2008

Structure-function Analysis Of The Drosophila Stubble Type Ii Transmembrane Serine Protease

Rachel Morgan
University of Central Florida

STRUCTURE-FUNCTION ANALYSIS OF THE *DROSOPHILA* STUBBLE TYPE II TRANSMEMBRANE SERINE PROTEASE

by

RACHEL ELIZABETH MORGAN
B.S. Erskine College, 2005
B.A. Erskine College, 2005

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2008

ABSTRACT

Hormonally-triggered regulatory hierarchies play a major role in organismal development. Disruption of a single member of such a hierarchy can lead to irregular development and disease. Therefore, knowledge of the members involved and the mechanisms controlling signaling through such pathways is of great importance in understanding how resulting developmental defects occur.

Type II transmembrane serine proteases (TTSPs) make up a family of cell surface-associated proteases that play important roles in the development and homeostasis of a number of mammalian tissues. Aberrant expression of TTSPs is linked to several human disorders, including deafness, heart and respiratory disease and cancer. However, the mechanism by which these proteases function remains unknown.

The ecdysone-responsive Stubble TTSP of *Drosophila* serves as a good model in which to study the functional mechanism of the TTSP family. The Stubble protease interacts with the intracellular Rho1 (RhoA) pathway to control epithelial development in imaginal discs. The Rho1 signaling pathway regulates cellular behavior via control of gene expression and actin cytoskeletal dynamics. However, the mechanism by which the Stubble protease interacts with the Rho1 pathway to control epithelial development, in particular leg imaginal disc morphogenesis, has yet to be elucidated.

The Stubble protein consists of several conserved domains. One approach to a better understanding of the mechanism of action of Stubble in regulating Rho1 signaling is to define which of the conserved domains within the protease are required for proper function. Sequence

analysis of twelve recessive *Stubble* mutant alleles has revealed that the proteolytic domain is essential for proper function. Alleles containing mutations which disrupt regions of the protease domain necessary for protease activation or substrate binding, as well as those with deletions or truncations that remove some portion of the proteolytic domain, result in defective epithelial development *in vivo*. In contrast, mutations in other regions of the Stubble protein, including the disulfide-knotted and cytoplasmic domains, were not observed.

Another important step for defining the connection between Stubble and Rho1 signaling is to identify a Stubble target that acts as an upstream regulator of the Rho1 pathway. We performed a genetic screen in which 97 of the 147 *Drosophila* non-olfactory and non-gustatory G-protein-coupled receptors (GPCRs), a family of proteins that has been shown to be protease-activated and to activate Rho1 signaling, were tested for interactions with a mutant allele of *Stubble*. We found 4 genomic regions uncovering a total of 7 GPCRs that interact genetically when in heterozygous combination with a *Stubble* mutant. Further analysis of these genes is necessary to determine if any of these GPCRs is targeted by Stubble during activation of the Rho1 pathway.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Studies of human disease have typically been undertaken in cell culture and mammalian model systems. Although these systems are more amenable to genetic manipulation than humans themselves, they have significant limitations. Studies in cell culture suffer from the disadvantage that cells are usually not observed in a natural context with cell-cell adhesion contacts maintained. The extent to which the absence of cell adhesion influences normal cellular behavior is unknown. Mammalian model systems permit *in vi*vo manipulation of genes but can be very costly to perform.

Over the past 10 years, *Drosophila* has emerged as an excellent model in which to study human development and disease. Genetic studies can be performed *in vivo* and at a fraction of the cost of mammalian model systems. Major advantages of *Drosophila* are the sophisticated genetic tools that have been developed over the past 100 years, limited genetic redundancy allowing gene function to be readily determined, and the discovery of a high degree of conservation of gene function with humans. For example, studies in *Drosophila* of biochemical signaling pathways highly conserved between *Drosophila* and vertebrates has greatly increased understanding of the major signaling pathways central to regulation of metazoan development and homeostasis (Wassarman et al., 1995; Cadigan and Nusse, 1997; Greenwald, 1998; Johnson and Scott, 1998). *Drosophila* models have also been developed toward identifying the mechanisms of a number of human neuropathologies, such as Alzheimer's, Parkinson's and Huntington's diseases (Jackson et al, 1998; Warrick et al, 1998; Bilen and Bonini, 2005). In

addition, many human genes known to play direct roles in cancers have highly conserved orthologs in *Drosophila* (Hariharan and Bilder, 2006).

Here I focus on a class of poorly characterized transmembrane serine proteases associated with a variety of human pathologies. I hypothesize that studies of these proteases in a *Drosophila* model will be important in establishing a direct link between the action of these proteases and human disease.

Vertebrate type II transmembrane serine proteases and pathology

Upon sequencing of vertebrate genomes, a surprising number of genes belonging to previously identified but sparsely populated protease families have been described. One of these families, the mammalian type II transmembrane serine protease (TTSP) family, has rapidly grown into a family of 21 membrane–associated proteases with known roles in development and homeostasis in a variety of mammalian tissues.

TTSPs have an atypical mode of insertion into the membrane (intracellular N-terminus and extracellular C-terminus; Singer, 1990). All members of this protease family have a similar structure, consisting of a short intracellular domain, an extracellular stem region and a highly conserved catalytic serine protease domain of the chymotrypsin (S1) fold subfamily (Hooper et al., 2001). Within the extracellular stem region, various conserved domains may be found (Szabo and Bugge, 2008). TTSPs have been divided into four subfamilies based on a phylogenetic analysis of the serine protease domain as well as domain structure of the extracellular stem region and chromosomal location of their genes (Figure 1; Szabo et al., 2003). The largest is the human airway trypsin (HAT)/differentially expressed in squamous cell carcinoma (DESC)

subfamily, followed by the hepsin/TMPRSS/enteropeptidase subfamily, the matriptase subfamily, and the corin subfamily.

The catalytic activity of TTSPs is dependent on the presence of three residues in the proteolytic domain, histidine, aspartate, and serine, which make up the catalytic triad. Enzymatic activity is further modulated by the structural organization of the substrate binding pocket which is responsible for conferring substrate specificity. All TTSPs contain a highly conserved activation motif preceding the catalytic domain, suggesting that they are synthesized as single chain zymogens and must be activated by proteolytic cleavage after a lysine or arginine residue. Upon activation, the catalytic domain remains covalently linked to the membrane-bound stem region, and thus to the cell surface, by a conserved disulfide bond (Hooper et al., 2001). The specific mechanism of activation of individual TTSPs is unknown. However, all members of the TTSP family are trypsin-like, containing Gly216 and Gly226 (chymotrysinogen numbering) at the top of the S1 pocket, or primary specificity site, and Asp189 at the bottom of the pocket. TTSPs should, therefore, exhibit strong preference for substrates containing an Arg or Lys residue in the P1 position (Perona and Craik, 1997), and indeed, all tested members of the TTSP family exhibit this Arg/Lys specificity (Takeuchi et al., 2000; Velasco et al., 2002; Hobson et al., 2004; Szabo et al., 2005). These data suggest that activation may occur upon cleavage by other TTSPs or by autoactivation.

Figure 1. Predicted domain structures of mammalian type II transmembrane serine proteases.

Numbers indicate the position of each domain in the amino acid sequence of the zymogen. Figure and legend taken in modified form from Szabo and Bugge, 2008.

The TTSP family is of particular clinical interest because misexpression of several of its members has been linked to a number of pathologies. At least 15 different mutations in the *TMPRSS3* gene are associated with non-syndromic congenital and childhood onset autosomal recessive deafness in human patients (Ben-Yosef et al., 2001; Masmoudi et al., 2001; Scott et al., 2001; Wattenhofer et al., 2001; Ahmed et al., 2004; Wattenhofer et al., 2005; Elbracht et al., 2007; Guipponi et al., 2008). Interestingly, mice deficient for another TTSP, hepsin, also display hearing loss and developmental defects in the inner ear (Guipponi et al., 2007), suggesting that several TTSPs may play a cooperative role in normal inner ear development.

Several TTSP family members have been implicated in cancer. Hepsin is consistently misexpressed and the most upregulated gene in prostate cancer, expressed solely on the surface of prostate tumor cells (Dhanasekaran et al., 2001; Luo et al., 2001; Magee et al., 2001; Stamey et al., 2001; Welsh et al., 2001). Upregulation of hepsin *in vivo* causes disruption and disorganization of the prostate basement membrane and leads to metastasis and cancer progression in mice (Klezovitch et al., 2004). Misexpression of another TTSP, TMPRSS2, is also associated with prostate cancer progression. One study identified two genetic abnormalities in a subset of prostate cancer patients. Approximately 38% of patients showed a duplication of the *TMPRSS2* gene, while a single individual with aggressive prostate cancer showed a 7-bp deletion which resulted in a frameshift mutation and truncation of the protein after incorporation of only one catalytic triad residue (Vaarala et al., 2001). Other studies have shown that a majority of prostate cancer patients exhibit chromosomal rearrangements leading to fusion of the *TMPRSS2* gene with the *ERG*, *ETV1*, and *ETV4* genes which encode oncogenic transcription factors of the ETS family (Soller et al., 2006; Tomlins et al., 2005, 2006). *TMPRSS2* is an androgen-responsive

gene (Lin et al., 1999; Jacquinet et al., 2001), thus fusions with *TMPRSS2* result in androgen-dependent regulation of ETS transcription factors and subsequent overexpression in androgen-sensitive prostate tissue.

Misexpression of matriptase, was first noted in breast carcinoma (Shi et al., 1993; Lin et al, 1997) and has since been described in benign and malignant epithelial tumors of diverse origins (reviewed in Bugge et al., 2007). Matriptase is expressed in a variety of embryonic and adult tissues, but its expression is limited to the epithelial compartments of these tissues (Takeuchi et al., 1999; Oberst et al., 2003; List et al., 2006; Netzel-Arnett et al., 2006; Fan et al., 2007; Szabo et al., 2007). Tight regulation of matriptase activity in these tissues is essential for normal epithelial development. As shown by List et al. (2005), even a modest overexpression in the epidermis of transgenic mice leads to induction of spontaneous squamous cell carcinomas. In addition to an association with epithelial cancers, misexpression of matriptase is linked to additional pathologies involving the skin and hair, further strengthening the necessity for proper regulation of matriptase activity during development. Matriptase-deficient mice develop to term but die shortly after birth due to severe dehydration resulting from greatly impaired epidermal barrier function. These mice also display absence of whiskers, hair follicle hypoplasia, and accelerated apoptosis of immature T cells in the thymus (List et al., 2002). Recently, Basel-Vanagaite et al. (2007) reported a single amino acid substitution (G827R) in a highly conserved glycine hinge residue in the protease domain of human *ST-14*, which encodes matriptase. This mutation is found in patients with ichthyosis with hyptrichosis (ARIH) syndrome, a rare autosomal recessive form of skin disease. Patients present with mild to moderate ichthyosis,

indicative of impaired barrier function, and follicular hypoplasia characterized by sparse, fragile, brittle, dry and lusterless hair that shows slow growth.

Finally, the corin subfamily of TTSPs has been linked to heart development and cardiovascular disease. A minor allele of corin, a TTSP confined to the cardiac myocytes of the atrium and ventricle (Yan et al., 1999; Hooper et al., 2000), was recently associated with a human population with high systolic blood pressure and an increased risk for chronic hypertension (Dries et al., 2005).

Because misexpression of TTSPs is associated with a variety of human pathologies, understanding the mechanism of action of this class of proteins is crucial for disease treatment and prevention. One approach to understanding the mechanism of action of TTSPs is to determine the relationship between the conserved domains contained within the proteases and *in vivo* function of the protease. It has been established that a protease domain with an intact catalytic triad is necessary for activation and subsequent catalytic activity (Oberst et al. 2003; Hammonds and Fristrom, 2006). The functions of the various conserved domains found in TTSPs are generally not well understood; however, for several proteases, a requirement for these domains has been demonstrated. For example, matriptase is a multi-domain TTSP containing a SEA (sperm protein, enterokinase, and agrin) domain, two tandem CUB (C1r/s, Uegf, and bone morphogenic protein-1) domains, and four tandem LDLRA (low density lipoprotein receptor class A) domains within the extracellular region (see Figure 1). A series of mutations made in each of these domains indicates that intact function of each domain type is essential for matriptase activation (Oberst et al., 2003). Upon mutation of a putative glycosylation site in the first CUB domain, a dramatic reduction of protease-inhibitor complex formation and

immunoreactivity with M69 mAb, an antibody which recognizes only the two-chain activated matriptase, was observed. This indicates that glycosylation of the first CUB domain is necessary for protease activation. The LDLRA domain contains a "calcium cage," an octahedral arrangement of a Ca^{2+} atom, which is involved in the binding of this domain to substrates. Point mutation of critical residues of the calcium cage previously shown to be necessary for substrate binding (Esser et al., 1988) was made in each of the four LDLRA domains of matriptase. For each domain separately and all four domains simultaneously, this mutation inhibited protease-inhibitor complex formation and M69 immunoreactivity, suggesting that all of the LDLRA domains are essential for activation. Interestingly, efficient matriptase activation was observed following deletion of all four LDLRA domains. This suggests that the LDLRA domains may serve dual functions as binding domains for activators of matriptase and as autoinhibitory domains when no activators are present. Finally, the SEA domain contains a consensus cleavage site which serves as a target for proteolytic processing. Cleavage at this site converts matriptase into a smaller form lacking the N-terminal and transmembrane regions. When N-terminal proteolytic processing is inhibited via mutation of the cleavage site, no protease activation occurs in *in vitro* assays, indicating that N-terminal cleavage is necessary for enzyme activation (Oberst et al., 2003).

