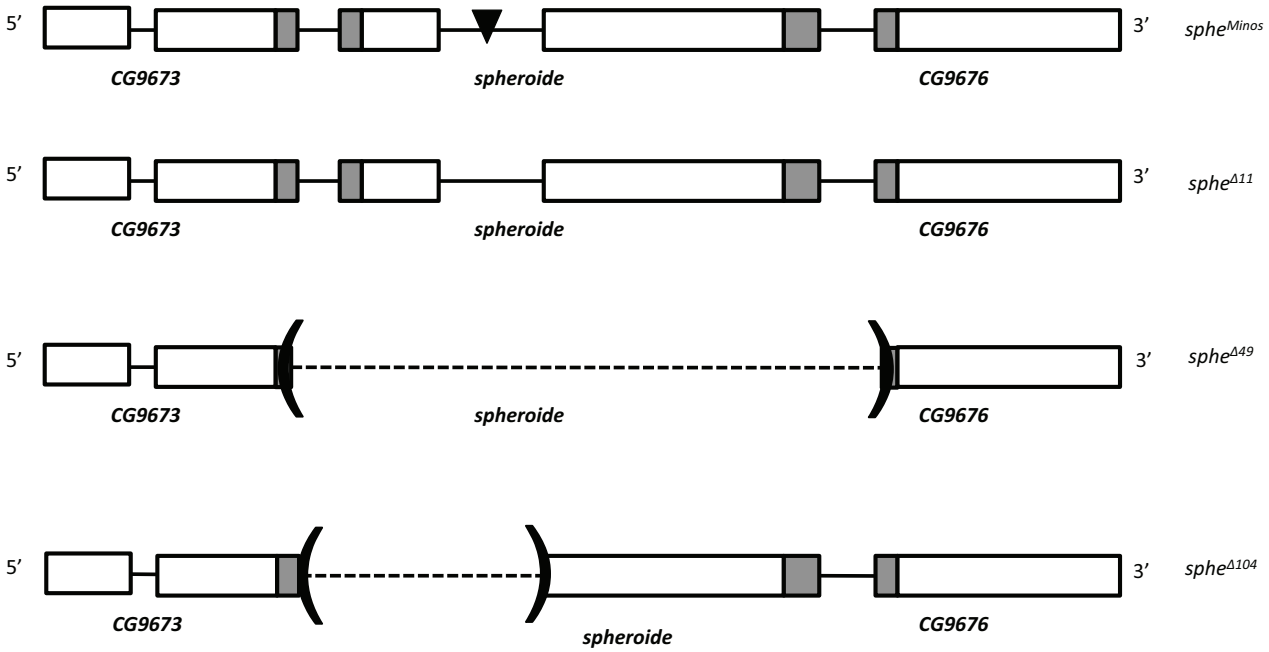


Figure 16 : création d'un mutant dans *spheroide* par excision d'un élément Minos

Expression de *spheroide* 24 heures PI, normalisée par rapport à *rp49* et en pourcentage de son expression dans les mouches sauvages w^{1118} (ctrl) (**A**). Expression de la *drosomycin* 24 heures PI, normalisée par rapport à *rp49* et en pourcentage de son expression dans les mouches sauvages $w^{1118\ 1}$ (ctrl) après infection par piqure septique avec *E. faecalis* (w^{1118} vs. $sphe^{minos}$ p<0.0335) (**B** et **C)** ou infection naturelle avec *B. bassiana* 48 heures PI (**E**) ou *M. luteus* (**G**). Courbe de survie infection par *E. faecalis* (**D**) ou infection naturelle par *B. bassiana* (**F**). Représentation schématique des délétions obtenues par excision de l'élément Minos (**H**).