TMPRSS3, another multi-domain protease, consists of a single LDLRA domain followed by a single SRCR (scavenger receptor cysteine rich) domain in its extracellular stem region (see Figure 1). Two missense mutations have been identified in the LDLRA domain. One of these mutations affects a highly conserved residue and is believed to impair the calcium binding site of the domain (Wattenhofer et al., 2002). The second mutation affects the last residue of the

LDLRA domain, which is potentially involved in binding with extracellular molecules (Ben-Yosef et al., 2001). These mutations result in almost complete loss of protease activity *in vitro* and are associated with hearing loss (Guipponi et al., 2002). In addition, two missense mutations in highly conserved residues of the TMPRSS3 SRCR domain have been identified in patients with congenital and childhood onset hearing loss (Wattenhofer et al., 2002; Hutchin et al., 2005). One of these SRCR mutations also results in almost complete loss of protease activity in *in vitro* assays (Guipponi et al., 2002).

It is important to note that in most cases the linkage between mutant TTSP alleles and human disease has not been directly established. Thus, an assay system that could determine the activity of mutant TTSPs *in vivo*, and provide functional genetic rescue capability would be of great benefit to this field. In the section on future studies (see discussion) I describe how the *Drosophila* model system used in these studies could be utilized to this end.

The *Drosophila* Stubble TTSP as a model to study the mechanism of action of the TTSP family

The only non-vertebrate TTSPs that have been characterized are the *Drosophila* Stubble and corin proteases and their orthologues in other arthropod species (Beaton et al., 1988; Appel et al., 1993; Irving et al., 2001; Zou et al., 2006). The *Stubble* locus encodes a typical member of TTSP family and is required for normal leg, wing and bristle morphogenesis. Based on sequence analysis of the protease domain, the Stubble protease shows greatest sequence similarity to the matriptase subfamily in vertebrates. The Stubble protease ends in a trypsin-like serine protease domain of 244 amino acids and contains a 35-residue disulfide knotted domain, or clip domain,

specific to arthropods, within its extracellular stem (Figure 2; Appel et al., 1993). The cytoplasmic domain is 58 amino acids in length and contains one putative protein kinase C (PKC) phosphorylation site, one putative cAMP phosphorylation site, and 3 putative N-myristoylation sites. The PKC phosphorylation consensus pattern, SRR, where serine is the phosphorylation site, is found between residues 49 and 51 in the Stubble protein. The cAMP phosphorylation site has the consensus pattern RRsT and is found between residues 14 and 17. The consensus patterns for N-myristoylation present in Stubble are GSrgSG at residues 38-43, GSggAA at residues 41-46 and GGaaAR at residues 43-48. The N-terminus of the Stubble protein is hydrophilic, preceding a hydrophobic transmembrane anchor sequence. The transmembrane domain is flanked by charged regions (positive intracellular and negative extracellular) that are required for correct orientation of the protein in the membrane (see figure 2; Singer, 1990).

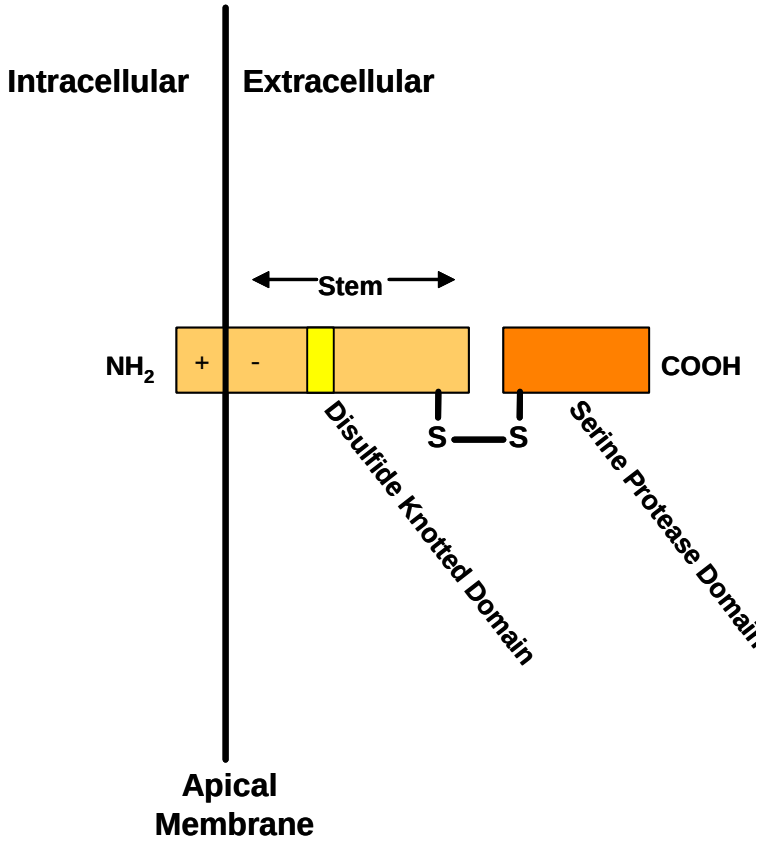

Figure 2. Structure of the active Stubble protease.

The Stubble protease is composed of four major regions: the intracellular (cytoplasmic) domain, the transmembrane domain, the stem region, and the proteolytic domain. A disulfide knotted domain, specific to arthropods, is found within the stem region.

The disulfide knotted domain, the only conserved domain contained within the extracellular stem of the Stubble protease, is connected to the protease domain by a linking sequence of 329 amino acids, substantially longer than the 23 to 101 amino acid residues found in most other clip domain-containing proteins (Jiang and Kanost, 1999). Disulfide knotted domains are held together by a series of three disulfide bonds which result in a compactly folded structure which can be schematically drawn in the shape of a paper clip, hence the alternative

designation of clip domain (Iwanaga et al., 1998). In addition to disulfide bridges formed within the clip and proteolytic domains an additional disulfide bond links the proteolytic and stem regions in the activated protease ensuring that the protease remains tethered to the membrane upon activation cleavage (Figure 3; Muta et al., 1990, 1993). Several functions have been proposed for disulfide knotted domains. They may function as mediators of interactions between proteases and their activators or substrates. It is also possible that disulfide knotted domains act as inhibitors of the protease domains to which they are linked. It is known that certain proteases, such as the cysteine protease cathepsin L and convertases of the serine protease subtilisin family, can be inhibited by their own propeptides (Carmona et al., 1996; Boudreault et al., 1998). Supporting this idea is the fact that disulfide knotted domains are small structures stabilized by disulfide bonds, common features of many canonical serine-type protease inhibitors (Bode and Huber, 1992). However, mutational analysis of the cysteine residues in the disulfide knotted domain of snake, a serine protease involved in dorsoventral patterning in *Drosophila*, argues against an autoinhibitory role (Smith et al., 1993). Alteration of five of the six individual cysteine residues to serine resulted in a failure of the protease to phenotypically rescue *snake*-deficient embryos. The remaining cysteine residue exhibited some phenotypic rescue upon alteration to serine, but activity was greatly reduced from wild type. Simultaneous alteration of cysteine pairs involved in disulfide bridge formation as well as deletion of the entire disulfide knotted domain also resulted in a failure to rescue *snake*-deficient animals. These data suggest that conserved disruption or removal of the disulfide knotted domain leads to a loss of snake activity. The opposite effect, constitutive activation, would be expected if the disulfide knotted domain acted as an inhibitor of the snake protease.

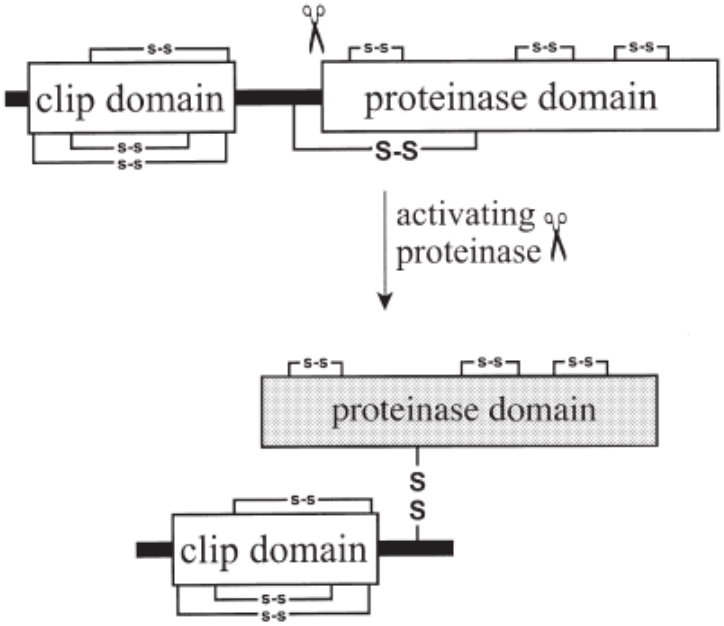

Figure 3. Structural organization of disulfide knotted (clip) domain proteases.

Clip domain proteases contain a linker region of variable length connecting the clip domain to a carboxy-terminal serine protease domain. A disulfide bond between conserved cysteine residues in the stem and proteolytic domains ensures that these domains remain covalently attached following activation of the zymogen. Figure and legend taken from Jiang and Kanost, 2000.

In this study I have chosen to use the *Drosophila* Stubble protease as a model to study TTSP action in epithelia. The *Stubble* locus is transcriptionally induced in response to the steroid hormone 20-hydroxyecdysone, a major mediator of morphogenetic events (Appel et al., 1993). Upon induction by ecdysone, the Stubble protein is localized to the apical cell membrane of developing leg and wing imaginal discs where it functions to control epithelial morphogenesis (von Kalm et al., 1995; Hammonds and Fristrom 2006). Studies in our laboratory have shown that in leg imaginal discs Stubble acts upstream of the monomeric GTPase Rho1, a major regulator of the actin cytoskeleton, and that this interaction occurs via an outside-in signaling

mechanism (Bayer et al., 2003). Thus, further studies of the mechanism underlying the interaction between Stubble and Rho1 are likely to improve our understanding of vertebrate TTSP action in development, homoestasis and disease.

In the following sections, I will describe in more detail the process of leg imaginal disc morphogenesis, the involvement of the Stubble protease in this process and the interaction between Stubble and Rho1 signaling in control of epithelial morphogenesis.

Leg imaginal disc development

In *Drosophila*, the steroid hormone 20-hydroxyecdysone (ecdysone) is necessary to mediate a major portion of the developmental process, metamorphosis. During metamorphosis, adult structures such as legs, eyes and wings arise from small epithelial organs called imaginal discs. Imaginal discs arise from invaginations of the embryonic epidermis (Cohen, 1993). During larval development, imaginal discs grow by mitotic division. Proliferation of the discs proceeds until just prior to metamorphosis, at which point the leg disc, for example, is simply a sac-like epithelial structure connected to the interior surface of the epidermis by a stalk where the original invagination occurred. This sac is composed of a thick folded columnar epithelium on one side, resembling a Danish pastry, and a flat squamous epithelium on the other. (Fristrom and Fristrom, 1993). At the end of larval development, a sharp increase in the titer of ecdysone occurs. This increase inhibits cell proliferation and triggers morphogenesis of the imaginal discs (Figure 4). It is during the subsequent 12-hour prepupal period of development that the leg and wing discs, for example, give rise to structures which resemble the adult appendages, and then evert to the outside of the body wall. The prepupal development of leg imaginal discs is depicted in Figure 5.

Following stimulation by the metamorphic pulse of hormone, leg discs elongate and evert to the outside of the animal. Elongation of the presumptive leg is driven by changes in apical cell shape leading to unfolding of the disc and shaping of the epithelium into a tube-like structure. Contraction of the actin-myosin contractile belt is the primary force underlying apical cell shape changes in elongating leg discs (von Kalm et al., 1995). When actin-myosin contraction is mis-regulated, apical cell shape changes are abnormal resulting in a malformed leg and wing phenotype, exemplified by legs that are shortened, thickened and twisted compared to normal legs, and wings that are crumpled and frequently blistered (Figure 6).

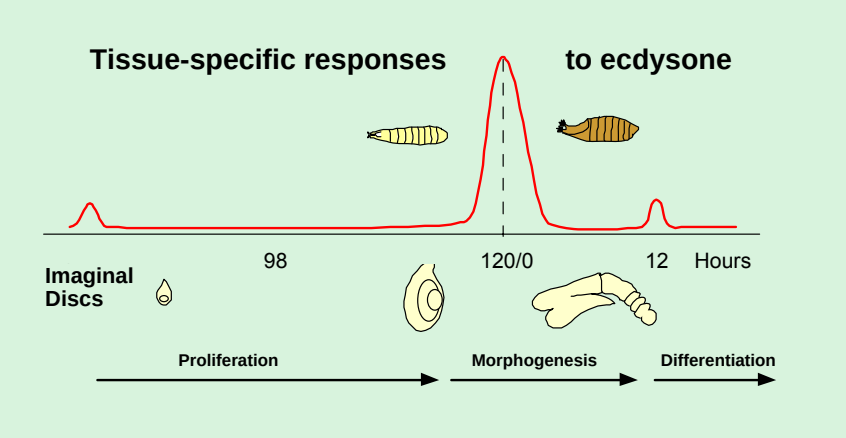

Figure 4. Fluctuations in ecdysone titer trigger imaginal disc morphogenesis at the onset of metamorphosis.

Proliferation of the leg imaginal discs occurs throughout larval development. The sharp rise in ecdysone titer (red line) occurring at 120 hours post egg laying inhibits further proliferation and triggers the elongation of the disc and subsequent eversion. Time 120h represents the end of larval development and the beginning of the 12 hour prepupal period and metamorphosis. At this point, the time line resets to 0h. A major phase of prepupal leg morphogenesis occurs during the first 6 hours of the prepupal period. Further morphogenesis and differentiation occur during the following 4-day pupal period.

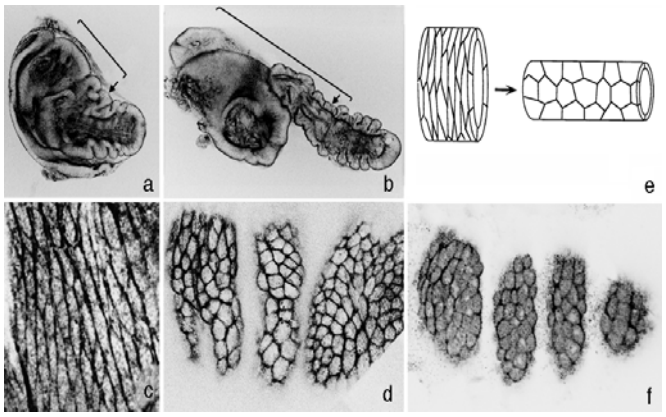

Figure 5. Cell shape changes in leg imaginal discs during elongation.

(a-d) Confocal micrographs of prepupal discs stained with phalloidin to visualize filamentous actin. (a) A leg at the beginning of the prepupal period. (b) A leg after 6 hours of prepupal development. (c and d) Apical view in cross-section indicated by arrows in a and b. Note the change in apical cell shape from anisometric to isometric in the elongating disc. (e) A schematic of cell shape changes that occur during elongation of the leg disc. (f) Apical view of a leg disc stained for Drosophila non-muscle myosin indicating the location of the actin-myosin contractile belt. (Photographs a-d from Condic et al., 1990; e-f from von Kalm et al., 1995).

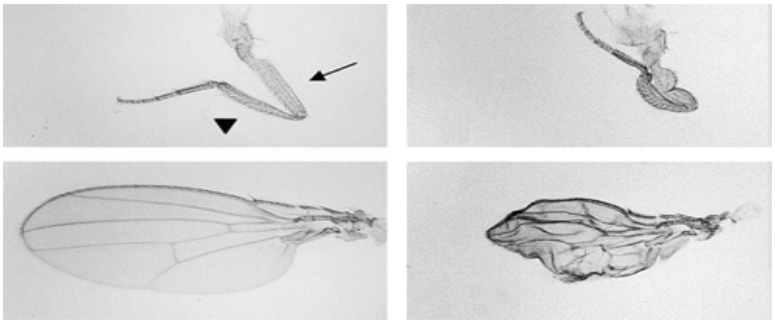

Figure 6. The malformed phenotype.

Defects in genes controlling cell shape changes result in malformed legs and wings (right panels) compared to wild-type (left panels) (from Bayer et al., 2003). Note that the mutant legs are shortened, thickened and twisted compared to normal legs. Wings are crumpled and frequently display blisters (separation of the wing epithelial bilayer).

Stubble function in normal development

The Stubble protein is expressed at the apical surface of leg imaginal disc cells, and multiple mutations in the *Stubble* locus have been identified which lead to the malformed leg phenotype. The malformed phenotype observed in *Stubble* mutant animals is caused by a failure of cell shape change in leg disc epithelia (Condic et al., 1991). The Stubble protease has also been linked to the Rho1 signaling pathway and control of actin cytoskeletal dynamics (Bayer et al., 2003; see below). Thus, a major role for the Stubble protease in leg disc development may be control of cell shape changes through regulation of the actin cytoskeleton.

Mutations in *Stubble* interact genetically with mutations in several members of the Rho1 signaling pathway in leg imaginal discs. Rho1 pathway members showing moderate to strong interaction with *Stubble* include *DRhoGEF2*, *Rho1*, *Rho kinase* (*drok*), *zipper* (non-muscle myosin II heavy chain), cofilin phosphatase, and myosin phosphatase (shown in red in Figure 7). In addition to genetic interaction data, leg malformation associated with overexpression of Stubble during leg morphogenesis can be suppressed by reducing *Rho1* gene dosage (Bayer et al., 2003). These data suggest that Stubble acts upstream of the Rho1 pathway and may be required for temporal and spatial activation of Rho1 signaling during leg disc elongation.

The Rho family of small guanosine triphosphates (GTPases) plays an important role in the regulation of epithelial morphogenesis. Rho1, Rac, and Cdc42 are three well-studied members of this family that act to regulate a variety of epithelial processes including but not limited to cell growth, cell division, cell adhesion and apical-basal and planar cell polarity. Regulation occurs by control of actin cytoskeletal dynamics and nuclear gene transcription through nucleation, polymerization, and depolymerization of actin filaments, and by control of

the actin-myosin contractile apparatus. Rho1, Rac, and Cdc42 regulate actin dynamics in part through involvement in intracellular signaling complexes in vertebrate and invertebrate epithelia (reviewed in Van Aelst and Symons, 2002).

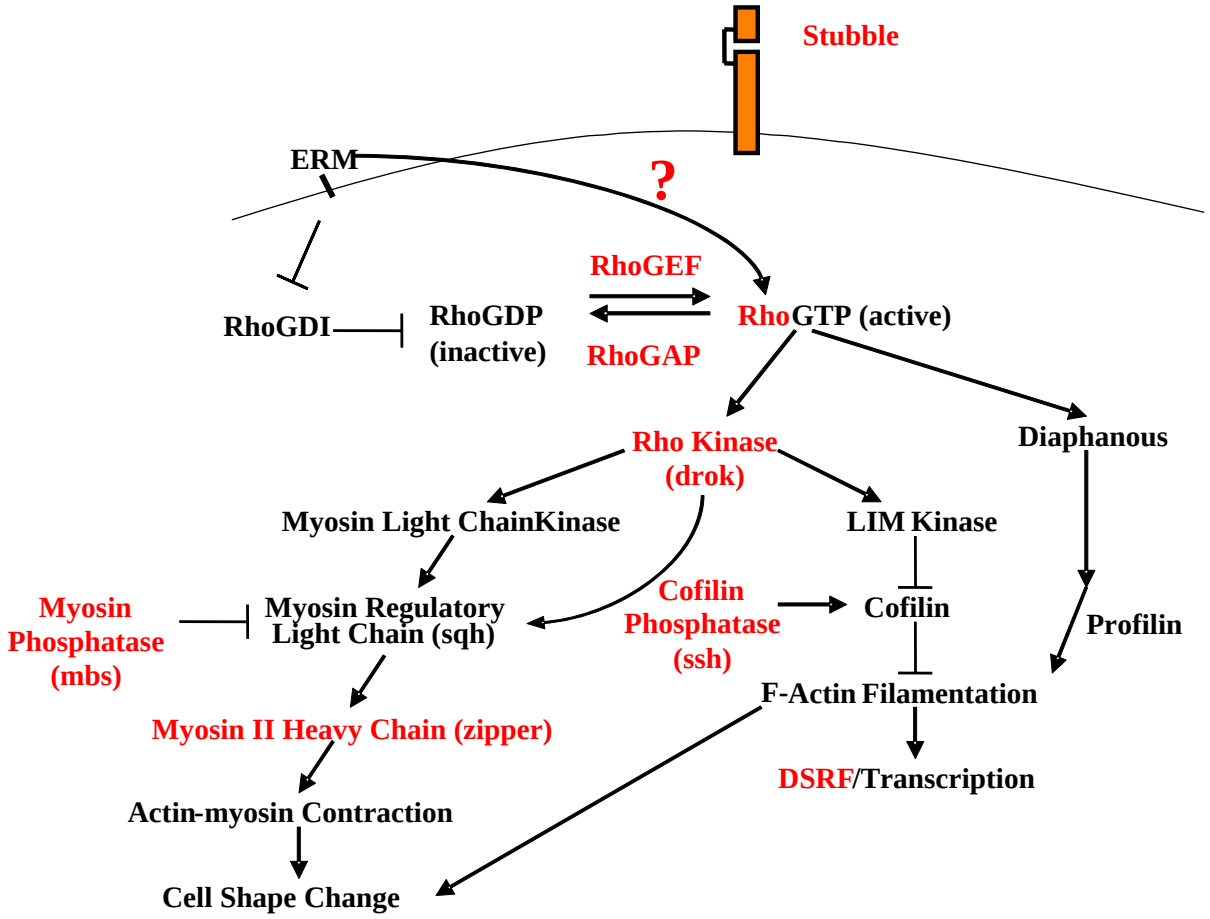

Figure 7. Rho1 regulation of actin contraction and filamentation. See text for details.

Drosophila orthologs of the vertebrate Rho kinase (drok), myosin regulatory light chain (sqh; Spaghetti Squash), non-muscle myosin II heavy chain (zipper), myosin phosphatase (mbs; myosin binding subunit), and cofilin phosphatase (ssh; slingshot) are indicated. Mutations in Rho1 pathway genes that interact genetically with Stubble mutations in leg imaginal discs are indicated in red.

Studies in vertebrate systems have provided a model for Rho1 signaling which leads to contraction of the actin-myosin contractile belt in epithelia (Figure 7). While only one Rho protein, Rho1, is present in *Drosophila*, three equivalent proteins, RhoA, RhoB and RhoC are present in vertebrates. Functionally, Rho1 is believed to be most similar to the vertebrate RhoA, though at the amino acid level, percent identity between Rho1 and human RhoA, B and C is identical (Hariharan et al., 1995). Rho1-specific guanine nucleotide exchange factors (RhoGEFs) are activated by extracellular signals that trigger a relocation of the GEF to the plasma membrane where it forms a complex with its target Rho1 GTPase (Zheng et al., 1996). Activated RhoGEF promotes cycling of the inactive form of Rho1 (GDP-bound) to the active form (GTP-bound). ERM (ezrin, radixin, and moesin) family proteins on the plasma membrane inhibit the association of guanine dissociation inhibitors (GDIs) with the inactive Rho1, thus reinforcing Rho1 activation through a positive feedback loop (Tsukita et al., 1997; Matsui et al., 1998). ERM proteins may also be involved in a negative feedback loop regulating Rho1 activity, thus fine tuning the rate of cycling of Rho1 between the GDP- and GTP-bound states (Speck et al., 2003). Activated Rho1 then stimulates the downstream effectors Rho kinase and Diaphanous. Activated Rho kinase stimulates at least two different substrates, splitting the pathway into a myosin-dependent branch and a Lim kinase-dependent branch. In the myosin-dependent branch, Rho kinase phosphorylates and thus activates myosin regulatory light chain (Sqh) (Amano et al., 1996) and inhibits myosin light chain phosphatase (Kimura et al., 1996). Activation of the myosin regulatory light chain leads to activation of myosin II heavy chain (Zipper) and contraction of the actin-myosin belt. In the Lim kinase-dependent branch of the pathway, Rho kinase activates Lim kinase, which inhibits the turnover of F-actin by cofilin, leading to

stabilization and accumulation of F-actin and subsequent cell shape change (Maekawa et al., 1999). Activated Diaphanous must cooperate functionally with Lim kinase to induce accumulation of F-actin (Geneste, Copeland and Treisman, 2002). While Rho kinase-Lim kinase signaling inhibits F-actin turnover, Rho1 promotes F-actin assembly by binding to and thus relieving autoinhibition of Diaphanous, which, in conjunction with profilin, nucleates actin filaments resulting in formation of unbranched F-actin filaments (Watanabe et al., 1999; Pruyne et al., 2002; Sagot et al., 2002).

Rho signaling and disease

Studies have recently shown Rho family proteins to be involved in tumorigenesis. The mechanisms by which these proteins are involved in cancer progression have not been defined but the links existing between Rho proteins and cancer are considerable. Misexpression of members of the RhoA signaling pathway in vertebrates has been associated with a variety of cancers of epithelial origin (reviewed in Benitah et al., 2004). Given that TTSPs are thought to be involved in disease progression via interaction with intracellular signaling pathways and that the RhoA pathway in vertebrates has been linked to disease, the interaction of the Stubble protease with the Rho1 signaling pathway in *Drosophila* serves as an excellent model in which to study the mechanistic role of TTSPs in disease.

G protein-coupled receptors are a potential link between Stubble proteolytic activity and Rho1 signaling

The genetic data presented above strongly implicate Stubble in Rho1 signaling. Interactions have been detected between mutations in the *Stubble* gene and *DRhoGEF2*, *Rho1*,

drok, cofilin phosphatase, myosin phosphatase, *zipper* and *DSRF* (shown in red in Figure 7). Also, a reduction in the gene dosage of *Rho1* suppresses the severe malformed leg phenotype induced by overexpression of Stubble (Bayer et al., 2003). As a result of these studies, it has been proposed that the Stubble protease operates through an "outside-in" signaling mechanism during leg development to either act in parallel with or directly regulate Rho1 signaling, though the mechanism by which this occurs is unknown.

In an attempt to better understand the mechanism by which Stubble acts, a genome-wide screen for genes that interact with Stubble to control leg development was initiated (von Kalm et al., unpublished). Using a deletion strategy, 80Mb (85%) of the *Drosophila* autosomal genome was screened for chromosomal intervals containing genes whose products interact with Stubble during epithelial morphogenesis to regulate cell morphology. Twelve intervals containing interacting gene products were identified. Six of these intervals uncover heterotrimeric guanine nucleotide-binding protein (G protein)-coupled receptors (GPCRs), a family of proteins through which Stubble could potentially link to intracellular Rho1-mediated signaling events.