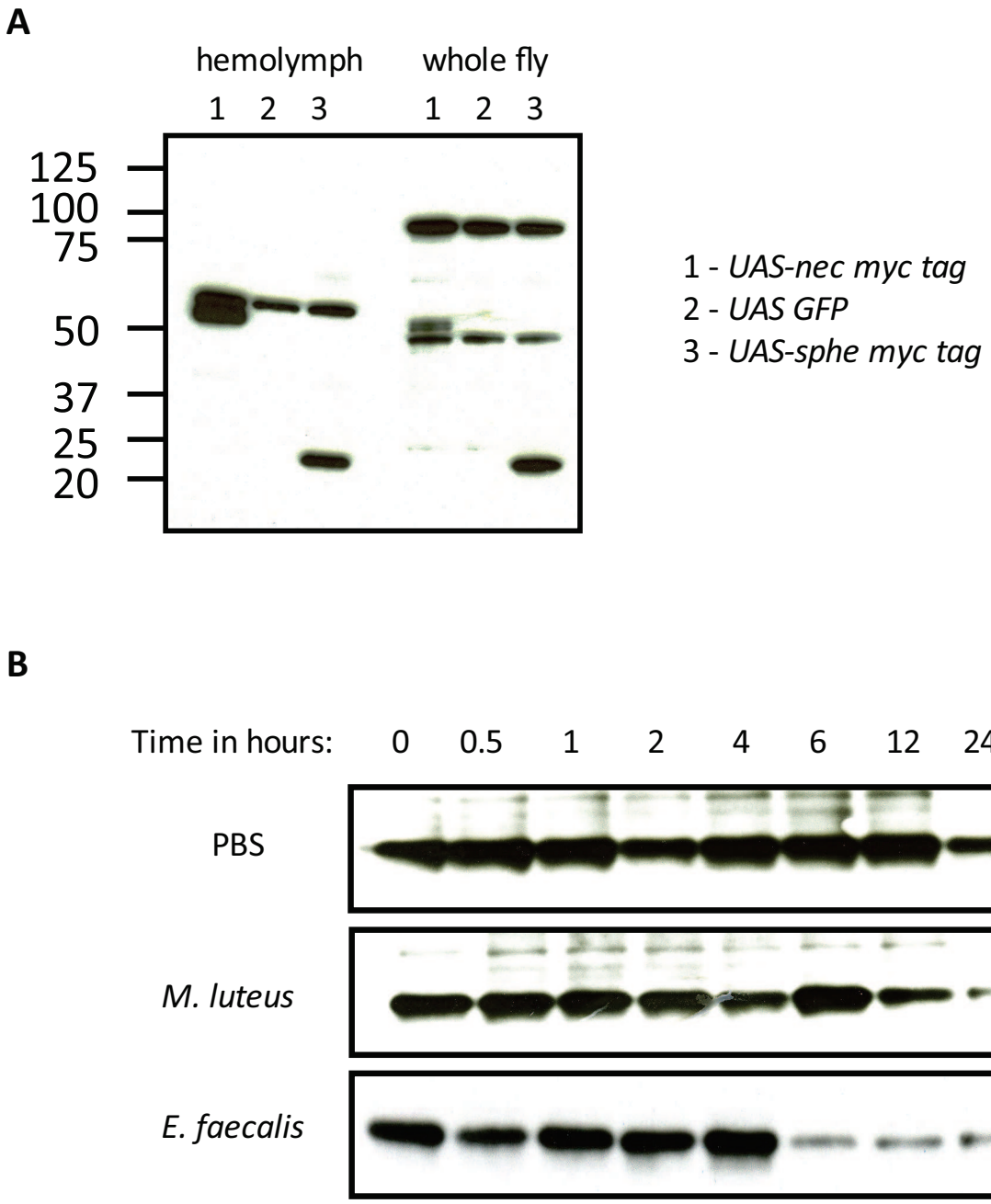


Figure 17 : la protéine Spheroide est secrétée dans l'hemolymphe et dérardée après infection

Surexpression de Spheroide à l'aide de *Act5C>Gal4* (2 jours à 29°C) détectée par des anticorps anti-Myc : localisation dans l'hémolymphe (**A**); Expression au cours du temps après piqûre propre (PBS) ou septique avec *M.luteus* ou *E. faecalis* (**B**). Les extraits de mouches entières sont préparés avec un tampon de lyse 30 mM Hepes pH 7.4, 150 mM NaCl, 2 mM MgAc, 1 % NP-40.

Jelena PATRNOGIĆ

Serine proteases and serine protease homologs : genetic analysis of their involvement in immune response activation in *Drosophila*

Lors de la réponse immunitaire de la drosophile, la voie Toll est activée lors d'un challenge immunitaire par des bactéries à Gram positif ou des champignons. Ce mécanisme est initié soit par la reconnaissance de motifs moléculaires associés aux pathogènes (PAMPs) qui activent la voie de reconnaissance, ou par des facteurs de virulence et des protéases produits par les agents pathogènes qui activent la voie des signaux de danger. Le travail que j'ai effectué a pour but de caractériser les différentes molécules impliquées dans ces cascades protéolytiques en amont de Toll. Cela permettra de reconstituer ces cascades *in vitro* et de comprendre comment elles sont organisées, comment et où des complexes peuvent être formés. La première partie concerne les approches génétiques utilisées pour générer des mutants des gênes pouvant être impliqués dans l'activation de la voie Toll par la voie des PAMPs. La deuxième partie se concentre sur un homologue inactif de protéase à sérine appelé *spheroide* et sur son implication dans la voie de reconnaissance des signaux de danger. Pour la première fois, nous avons pu démontrer qu'une protéase inactive est requise dans la cascade protéolytique, et plus particulièrement dans la détection des signaux de danger après un challenge immunitaire par des bactéries pathogènes à Gram positif.

Mots clés: immunité innée, *Drosophila*, protéase à sérine, signal de danger

The Toll pathway in *Drosophila* immune response is activated upon immune challenge with Gram-positive bacteria and fungi. This can be achieved either through recognition of Pathogen Associated Molecular Patterns (PAMPs), which triggers the recognition cascade; or by virulence factors and proteases produced by the pathogens, which triggers the danger signal cascade. The work I have done aimed to characterize the various molecules involved in proteolytic cascades upstream of Toll. This will help to reconstitute these cascades *in vitro* and understand how they are organized, how and where complexes could be formed. The first part focuses on genetic approaches used to generate mutants for genes suggested to be involved in the activation of Toll pathway via the recognition cascade. The second part focuses on an inactive serine protease, a serine protease homolog *spheroide* and its involvement in the danger signal cascade. For the first time, we could demonstrate that an inactive protease is required in the proteolytic cascade, involved in the sensing of danger signals upon immune challenge with pathogenic Gram-positive bacteria.

Key words: innate immunity, *Drosophila*, serine proteases, danger signal