In addition to Stubble, four vertebrate type II transmembrane serine proteases also appear to operate through an "outside-in" signaling mechanism. Enteropeptidase, HAT (human airway trypsin), Matriptase/MT-SP1, and TMPRSS2 have all been linked to intracellular signaling events which involve PAR (protease-activated receptor)-2 (Takeuchi et al., 2000; Cottrell et al., 2004; Chokki et al., 2005; Wilson et al., 2005). PARs are a subset of G protein-coupled receptors that are activated by proteolytic cleavage of the receptor, creating a tethered ligand that cannot diffuse away, rather than by simple ligand occupancy (Vu et al., 1991). Proteolytic activation of PARs is mediated by the serine protease family. PARs are expressed in a number of normal cell

types as well as in tumor cells and invasive cell lines and are found in carcinoma specimens. PAR-2, specifically, is expressed in a number of cell types including vascular endothelial cells, myocytes, fibroblasts, immune cells, neurons, glial cells and a variety of epithelial cells. The physiological role of the PARs in different cell types includes, but is not limited to, control of homeostasis, inflammation, pain and healing, and proliferation (MacFarlane et al., 2001).

G protein-coupled receptors, which are able to activate heterotrimeric G proteins, are known to be major upstream regulators of Rho signaling. G proteins couple activated GPCRs to the initiation of intracellular signaling responses by acting on downstream effector molecules. Several Rho-specific GEFs have been shown to act as effectors of various G protein α subunits in mammalian systems, thus coupling activation of Rho signaling with G protein signaling events. A number of distinct G protein α subunits have been identified and are divided into four families based on amino acid sequence similarity and function: G_s, G_i, G_q and G_{12} (Wilkie et al., 1992). PDZ-RhoGEF, p115 RhoGEF and leukemia-associated Rho guanine nucleotide exchange factor (LARG) interact with the G_{12} family subunits $G\alpha_{12}$ and $G\alpha_{13}$ through their regulator of G protein signaling (RGS) domains (Hart et al., 1998; Kozasa et al., 1998; Fukuhara et al., 1999, 2000). LARG has also been shown to cooperate with activated $G\alpha_{q/11}$, as well as $G\alpha_{12/13}$, to stimulate activation of RhoA, suggesting that LARG can also serve as an effector for G_q coupled receptors (Booden et al, 2002; Vogt et al., 2003). More recently, p63RhoGEF was shown to couple $G_{q/11}$- but not $G_{12/13}$-coupled receptors to Rho signaling (Lutz et al., 2005, 2007). It was further demonstrated that the p63RhoGEF-induced enhancement of RhoA activation occurred by direct protein-protein interaction between the C-terminus of p63RhoGEF and active $G\alpha_q$ or $G\alpha_{11}$.

Aims of this study

Localization to the cell surface puts type II transmembrane serine proteases in an excellent position to mediate signal transduction between intracellular signaling pathways and the extracellular environment and to regulate cellular responses. Indeed, a number of TTSPs have been shown to play important roles in the normal development of tissues and have been associated with a variety of human diseases, including heart disease and cancer. To understand the distinct roles that TTSPs play in tissue development and disease, it is necessary to define their function on a mechanistic level. Two approaches to a better understanding of the mechanism of action of TTSPs in regulating cellular response are to define which of the conserved domains within the proteases are required for proper function and to identify TTSP targets that could act as upstream regulators of intracellular signaling and response pathways. Because the Stubble TTSP is known to play a role in epithelial morphogenesis through interaction with the intracellular Rho1 signaling pathway, it serves as an excellent model in which to study the mechanism of action of the TTSP family.

In an attempt to identify the essential conserved domains of the Stubble protein, I sequenced twelve recessive mutant alleles of *Stubble*. Coding region mutations were identified in nine of these alleles and all but one localize to the protease domain. The remaining mutation localizes to the stem region and affects the amino acid residue directly preceding the disulfide knotted domain. Mutations in other regions of the Stubble protein, including the disulfide-knotted and cytoplasmic domains, were not observed. These data identify the protease domain as essential for function but do not rule out the remaining domains as necessary functional components.

Because G protein-coupled receptors have been identified as TTSP targets and have also been linked to the Rho1 pathway via activation of G protein α-subunits, I asked if GPCRs might be targeted by Stubble to regulate Rho1 signaling. These experiments involved a genetic screen of the non-gustatory and non-odorant GPCRs in the *Drosophila* genome for interactions with a mutant *Stubble* allele. At present, 97 of the 147 total non-gustatory, non-odorant receptors have been tested, and 4 genomic regions which interact genetically with the *Stubble* mutant have been identified. These regions uncover a combined total of 8 GPCRs. Further analysis is necessary to determine if any of these GPCRs are targeted by Stubble leading to activation of the Rho1 pathway.

CHAPTER TWO: METHODS

Drosophila stocks

Drosophila *sbd*, *Rho1*, and *zipper* mutant alleles used in this study are described in Table 1. All $sbd^{E(br)}$ stocks were obtained from R. Ward (University of Kansas; Ward *et al.* 2003). All deficiency stocks used in this study were obtained from the Bloomington Drosophila Stock Center at Indiana University (Bloomington, IN). Transgenic stocks are described below. All stocks were maintained at 25º on standard cornmeal/yeast/agar medium.

Genetic complementation analysis

To genetically test the activity of the *sbd* mutants, second-site noncomplementation (SSNC) assays were performed by mating animals heterozygous for *sbd* mutations with animals heterozygous for other mutant genes of interest. Leg malformation was scored in the doubly heterozygous F1 progeny class (*i.e., */+; sbd/+*). Genotypes generated to test for dominant interactions (*i.e., $Rho1^{72O}$/+; sbd*/sbd^{1}*) were obtained by mating animals carrying *sbd* mutations, *e.g., sbd*/TM6B, Tb Hu e*, to *$Rho1^{72O}$/CyO-CR2*, *P{sevRas1.V12}FK1; sbd^{1}/TM6B, Tb Hu e* animals. In each cross, 7-10 virgin females were mated with 5 males. All crosses were set up in duplicate. Cultures were incubated at 25º for either 3 or 5 days followed by transfer of adults to fresh medium. Subsequent cultures were incubated for 2 or 5 days, respectively, until a total of 6 cultures per cross were generated.

Table 1. Mutant alleles of stocks used in this study

Mutant Allele	Mutagen	Reference
sbd$^{E(br)20}$	EMS	Ward *et al*. (2003)
sbd$^{E(br)48}$	EMS	Ward *et al*. (2003)
sbd$^{E(br)228}$	EMS	Ward *et al*. (2003)
sbd$^{E(br)448}$	EMS	Ward *et al*. (2003)
sbd$^{E(br)536}$	EMS	Ward *et al*. (2003)
sbd$^{E(br)623}$	EMS	Ward *et al*. (2003)
sbd^{PNR11}	EMS	Heitzler and Simpson (1991)
sbd^{VE3}	EMS	Heitzler and Simpson (1991)
sbd^{1}	Spontaneous	Dobzhansky (1930)
sbd^{206}	EMS	Hecht (1989)
sbd^{46}	EMS	R. Ruggiero (unpublished data)
sbd^{173}	EMS	R. Ruggiero (unpublished data)
sbd^{241}	EMS	R. Ruggiero (unpublished data)
sbd^{258}	EMS	R. Ruggiero (unpublished data)
sbd^{266}	EMS	R. Ruggiero (unpublished data)
sbd^{277}	EMS	R. Ruggiero (unpublished data)
Df(3R)sbd^{105}	X ray	Beaton *et al*. (1988)
Rho1$^{J3.8}$	EMS	Halsell *et al*. (2000)
Rho1$^{E3.10}$	EMS	Halsell *et al*. (2000)
Rho1^{72O}	*P*-element insertion/excision	Strutt *et al*. (1997)
Rho1^{k02107b}	*P*-element insertion	Magie *et al*. (1999)
zip^{Ebr}	EMS	Gotwals and Fristrom (1991); Gotwals (1992); Halsell *et al*. (2000)

To screen for G-protein-coupled receptors (GPCRs) potentially involved in leg morphogenesis, SSNC assays were performed by mating heterozygous *Stubble* mutant animals (*i.e., red Sb*^{63b}*e/TM6B, Tb Hu e*) with heterozygous or homozygous animals carrying a deficiency uncovering the GPCR, or a transposable element insertion or mutation in the GPCR. Leg malformation was scored in the doubly heterozygous F_1 progeny class (*i.e., */+; Sb/+*). Secondary SSNC assays of mutants found to interact in the primary screen were conducted by

mating animals carrying interacting mutations with animals heterozygous for other mutant alleles of interest (*i.e., Sb^{70}, $Rho1^{E3.10}$, $Rho1^{J3.8}$, Np^{125A}* and *Np^{2}*) and by mating heterozygous *Stubble* mutant animals (*i.e., red Sb^{63b}e/TM6B, Tb Hu e*) with animals carrying deficiencies overlapping the region uncovered by the original interacting deficiency. All crosses were set up in duplicate. Cultures were incubated at 25º for 5 days followed by transfer of adults to fresh medium. Subsequent cultures were incubated for 5 days, until a total of 6 cultures per cross were generated.

Lethal-phase analysis of *sbd* mutant alleles

Embryonic lethality was assessed by collecting 0- to 2-hr embryos from *sbd/TM6B, P{w^{+}, UbiGFP}* stocks. Embryos were aged to 15 hours at 25ºC and dechorionated in 50% bleach for 3 minutes. Homozygous *sbd* embryos were then selected based on absence of green fluorescent protein (GFP) expression. Homozygous embryos were aged at 25ºC to 48 hours post-egg laying, and dead embryos were counted. Embryonic lethality was calculated as (number of dead embryos/total number of mutant embryos) x 100. Dependent on the number present, 34-100 *sbd* mutant embryos were assessed in each experiment, and each experiment was done in triplicate. Larval lethality was determined by collecting homozygous mutant *sbd* first instar larvae derived from *sbd/TM6B, P{w^{+}, UbiGFP}* stocks on the basis of absence of GFP expression. Mutant larvae were allowed to develop for 7-10 days at 25ºC, and pupae were counted. Larval lethality was calculated as [(total number of mutant larvae – number of pupae)/total number of mutant larvae] x 100. Dependent on the number of hatchers, up to 100 larvae were assessed in each experiment, and each experiment was done in triplicate. Pupal

lethality was determined by allowing pupae from the larval lethality assay to eclose. Adults were counted and pupal lethality was calculated as [(total number of pupae – number of adults)/total number of pupae] x 100.

Genetic characterization of *sbd* alleles

To genetically characterize the *sbd* alleles, 150 virgin *sbd*^{105} *e ca/TM6B*, *P{w*$^{+}$*, UbiGFP}* females and 150 *sbd*/TM6B*, *P{w*$^{+}$*, UbiGFP}* females were mated separately to 75 *sbd*/TM6B*, *P{w*$^{+}$*, UbiGFP}* males. 0-2 hour embryos were collected on grape plates and aged to 15-17 hours at 25º, at which time they were dechorionated in 50% bleach for 3 minutes. Homozygous *sbd** and *trans*-heterozygous *sbd*^{105} *e ca/sbd** mutant embryos were then selected based on absence of green fluorescent protein (GFP) expression and placed on a fresh grape plate. Embryos were then aged at 25º to 48 hours post-egg-laying, at which time dead embryos were counted and hatched larvae were transferred to a fresh plate with a strip of thin yeast paste and stored at 25º in a humid chamber. Every 24 hours, dead larvae were counted, larval progress was observed and noted, and larvae were transferred to a fresh plate containing a strip of thin yeast paste. Upon reaching the third instar stage, larvae were placed in a vial containing *Drosophila* medium and aged at 25º for a period of time sufficient for pupariation and eclosion of any viable adults. Development and lethal-phase of homozygous *sbd** mutants was compared to that of *sbd*^{105} *e ca/sbd** mutant *trans*-heterozygotes. Any *sbd** allele exhibiting an identical developmental pattern and lethal-phase as *trans*-heterozygotes is considered to be a null allele.

Sequence analysis of *sbd* mutant alleles

To identify mutations in *sbd* alleles, genomic DNA from mutant animals in homozygous and various heterozygous conditions (Table 2) was isolated via phenol-chloroform extraction followed by ethanol precipitation. Twenty adults were collected for each mutant genotype analyzed and ground using a motorized pestle in a 1.5 mL microcentrifuge tube containing 200µl homogenizing buffer (100mM NaCl, 200mM sucrose, 100mM Tris-Cl, pH 9.1, 50mM EDTA, 0.5% SDS). Upon homogenization, an additional 500µl homogenizing buffer was added and used to rinse the pestle. Phenol-chloroform extraction was performed by adding 250µl phenol and 250µl chloroform (chloroform:isoamyl alcohol 24:1) and vortexing for 30 seconds, followed by centrifugation at 4º for 5 minutes at 14,000 rpm. After spinning, the aqueous phase was removed and a second phenol-chloroform extraction was performed as above. 500µl of the final aqueous phase was placed in a clean microcentrifuge tube and prepared by ethanol precipitation with 50µl 3M NaAc (pH 4.8-5.2) and 1mL cold 100% ethanol. Samples were mixed and spun at 4º for 5 minutes at 14,000 rpm. After centrifugation, the supernatant was decanted and pellets were dried in a vacuum aspirator for 15 minutes, followed by resuspension in 200µl TE. A second ethanol precipitation was performed using 20µl 3M NaAc (pH 4.8-5.2) and 400µl cold 100% ethanol. Samples were mixed and centrifuged at 4º for 5 minutes at 14,000 rpm. Supernatant was removed and pellets were washed with 400µl 70% ethanol and centrifuged at 4º for 5 minutes at 14,000 rpm. Supernatant was carefully removed by pipetting, and pellets were dried in a vacuum aspirator for 15 minutes, followed by resuspension in 25µl TE containing 2mg/mL RNase. For homozygous lethal genotypes, 50 homozygous first instar larvae derived

from *sbd/TM6B, P{w$^+$, UbiGFP}* stocks were collected based on absence of GFP expression, and DNA was prepared as described above.

The *Stubble* locus was divided into six regions (Figure 8) for amplification and sequence analysis. A series of intron- and exon-specific primers (Table 3) was used to amplify each region. Amplification of all regions was conducted using a reaction mix containing 1X reaction buffer (10mM Tris-HCl, pH 8.3, 50mM KCl), 10µM 3'-primer, 10µM 5'-primer, 0.8mM dNTPs, 1.5mM $MgCl_2$ and 1U/rxn Jump-Start Taq polymerase. Stock solutions of reaction buffer (10X), $MgCl_2$ (25mM) and Jump-Start Taq were obtained from Invitrogen (Carlsbad, CA). All reactions were done in a 20µl volume in 0.5mL thin-walled tubes in the UNOII Biometra Thermocycler. Thermal cycler settings used for amplification were the same for all primer pairs and are shown in Table 4. Amplification products were run on 1.0% Low EEO agarose, excised and purified using the QIAquick Gel Extraction kit (Qiagen, Valencia, CA). DNA concentrations were estimated by running 5ul of the purified product on 1.5% Low EEO agarose and comparing band intensity to the quantitative Hyperladder II (Bioline USA Inc., Taunton, MA). Purified amplification products and primers were sent to the Nevada Genomics Center for sequencing using the ABI BigDye Terminator Cycle Sequencing Ready Reaction Kit v3.1 and the ABI3730 DNA Analyzer. Mutations were verified by sequencing independent PCR products amplified from genomic DNA samples originating from different sets of animals of the same mutant genotypes. DNA sequence analysis was facilitated by the use of the Lasergene 6 DNA and protein analysis software (DNASTAR, Inc., Madison, WI).

Table 2. Genotype and regions of all individual mutant stocks used for sequence analysis

Genotype	*Sb* Regions Amplified and Sequenced					
	1	2	3	4	5	6
sbd^{46}/TM6B, Tb Hu e	x	x	x	x	x	x
w^{1118}/+; iso2/+; iso3/sbd^{46}			x	x		
sbd^{173}/TM6B, Tb Hu e	x	x	x	x	x	x
sbd^{173}/sbd^{173}			x	x		
sbd^{241}/TM6B, Tb Hu e	x	x	x	x	x	x
w^{1118}/+; iso2/+; iso3/sbd^{241}			x	x		
sbd^{258}/TM6B, Tb Hu e	x	x	x	x	x	x
w^{1118}/+; iso2/+; iso3/sbd^{258}			x	x		x
sbd^{266}/TM6B, Tb Hu e	x	x	x	x	x	x
w^{1118}/+; iso2/+; iso3/sbd^{266}		x	x	x	x	
sbd^{277}/TM6B, Tb Hu e	x	x	x	x	x	x
w^{1118}/+; iso2/+; iso3/sbd^{277}			x	x		
sbd$^{E(br)20}$/TM6B, Tb Hu e	x	x	x	x	x	x
sbd$^{E(br)20}$/sbd$^{E(br)228}$			x	x		x
sbd$^{E(br)48}$/TM6B, Tb Hu e	x	x	x	x	x	x
sbd$^{E(br)48}$/sbd$^{E(br)228}$			x	x		
sbd$^{E(br)228}$/TM6B, Tb Hu e	x	x	x	x	x	x
sbd$^{E(br)448}$/TM6B, Tb Hu e	x	x			x	x
sbd$^{E(br)448}$/sbd$^{E(br)228}$			x	x		x
sbd$^{E(br)536}$/TM6B, Tb Hu e	x	x	x	x	x	x
sbd$^{E(br)536}$/ sbd$^{E(br)228}$			x	x	x	
sbd$^{E(br)623}$/TM6B, Tb Hu e	x	x			x	x
sbd$^{E(br)623}$/sbd$^{E(br)228}$			x	x		
sbd$^{E(br)623}$/sbd$^{E(br)623}$			x			x
sbd^{1} st ro e ca/ sbd^{1} st ro e ca	x	x	x	x	x	x
red sbd^{201}e/TM6B, Tb Hu e	x	x	x	x	x	x
red sbd^{201}/ red sbd^{201}	x	x	x	x	x	x
sbd^{206} ru cu e ca/ TM6B, Tb Hu e	x	x	x	x	x	x
sbd^{206} ru cu e ca/ sbd^{206} ru cu e ca			x	x		
st sbd^{PNR11}e/ TM6B, Tb Hu e	x	x	x	x	x	x
st sbd^{PNR11}e/ st sbd^{PNR11}e			x	x		x
st sbd^{VE3}e/ TM6B, Tb Hu e	x	x	x	x	x	x
st sbd^{VE3}e/ st sbd^{VE3}e			x	x	x	
w^{1118}/ w^{1118}; iso2/iso2; iso3/iso3 (*isoline*)	x	x	x	x	x	x
br^{1}/ br^{1}	x	x	x	x	x	x

Regions 1–6 refer to the six regions into which the *Stubble* genetic locus was split for sequencing purposes (see Figure 8). A cross (x) denotes successful amplification and sequencing of the specified region for the given genotype.

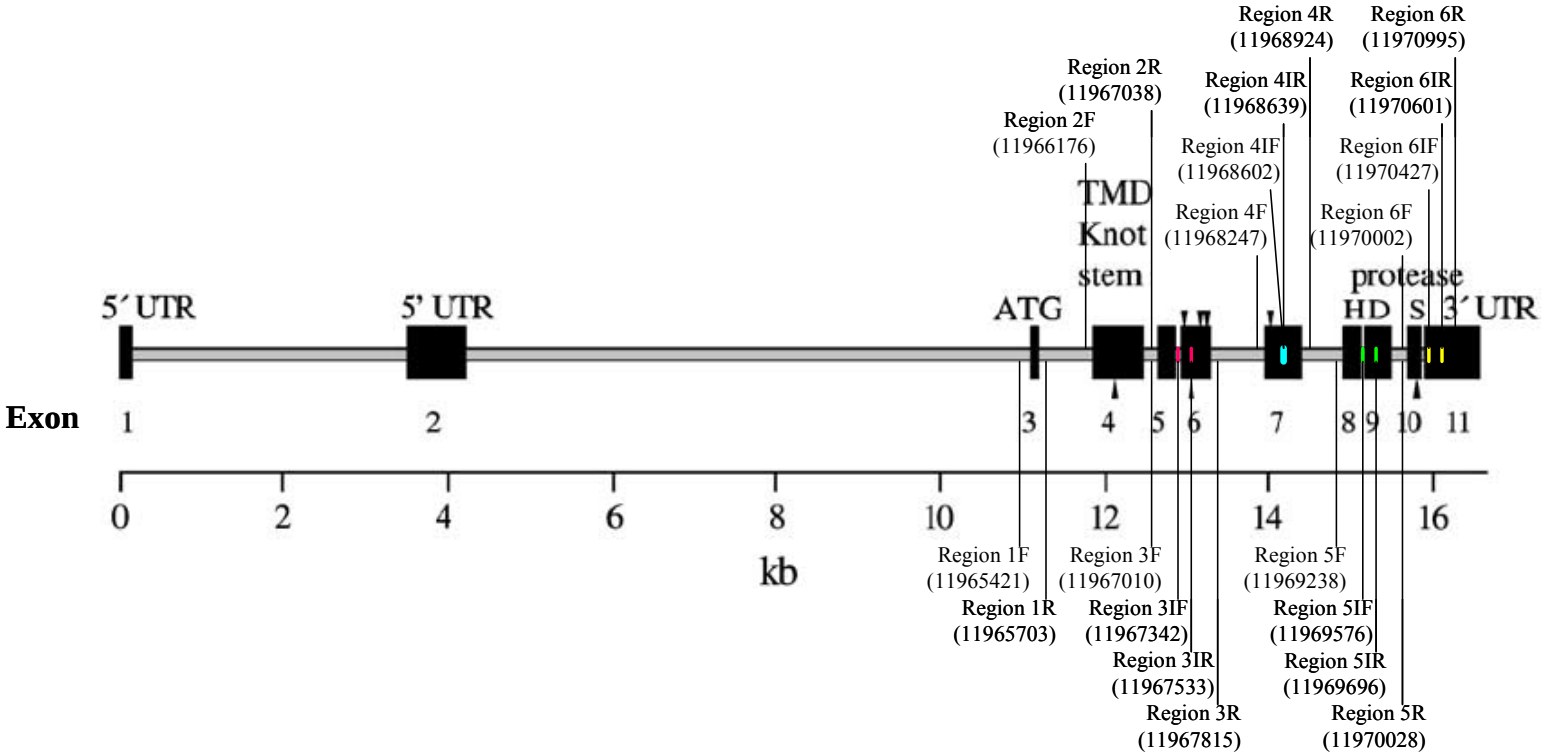

Figure 8. Stubble genomic structure and locations of primers used to sequence mutant sbd alleles.

Sizes of introns and exons are shown to scale and indicated in kilobases. Exons are shown as black boxes with exon number indicated below. Features from the cDNA (5'- and 3'-UTR, ATG translation start) are indicated above the exons in which they are found. The protease domain is split between the last four exons. Each of the residues forming the catalytic triad (H, D, S) is encoded in a separate exon, each indicated above the corresponding exon. The transmembrane domain (TMD), cysteine-knotted domain (knot) and the first part of the stem are in exon 4. For sequencing purposes, the coding domain of the *Stubble* locus was split into 6 regions. Forward (F) and reverse (R) primers for each region are indicated. Specific primer sequences are shown in Table 3. Numbers given with the primers denote the molecular position of its 5' end. Colored dashes within the exons mark the positions of the internal primers in regions 3 (red), 4 (blue), 5 (green), and 6 (yellow). Figure and legend are taken in modified form from Hammonds and Fristrom, 2006.

Table 3. Primers used to amplify and sequence the coding region of the Stubble locus

Primer Name	Alternate Name	Region Amplified	Exon(s) Amplified	5' Molecular Coordinate (3R)	Primer Sequence (5'→3')
5' Sb Exon 1(1)	Region 1F	1	3	11965421	CATTCTCGATGCCATTACAAACATA
3' Sb Exon 1(1)	Region 1R	1	3	11965703	AACTGGTAGCACACATCGCTGCCTCC
5' Sb Region 2	Region 2F	2	4	11966176	AAGTGCGCTTAATTAGTGGTGGACT
3' Sb Region 2	Region 2R	2	4	11967038	GATATATTATGCTATCTTTCAAGCCGCTT
5' Sb Region 3	Region 3F	3	5,6	11967010	AAGCGGCTTGAAAGATAGCATAATATATCT
3' Sb Region 3	Region 3R	3	5,6	11967815	GTTTGTTTCGCTTTCGACATCAAT
Sb Region 3F seq	Region 3IF	3 (internal)	5,6	11967342	GATGCATTGTTAGCATGA
Sb Region 3R seq	Region 3IR	3 (internal)	5,6	11967533	GCAAATTGACGATGCCGG
5' Sb Region 4	Region 4F	4	7	11968247	TTCTTAGTTAATGCAGATTTAGCGGC
3' Sb Region 4	Region 4R	4	7	11968924	CTAATAAGCTATCATCTTCAATGAGCAAGC
Sb Region 4F seq	Region 4IF	4 (internal)	7	11968602	CTCCAGCTCGAGCACCTC
Sb Region 4R seq	Region 4IR	4 (internal)	7	11968639	GAGTCTTGCTCGATGTGG
5' Sb Region 5	Region 5F	5	8,9	11969238	TCAATTTGTGAGGGAATTCGAGATACAGTT
3' Sb Region 5	Region 5R	5	8,9	11970028	TAATTGCAATCGCCCCATTTAGAACC
Sb Region 5F seq	Region 5IF	5 (internal)	8,9	11969576	GATCCAGATCGAGCTAAT
Sb Region 5R seq	Region 5IR	5 (internal)	8,9	11969696	GCTCTATGTAGGGCAGCT
5' Sb Region 6	Region 6F	6	10,11	11970002	TGGTTCTAAATGGGGCGATTGCAATTA
3' Sb Region 6	Region 6R	6	10,11	11970995	TATTGCTCTCGAGCTTTCGCATCGCTAT
Sb Region 6F seq	Region 6IF	6 (internal)	10,11	11970427	CGGAGTCTGTACGAGAAT
Sb Region 6R seq	Region 6IR	6 (internal)	10,11	11970601	TGAAGCTAACTACGCT

Table 4. Thermal cycling parameters for amplification of sbd alleles

Segment	Cycles	Temperature	Time
1	1	94ºC	3 minutes
2	31	94ºC – Denaturation	1 minute
		55ºC – Annealing	1 minute
		72ºC – Extension	1 minute
3	1	72ºC	15 minutes
4	1	25ºC	30 minutes

Site-directed mutagenesis

Site-directed mutagenesis was performed on a full-length *Stubble* cDNA including 838 nucleotides of 5'-(UTR) and 523 nucleotides of 3'-(UTR) cloned into the pLitmus29 vector (pL29 – Sbc10h7). Site-directed mutagenesis was done using the QuikChange II Site-Directed Mutagenesis Kit according the manufacturer's specifications (Stratagene, La Jolla, CA). Briefly, complimentary oligonucleotides containing the desired mutations were designed using Stratagene's web-based QuikChange Primer Design Program (http://www.stratagene.com/qcprimerdesign) and ordered through Sigma-Genosys with desalting for purification (Sigma-Aldrich Co., The Woodlands, TX) (Table 5). Further purification by FPLC or PAGE, as recommended in the QuikChange II protocol, was not performed. Mutagenesis reactions were performed by thermal cycling using *PfuUltra* High-Fidelity DNA polymerase. All reactions were done in a 50µl volume in 0.5mL thin-walled tubes in the UNOII Biometra Thermocycler. Cycling parameters are listed in Table 6. To make mutations analogous to *sbd* alleles, the oligonucleotides used were H573R which alters histidine 573 to arginine (CAC to CGC), P137L which alters proline 137 to leucine (CCA to CTA), and G760S which alters glycine 760 to serine (GGC to AGC). To make mutations analogous to human disease-associated

mutations, the oligonucleotides used were G760R which alters glycine 760 to arginine (GGC to CGC) and S549I which alters serine 549 to isoleucine (AGC to ATC). The serine at 549 was also converted to asparagine (AGC to AAC), as found in wild-type human trypsinogen, using oligonucleotide S549N. For mutations changing cysteines in the disulfide-knotted domain, oligonucleotides Cys2→Ser which alters cysteine 147 to serine (TGC to TCC), Cys3→Ser which alters cysteine 153 to serine (TGC to TCC) and Cys5→Ser which alters cysteine 173 to serine (TGC to TCC) were used. For mutations which alter each of the two putative start codons, the oligonucleotides used were Met1→Leu and Met2→Leu which change methionines 1 and 24, respectively, to leucine (ATG to TTG). To alter serine 49 to asparagine (AGC to AAC) and thus remove the putative protein kinase C (PKC) phosphorylation site in the intracellular domain, oligonucleotide PKC phos S49N was used. Following thermal cycling, mutagenesis reactions were digested with the *Dpn* I endonuclease, specific for methylated and hemimethylated DNA, to digest the parental DH5α-derived DNA template and thus select for newly synthesized mutation-containing DNA. Upon digestion with *Dpn* I, mutagenesis reaction DNA was transformed into XL1-Blue supercompetent cells. Presence of mutations was verified by sequencing the entire Stubble coding region.

Table 5. Sequences of oligonucleotides used in site-directed mutagenesis reactions

Oligonucleotide Name	Sequence (5'→3')
F H573R	TCTCGAGCACCC**G**CCGTTGCGGTGG
F P137L	GATCAGCCCCAAGC**T**ATGCTCCTTTGGCC
G760S F	GGGATCATCTCCTGG**A**GCATTGGCTGTGCCG
G760R F	GGGATCATCTCCTGG**C**GCATTGGCTGTGCCG
S549I F	CGTGGGCGGTAAGA**T**CGCGGCCTTCGGTCG
S549N F	CGTGGGCGGTAAGA**A**CGCGGCCTTCGGTCG
F Cys2→Ser	GTCGAGGGCACCT**C**CATGTTCGTGTGG
F Cys3→Ser	GTTCGTGTGGGAGT**C**CATCAAGTCCGAGG
F Cys5→Ser	CATGTTCGGCTCCT**C**CTGCACGCACAACT
F Met1→Leu	CCGGCACGTTACCG**T**TGAAGCAGCCAACT
F Met2→Leu	CGGCAGCCACCAAG**T**TGTGTCCCAAAAGG
F PKC phos S49N	TGCGGCGCGGA**A**CAGGCGCAGTC

Only forward oligonucleotides and sequences are given. The reverse compliment of each forward oligonucleotide sequence gives the corresponding reverse oligonucleotide sequence. Point mutations induced by site-directed mutagenesis are shown in bold.

Table 6. Cycling parameters for site-directed mutagenesis

Segment	Cycles	Temperature	Time
1	1	95ºC	30 seconds
2	12	95ºC	30 seconds
		55ºC	1 minute
		68ºC	7 minutes

Transgenic constructs

Construction of the wild-type *hs-Sb*$^{+}$ transgene was described previously (Hammonds 2002; Bayer 2003). Briefly, a full-length 3.8 kb *Stubble* cDNA was cloned into the 2.8 kb pLitmus29 vector and subsequently into the 8.9 kb pCaSpeR-hs heat-shock-inducible *P*-element transformation vector to allow inducible expression of the transgene under heat-shock control.

Mutant transgenic constructs were generated by sub-cloning the mutagenized *Stubble* cDNA (described above) from the pLitmus29 vector into pCaSpeR-hs. Mutagenized *Stubble* DNA fragments were removed from pLitmus29 by digestion with BamHI/SpeI and were gel purified following electrophoresis in 0.8% SeaPlaque GTG agarose (BioWhittaker Molecular Applications, Rockland, ME) in 1X TAE buffer. Bands corresponding to the *Stubble* mutant constructs were excised from the agarose and purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and cloned into BglII/XbaI-digested pCaSpeR-hs using the Rapid DNA Ligation Kit (Roche Diagnostics Corp., Indianapolis, IN). Ligation products were then transformed into MAX Efficiency DH5α or TOP10 One Shot competent cells (Invitrogen, Carlsbad, CA) according to manufacturer's specifications with the following exceptions: MAX Efficiency competent cells were used in 50µl aliquots and 1µl of undiluted ligation reaction (1-10 ng) was added for transformation. The resulting *hs-Stubble* mutant constructs were injected into w^{1118} embryos to generate transgenic lines.

CHAPTER THREE: RESULTS – CHARACTERIZATION OF *STUBBLE* MUTANT ALLELES

At present, the mechanism by which Stubble-mediated extracellular proteolysis influences the intracellular Rho1 signaling pathway to control of epithelial morphogenesis is poorly understood. In order to better understand the role played by the Stubble protease in epithelial morphogenesis, it is necessary to define Stubble action on a mechanistic level. One approach to a better understanding of the mechanism of action of Stubble is to determine which of the conserved domains found within this protease are necessary for its function. Using sequence analysis and genetic interaction assays, I have genetically characterized and identified functional domains of sixteen mutant alleles of the Stubble TTSP.

Stubble mutant alleles used in this study

In an attempt to determine which of the domains contained within the Stubble protease are necessary for its function, we conducted sequence analysis of 16 recessive Stubble (*sbd)* mutant alleles having several origins (see Table 1). The six $sbd^{E(br)}$alleles were generated by Ward et al. (2003) through EMS mutagenesis of a stock carrying the br^{1} allele. Mutagenized chromosomes were allowed to undergo free recombination for several generations to remove unlinked second-site mutations before being balanced.

The sbd^{46}, sbd^{173} sbd^{177}, sbd^{241}, sbd^{258}, sbd^{266}, and sbd^{277} alleles were produced by EMS mutagenesis of w^{1118} animals carrying isogenic second and third chromosomes (Ruggiero, 2006). These alleles were outcrossed after balancing to remove extraneous mutations on the X and

second chromosomes; however, this process did not allow free recombination and thus would not have removed any second-site EMS-induced mutations carried on the third chromosome. It was previously stated that *sbd*^{177} no longer carries the *sbd* mutation (Ruggiero, 2006). This was confirmed by generating *sbd*^{177}/*sbd*^{201} *trans*-heterozygotes. These animals are expected to have reduced bristles and malformed legs and wings. In contrast, a wild-type phenotype was observed indicating that the *sbd*^{177} allele had been lost from the stock. Thus, this allele was excluded from further analysis.

The alleles *sbd*^{PNR11} and *sbd*^{VE3} were produced by EMS mutagenesis of *st e* animals and *sbd*^{206} was produced by EMS mutagenesis of animals of the genotype *ru cu ca* (Appel, 1992). Further information on these alleles is unavailable. Finally, *sbd*^{1} is a spontaneous mutation first discovered by Sturtevant in 1926 and published by Dobshansky in 1930.

Sequencing of *sbd* alleles

sbd$^{E(br)}$ alleles

Genomic DNA from *sbd*$^{E(br)}$*/TM6B, Tb Hu e* or *sbd*$^{E(br)}$/*sbd*$^{E(br)228}$ heterozygotes and the *br*^{1} progenitor stock for the *sbd*$^{E(br)}$ alleles was sequenced on both strands through the entire protein-coding region. Four differences from either of the published wild-type *Stubble* sequences (Appel et al., 1993; gi (GenBank ID) 158511 and Adams et al., 2000; gi 7300108) were found in all six *sbd*$^{E(br)}$ alleles (Table 7). In an S/T-rich portion of exon 7, the *sbd*$^{E(br)}$ alleles contain sequence variation del2129_2131, an in-frame 3-bp deletion which eliminates the threonine residue at amino acid 421, thereby reducing a string of 8 threonines to 7. Also found in the S/T-rich region of exon 7 is sequence variation 2167A>G, a silent A-to-G single base change at

cDNA nucleotide 2167. In exon 9, sequence variation 2752G>T, a silent G-to-T substitution at nucleotide 2752, is found in all $sbd^{E(br)}$alleles. Finally, in exon 10, sequence variation 3037C>T, a single base change (C to T) at nucleotide 3037, produces no amino acid change. These four variations are also found in the br^{1} progenitor stock and thus are considered polymorphisms and are not attributed to the *sbd* phenotype observed in animals carrying these alleles. Four previously published polymorphisms, all found in exon 6, were also identified in all six $sbd^{E(br)}$ alleles and in the progenitor stock (Table 7). These include sequence variations del(1727_1744)CAG, a 3-bp CAG in-frame deletion which reduces a string of seven glutamines to six between amino acids 287 and 293; 1862C>A, a silent C-to-A point mutation at nucleotide 1862; 1913A>T, an A-to-T change at nucleotide 1913 that results in a serine-for-threonine substitution at amino acid 349 and 1928C>T, a serine-for-proline substitution at amino acid 354 caused by a single C-to-T base change at nucleotide 1928 (Hammonds and Fristrom, 2006).

Table 7. Polymorphisms of the Stubble locus

Sequence Variation[a, b]	Exon	Amino Acid Level[b]	Description	Alleles[c]
1025T>C	4	silent	Silent T-to-C single base substitution at nucleotide 1025	*sbd*^{1}
1150T>C	4	silent	Silent T-to-C single base substitution at nucleotide 1150	*sbd*^{VE3}, *sbd*^{PNR11}
del(1727_1744) CAG	6	delQ(287_293)	In-frame 3-bp deletion of a CAG repeat reduces a string of 7 glutamines to 6 between amino acids 287 and 293	*sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{241}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *sbd*^{VE3}, *sbd*^{PNR11}, *br*^{1}
1862C>A	6	silent	Silent C-to-A single base substitution at nucleotide 1862	*sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{241}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *sbd*^{VE3}, *sbd*^{PNR11}, *br*^{1}
1913A>T	6	T349S	A-to-T single base substitution at nucleotide 1913 results in a threonine-to-serine substitution at amino acid 349	*sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{241}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *sbd*^{VE3}, *sbd*^{PNR11}, *br*^{1}
1928C>T	6	P354S	C-to-T single base substitution at nucleotide 1928 results in a proline-to-serine substitution at amino acid 354	*sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{241}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *sbd*^{VE3}, *sbd*^{PNR11}, *br*^{1}
del2090_2098	7	delSTT408_410	In-frame 9-bp deletion from nucleotide 2090 to nucleotide 2098 removes a serine, threonine, threonine (STT) repeat from amino acids 408 to 410	*sbd*^{206}, *sbd*^{258}, *sbd*^{266}, *sbd*^{VE3}, *sbd*^{PNR11}, *isoline*
del2129_2131	7	delT421	In-frame 3-bp deletion of ACA reduces a string of 8 threonines to 7 at amino acid 421 in an S/T-rich region	*sbd*^{1}, *sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{241}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *br*^{1}
2167A>G	7	silent	Silent A-to-G single base substitution at nucleotide 2167	*sbd*^{1}, *sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{206}, *sbd*^{241}, *sbd*^{258}, *sbd*^{266}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *sbd*^{VE3}, *sbd*^{PNR11}, *br*^{1}, *isoline*
2377C>T	7	silent	Silent C-to-T single base substitution at nucleotide 2377	*sbd*^{258}, *sbd*^{266}, *isoline*
2430A>G	7	H521R	A-to-G single base substitution at nucleotide 2430 results in a histidine-to-arginine substitution at amino acid 521	*sbd*^{VE3}, *sbd*^{PNR11}
2752G>T	9	silent	Silent G-to-T single base substitution at nucleotide 2752	*sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{241}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *br*^{1}
2812G>A	9	silent	Silent G-to-A single base substitution at nucleotide 2815	*sbd*^{1}
2827G>T	9	silent	Silent G-to-T single base substitution at nucleotide 2827	*sbd*^{1}, *sbd*^{206}
3001T>C	10	silent	Silent T-to-C single base substitution at nucleotide 3001	*sbd*^{258}, *sbd*^{266}, *sbd*^{VE3}, *sbd*^{PNR11}, *isoline*
3037C>T	10	silent	Silent C-to-T single base substitution at nucleotide 3037	*sbd*^{1}, *sbd*^{46}, *sbd*^{173}, *sbd*^{201}, *sbd*^{206}, *sbd*^{241}, *sbd*^{277}, *sbd*$^{E(br)20}$, *sbd*$^{E(br)48}$, *sbd*$^{E(br)228}$, *sbd*$^{E(br)448}$, *sbd*$^{E(br)536}$, *sbd*$^{E(br)623}$, *br*^{1}

[a] Nucleotide positions are in cDNA numbering corresponding to the published genomic sequence from a *y; cn bw sp* strain (Adams et al., 2000; gi 7300108).

[b] Parentheses indicate uncertainty in the position of the variation.

[c] Isoline refers to *w1118; iso2; iso3* stock used as the progenitor for the *sbd* alleles produced by Ruggiero (2006).

In addition to the eight polymorphisms identified, four of the six *sbd*$^{E(br)}$ alleles also contain additional changes within the protease domain that are unique to each allele (Table 8; Figure 9). Exon 11 of the *sbd*$^{E(br)20}$ and *sbd*$^{E(br)228}$ alleles contains single amino acid substitutions resulting in missense mutations. In *sbd*$^{E(br)20}$, a C-to-T change at nucleotide 3141 results in a phenylalanine-for-serine substitution (TCC to TTC) at amino acid 758, and in *sbd*$^{E(br)228}$, a G-to-A change at nucleotide 3146 results in an arginine-for-glycine substitution (GGC to AGC) at amino acid 760.

Table 8. Mutations identified in sbd alleles

Allele	Mutation	Allele	Mutation
sbd$^{E(br)20}$	S758F	*sbd*^{PNR11}	D737N
sbd$^{E(br)48}$	Regulatory?	*sbd*^{VE3}	P662L
sbd$^{E(br)228}$	G760S	*sbd*^{206}	Q446 Nonsense
sbd$^{E(br)448}$	Q714 Nonsense	*sbd*^{1}	Regulatory?
sbd$^{E(br)536}$	Regulatory?	*sbd*^{258}	G755E
sbd$^{E(br)623}$	24bp deletion in intron separating exons 10 and 11 – removes lariat sequence	*sbd*^{266}	P137L; T587N directly followed by Δnt2630-2922 – causes frameshift leading to truncation after addition of 14 residues

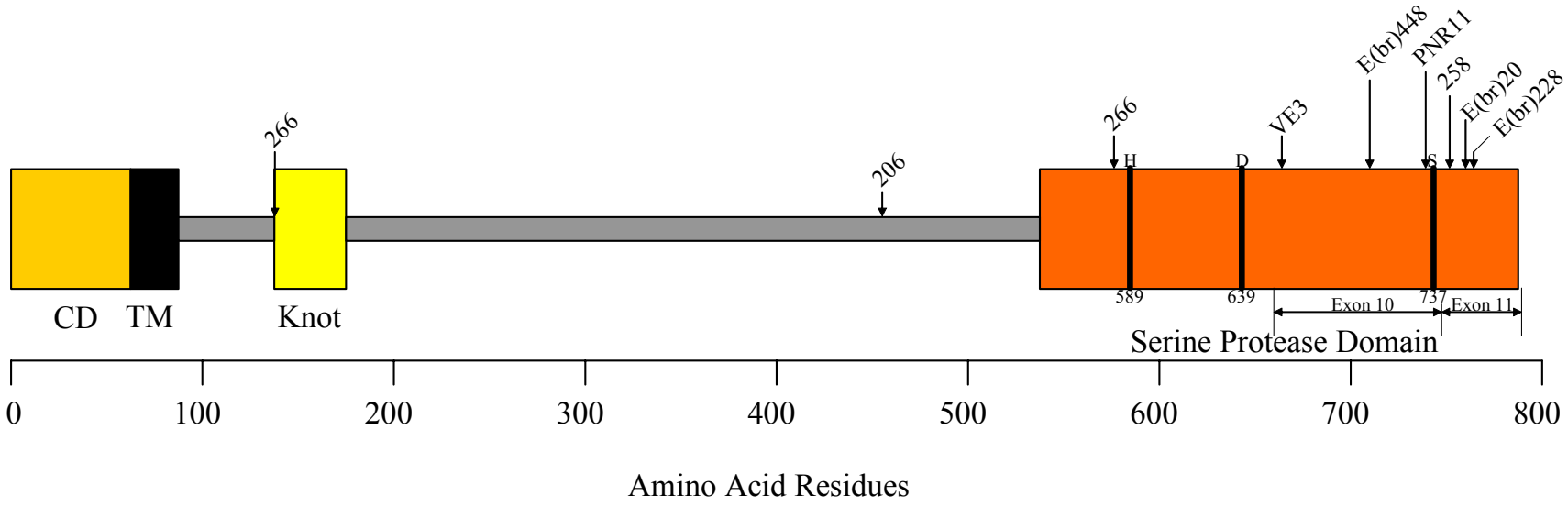

Figure 9. Locations of point mutations in sbd alleles.

Amino acid residue number is shown in the scale at the bottom. Stubble protein domains are shown: CD (cytoplasmic domain) aa 1-58, TM (transmembrane domain) aa 59-81, knot (disulfide knotted domain) aa 138-173, and serine protease domain from the activation cleavage site at aa 543 to the C terminus at aa 787. The residues of the catalytic triad are denoted by solid vertical bars. The boundary between exons 10 and 11 is also noted.

The allele *sbd*$^{E(br)448}$ contains a single amino acid substitution (C to T) at nucleotide 3008 which replaces the glutamine residue at amino acid position 714 with a stop codon, thus truncating the protein 74 residues before its C-terminus and removing the catalytic serine and the substrate binding site.

Finally, the allele *sbd*$^{E(br)623}$ contains a 24-bp deletion within the intron separating exons 10 and 11. The deleted portion of the intron contains the lariat branch point consensus sequence CTAAT, in which branch formation occurs at the underlined A (Mount et al., 1992). No additional mutations were found in the coding region of the remaining two alleles, *sbd*$^{E(br)48}$ and *sbd*$^{E(br)536}$, thus the mutations leading to the *sbd* phenotype in these mutants are presumed to be of a regulatory nature.

sbd^{46}, *sbd*^{173}, *sbd*^{241} and *sbd*^{277} are all *sbd*^{201} alleles

Four of the putative *sbd* mutant alleles identified by Ruggiero (2006) in his EMS screen are actually *sbd*^{201} alleles. Genomic DNA from *sbd*^{46}, *sbd*^{173}, *sbd*^{241}, and *sbd*^{277} in homozygous condition or in heterozygous condition with the *w*^{1118}*; iso2; iso3* progenitor line, as well as genomic DNA from the wild-type *w*^{1118}*; iso2; iso3* progenitor line itself, was sequenced on both strands through the entire protein-coding region (see Table 2). In each of these alleles, an identical set of four previously published polymorphisms as well as five variations distinct from either of the published wild-type sequences were found (Table 7). The four previously published polymorphisms found in *sbd*^{46}, *sbd*^{173}, *sbd*^{241}, and *sbd*^{277} (del(1727_1744)CAG, 1862C>A, 1913A>T and 1928C>T) as well as four of the five newly identified differences (del2129_2131, 2167A>G, 2752G>T and 3037C>T) are identical to the set of polymorphisms described for the *sbd*$^{E(br)}$ and *br*^{1} alleles. The *br*^{1} stock was also the progenitor for the *sbd*^{201} allele used by

Ruggerio (2006) to isolate new *sbd* alleles. The remaining variation is an A-to-G point mutation at nucleotide 2586 in exon 8, which results in an amino acid substitution of arginine (CGC) for histidine (CAC) at amino acid 573. This change has been attributed by Hammonds and Fristrom (2006) to the sbd^{201} phenotype. The polymorphisms found in the w^{1118}*; iso2; iso3* progenitor stock used in the screen from which sbd^{46}, sbd^{173}, sbd^{241}, and sbd^{277} were identified support the finding that these alleles are actually alleles of sbd^{201}. The w^{1118}*; iso2; iso3* stock (isoline in Table 7) contains three polymorphisms not observed in the sbd^{46}, sbd^{173}, sbd^{241}, and sbd^{277}alleles. These include sequence variations del2090_2098, a 9-bp in-frame deletion in exon 7 that results in the deletion of one unit of a tandem repeat of serine, threonine, threonine (STT) between amino acids 408 and 410; 2377C>T, a silent C-to-T point mutation at nucleotide 2377 in exon 7 and 3001T>C, a silent T-to-C point mutation at nucleotide 3001 in exon 10.

$sbd^{258}$and sbd^{266}

Genomic DNA from sbd^{258} and sbd^{266}*/TM6B Tb Hu e* or w^{1118}*; iso2; iso3/*sbd^{258} and sbd^{266} heterozygotes was sequenced on both strands through the entire protein-coding region (see Table 2). These alleles were produced from the w^{1118}*; iso2; iso3* progenitor stock, and carry all of the polymorphic changes associated with the progenitor (Table 7). In addition, sbd^{258} contains a single base change (G to A) at nucleotide 3132 which results in a glycine-to-glutamic acid amino acid substitution at residue 755 (Table 8; Figure 9).

In addition to the polymorphisms observed in the progenitor stock, sbd^{266} contains a C-to-A point mutation at nucleotide 2628, which results in a threonine-to-asparagine (ACC to AAC) substitution at residue 587, followed by an out-of-frame 292-bp deletion from nucleotide 2630 to nucleotide 2922 that adds 14 incorrect amino acids onto modified residue 587 (Asp) before

reading into a stop codon and effectively truncating the protein before any of the catalytic triad residues have been added (Table 8; Figure 9). In addition, *sbd*^{266} contains a C-to-T point mutation at nucleotide 1278, which results in an amino acid substitution of leucine-for-proline at amino acid 137.

sbd^{PNR11} and *sbd*^{VE3}

Genomic DNA from *sbd*^{PNR11} and *sbd*^{VE3}*/TM6B, Tb Hu e* heterozygotes or *sbd*^{PNR11} and *sbd*^{VE3} homozygotes was sequenced on both strands through the entire protein-coding region (see Table 2). The progenitor was not available for sequencing. Nine polymorphisms were identified in both alleles, consistent with their origin from a common progenitor (Table 7). These include sequence variants 1150T>C, a silent T-to-C point mutation at nucleotide 1150 in exon 2; 2430A>G, a single A-to-G base change at nucleotide 2430 resulting in an arginine-for-histidine substitution at amino acid 521 in exon 7 and del(1727_1744)CAG, 1862C>A, 1913A>T,1928C>T, del2090-2098, 2167A>G and 3001T>C. In addition to the polymorphic variations, *sbd*^{PNR11} contains a G-to-A point mutation at nucleotide 3077 which results in the substitution of an asparagine (AAT) for an aspartic acid (GAT) at amino acid 737 (Table 8; Figure 9). *sbd*^{VE3} also contains a sequence variation in addition to the polymorphisms identified. This variant is a single C-to-T point mutation at nucleotide 2853 which leads to substitution of leucine (CTC) for proline (CCC) at amino acid 662. This proline is a conserved residue in a conserved region of the protease domain (Table 8; Figure 9).

sbd^{206}

Genomic DNA from *sbd*^{206} *ru cu e ca/TM6B, Tb Hu e* heterozygotes or *sbd*^{206} *ru cu e ca* homozygotes was sequenced on both strands through the entire protein-coding region (see Table 2). Five differences from either of the published wild-type sequences were found. Four of the five differences were polymorphic sequence variations (Table 7): del2090_2098, 2167A>G, 2827G>T, a silent G-to-T single base change at nucleotide 2827 in exon 9 and 3037C>T. The fifth and more significant difference is a C-to-T single base change at nucleotide 2324 which results in a nonsense mutation, substituting a stop codon (TAA) for the glutamine (CAA) residue at amino acid position 446 in the stem region (Table 8; Figure 9). This mutation results in truncation of the *Stubble* protease in exon 7 prior to the start of the protease domain.

sbd^{1}

Genomic DNA from *sbd*^{1} *st ro e ca* homozygous animals was sequenced on both strands through the entire protein-coding region. Six differences from either of the published wild-type sequences were found (Table 7). Four are polymorphic sequence variations found in other *sbd* alleles (del2129_2131, 2167A>G, 2827G>T and 3037C>T). The remaining two variations are both silent changes not found in any of the other sequenced alleles. These include sequence variants 1025C>T and 2812G>A (Table 7), a C-to-T single base change at nucleotide 1025 in exon 4 and a G-to-A change at nucleotide 2812 in exon 9, respectively. Though these two variations have only been identified in one allele, they are most likely polymorphisms because they are silent changes. No other changes have been identified in the coding region of *sbd*^{1}, suggesting that the mutation which causes the *sbd* phenotype may be of a regulatory nature.

Second-site noncomplementation analyses between *sbd* mutations and mutations in Rho1 pathway members

Genetic interaction studies of molecularly characterized mutant alleles of a gene can be a useful tool to rank the relative severity of different mutations. This, in turn, can establish the relative importance of the conserved domains within the protein to protein function and potentially link individual domains to different developmental functions.

In order to genetically characterize the twelve mutant *sbd* alleles in Table 8, we first determined the strength of interaction between these alleles and mutations in Rho1 pathway members using second-site noncomplementation (SSNC) analyses. SSNC assays test pairs of mutant alleles in double heterozygous condition (*i.e. Rho1*/+; *Sb*/+). Strong genetic interactions observed in SSNC assays frequently indicate that two genes function in the same signaling pathway or act directly on the same developmental process. Previous work in our lab has shown that different combinations of *Rho1* and *Stubble* alleles can result in high frequencies of animals with leg malformation in SSNC assays (Bayer et al., 2003). For example, the dominant alleles Sb^{63b} and Sb^{70} interact very strongly with *Rho1* alleles and a *Rho1* deficiency, with malformation frequencies ranging from 59-95%. In contrast, the recessive sbd^{105} deficiency allele shows weak interaction with *Rho1* alleles and moderate interaction with a *Rho1* deficiency. No interactions were observed with the recessive sbd^{1} or sbd^{201} alleles. In these studies, interactions were defined as strong when 50% of animals of the relevant genotype had at least one malformed leg, moderate when 25 to 49% of animals show malformation, and weak when 5 to 24% malformation is exhibited (Bayer et al., 2003).

To extend these data, the 12 *sbd* alleles analyzed in this study were tested for SSNC interactions with mutations in *Rho1* and *zipper*, a member of the Rho1 signaling pathway, and 10

were tested for dominant interactions with the *Rho1*^{72O} allele. A majority of the SSNC assays between *sbd* mutant alleles and *Rho1* mutant alleles were carried out by an undergraduate assistant, Stephanie Stallworth, from Daytona Beach Community College participating in the NIH Bridges to the Baccalaureate Program who worked under my supervision.

Genetic interaction data for *sbd* and *Rho1* mutant alleles demonstrate that *sbd* alleles show either weak or no interaction with *Rho1*^{72O}, *Rho1*$^{J3.8}$, *Rho1*$^{E.3.10}$ and *Rho1*^{k02107b} (Table 9). Leg malformation frequencies range from 0 to 18%, with the majority of *Rho*^{1}*; sbd* heterozygous combinations exhibiting less than 5% malformation. In combination with the *zipper* mutant *zip*^{Ebr}, three *sbd* alleles (*sbd*^{258}, *sbd*^{PNR11} and *sbd*^{206}) exhibit moderate frequencies of malformation while the remaining nine alleles exhibit weak or no interaction (Table 9). Because all twelve *sbd* alleles are inviable as homozygotes, the *trans*-heterozygous combination *sbd*/sbd*^{1}, was used for dominant interaction assays. With the exception of *sbd*^{VE3}*/sbd*^{1}, *Rho1*^{72O} shows very strong dominant interactions with all ten *sbd*/sbd*^{1} *trans*-heterozygotes tested (Table 9). However, the background malformation rates for these *sbd*/sbd*^{1} combinations range from 59 to 100% (see *sbd*/*sbd* combinations in Table 9) and thus severely limit the usefulness of the data obtained from these dominant genetic interaction assays. In contrast, a weak interaction was observed between *Rho1*^{72O} and *sbd*^{VE3}*/sbd*^{1} *trans*- heterozygotes, which have a background malformation rate of 0% (see *sbd*^{VE3}*/sbd*^{1} in Table 9).

Table 9. Second-site noncomplementation analyses of sbd mutant alleles

	sbd^{201}*e*	*sbd*^{201}	*sbd*^{1}	*Rho1*^{72O}	*Rho1*$^{J3.8}$	*Rho1*$^{E3.10}$	*Rho1*^{k02107b}	*zip*^{Ebr}	*Rho1*^{72O}*; sbd*^{1}
sbd^{258}	89.1 (46)	97.6 (43)	78.4 (51)	0 (64)	1.1 (92)	2.4 (83)	0 (73)	27.8 (126)	98.5 (66)
sbd^{266}	100 (6) semi-lethal	100 (33)	80.7 (57)	0 (93)	0.8 (130)	1.2 (85)	2.2 (91)	21.3 (108)	97.6 (41)
sbd$^{E(br)20}$	100 (50)	100 (110)	100 (51)	1.0 (99)	0 (55)	4.2 (142)	7.3 (82)	18.9 (122)	100 (61)
sbd$^{E(br)48}$	100 (57)	100 (99)	90.0 (120)	0 (50)	0 (39)	4.4 (135)	2.4 (83)	4.6 (109)	97.5 (79)
sbd$^{E(br)228}$	100 (64)	100 (66)	80.9 (42)	1.3 (79)	0.9 (108)	0.8 (113)	2.3 (86)	1.0 (98)	83.3 (60)
sbd$^{E(br)448}$	94.1 (51)	100 (86)	97.1 (70)	0 (103)	0 (74)	1.6 (122)	0 (107)	11.2 (107)	100 (77)
sbd$^{E(br)536}$	100 (78)	100 (38)	97.3 (74)	2.5 (80)	17.5 (120)	7.0 (113)	8.7 (92)	22.2 (81)	100 (67)
sbd$^{E(br)623}$	97.9 (47)	100 (135)	95.4 (65)	0.9 (106)	8.0 (87)	3.8 (133)	3.2 (93)	11.8 (68)	98.8 (83)
sbd^{VE3}	0 (76)	0 (88)	0 (13)	0 (115)	0 (117)	1.6 (63)	1.9 (104)	0.8 (124)	16.7 (30)
sbd^{PNRII}	100 (44)	100 (34)	96.4 (28)	0 (122)	1.1 (90)	3.7 (107)	0.9 (116)	35.8 (120)	100 (35)
sbd^{206}	100 (64)	100 (115)	ND	2.3 (129)	4.0 (100)	11.9 (109)	5.4 (93)	33.6 (110)	ND
sbd^{1}	19.8 (116)	45.8 (205)	ND	0.7 (143)	1.9 (104)	1.7 (172)	1.6 (122)	6.7 (149)	ND

Animals are either doubly heterozygous, *e.g.* *Rho1*^{72O}/+*;* *sbd*^{258}/+, or *trans*-heterozygous, *e.g.* *sbd*^{258}/*sbd*^{201} *e*. The numbers shown indicate the percentage of animals with malformed legs, with the total number of animals of the indicated genotype scored shown in parentheses. ND, not determined.

Lethal-phase analysis of *sbd* mutant alleles

To characterize the effects of each mutation on Stubble protease function, lethal-phase studies were conducted. These experiments were done in the same manner as described by Ward et al. (2003) to characterize the $sbd^{E(br)}$ alleles (see Methods). These experiments reveal that sbd^{258} exhibits predominantly embryonic lethality, with only 10% of embryos hatching (Table 10). Those animals that do hatch die as early first instar larvae. Four of the six mutants tested, sbd^{266}, sbd^{PNR11}, sbd^{VE3}, and sbd^{206}, are predominantly larval lethal. The remaining allele, sbd^{1}, shows broad stage-specific lethality with ~45% viable adults.

Table 10. Lethal phase analysis of sbd mutants

	Embryonic lethality[a]	Larval lethality[b]	Pupal lethality[c]
sbd^{258}	89.6 ± 1.7	100	NA
sbd^{266}	4.0 ± 0.4	100	NA
sbd^{PNR11}	8.0 ± 1.9	100	NA
sbd^{VE3}	29.1 ± 4.1	100	NA
sbd^{206}	6.8 ± 0.1	100	NA
sbd^{1}	26.4 ± 6.4	31.0 ± 2.6	10.8 ± 2.5

[a] Mean ± SE of embryonic lethality from three independent experiments. Embryonic lethality was calculated as (number of dead embryos/total number of mutant embryos) x 100.
[b] Mean ± SE of larval lethality from three independent experiments. Mutant larvae were picked as newly hatched first instar larvae. Larval lethality was calculated as [(total number of mutant larvae – number of pupae)/total number of mutant larvae] x 100.
[c] Mean ± SE of pupal lethality from three independent experiments. Pupal lethality was determined by allowing those larvae from the larval lethality assay that pupariated to eclose. Adults were counted and pupal lethality was calculated as [(total number of pupae – number of adults)/total number of pupae] x 100. NA, not applicable.

Experiments in progress

Classification of *sbd* mutant alleles.

In 1932, Muller developed a naming system to classify mutations based on their behavior in various genetic situations (Muller, 1932). He coined the terms amorph, hypomorph, hypermorph, antimorph and neomorph to define the genetic behavior of various classes of alleles he observed. In Muller's terminology, an amorphic mutation results in a complete loss of gene function while a hypomorphic mutation causes a partial loss of gene function. Hypermorphic, antimorphic (also called dominant negative) and neomorphic mutations result in a gain of gene function. A hypermorph causes an increase in normal gene activity (e.g. transcription), an antipmorph is a dominant mutation that acts antagonistically to normal gene activity, and a neomorph is a mutation that results in a gain of novel gene activity. If the behavioral classification of an allele can be determined, it can prove useful in the interpretation of classical and molecular genetic data.

To classify the nine non-regulatory mutations as well as the three putative regulatory mutations in our *sbd* alleles (see Table 8) as amorphic, hypermorphic or gain-of-function, the developmental progress and lethal phase of *sbd* homozygotes was monitored and compared with that of *trans*-heterozygous *sbd*/Df(3R)sbd*^{105} animals. The *Df(3R)sbd*^{105} allele contains a deficiency which removes the entire *Stubble* locus. In *trans*-heterozygous condition with a deficiency, an amorphic mutant should exhibit identical developmental progress and lethal phase as the amorphic homozygote since both genotypes result in complete loss of gene function. A hypomorphic mutant should exhibit a more severe phenotype over a deficiency than as a homozygote. Hypermorphic mutations should exhibit a less severe phenotype over a deficiency

than as a homozygote. The phenotype of dominant negative homozygous mutations compared to dominant negative over a deficiency is hard to predict; however, the severity of phenotype of a dominant negative can be reduced by over-expressing the wild-type protein. In contrast, the phenotype of a hypermorphic mutation is expected to become worse if the the wild-type protein is over-expressed. Finally, the phenotype of neomorphic mutations as homozygotes or over a deficiency is expected to be similar; however, over-expression of wild-type protein will have no effect on the phenotype.

At present, only sbd^{258}, sbd^{PNR11}, sbd^{206}, and $sbd^{E(br)228}$ have been tested. All other alleles (see Table 8) remain to be analyzed. As homozygotes, sbd^{258} animals are predominantly embryonic lethal, whereas sbd^{PNR11}, sbd^{206}, and $sbd^{E(br)228}$ animals exhibit predominantly larval lethality (see Table 10). As *trans*-heterozygotes with $Df(3R)sbd^{105}$, all four alleles show improved development with at least half of the animals surviving to adulthood, suggesting that all are associated with gain-of-function mutations.

CHAPTER FOUR: RESULTS – SCREEN FOR G PROTEIN-COUPLED RECEPTORS THAT INTERACT WITH *STUBBLE* TO CONTROL LEG DEVELOPMENT

It has been proposed that the Stubble protease operates through an "outside-in" signaling mechanism during leg development to either act in parallel with or directly regulate Rho1 signaling. Four vertebrate TTSPs have also been shown to operate via an apparent outside-in signaling mechanism using the GPCR PAR-2. In addition, GPCRs are known to be major upstream regulators of vertebrate Rho signaling through activation of heterotrimeric G proteins. Because GPCRs have been linked to both extracellular proteolytic activity and intracellular Rho signaling, I asked if the known non-odorant and non-gustatory G protein-coupled receptors within the *Drosophila* genome interact genetically with *Stubble* in the control of leg imaginal disc morphogenesis. Though it is not inconceivable that an odorant or gustatory receptor could link Stubble proteolytic activity with Rho1 signaling in epithelial development, these receptor types were excluded based on their known involvement in the nervous system.

Excluding the 119 gustatory and odorant receptors, 147 genes encoding G protein-coupled receptors have been identified in the *Drosophila* genome. Four of these have not been mapped to the genome sequence and thus were also excluded. To screen for G-protein-coupled receptors potentially involved in leg morphogenesis, SSNC assays were performed by mating heterozygous *Stubble* mutant animals (*i.e., red* $Sb^{63b}e$*/TM6B, Tb Hu e*) with heterozygous or homozygous animals carrying a deficiency uncovering the GPCR, or a lethal transposable element insertion or mutation in the GPCR. Resulting adults of the doubly mutant heterozygous class, *e.g.*, *GPCR/+; red* $Sb^{63b}e$*/+*, were scored for frequency of leg malformation. *Stubble*

mutant animals exhibit a background malformation rate of ~2%. A minimum increase in malformation rate of 20-fold above background is considered significant, and thus only crosses producing greater than 40% malformation in the doubly heterozygous F1 class were deemed significant. At present, 98 of the 147 GPCR-encoding genes have been tested for genetic interactions with *Stubble* in a total of 155 individual assays. A majority of this work was done by Nikette Benjamin, an undergraduate research assistant who worked under my supervision. From this screen, we identified 4 interacting chromosomal regions which collectively uncover 8 GPCRs (Table 11). A list of all deficiencies and mutations tested is given in Appendix 1.

Table 11. GPCR deficiencies showing strong interactions with Sb^{63b}

Deficiency	Region Uncovered	GPCRs Uncovered	% mlf with Sb^{63b}
Df(2L)Exel7027	26F6-27B1	GRHR	57.1 (28)
Df(2R)Exel6068	56A1-56B5	5-HT1A, 5-HT1B	34.5[a] (29)
Df(3L)Exel6084	61B2-61C1	mthl9, mthl10, mthl14	80.4 (92)
Df(3R)red-P52	88A4-88B2	CG9918, DopR	87.5 (40)

[a] The percent malformation for *Df(2R)Exel6068/+; red* Sb^{63b} *e/+* animals was less than 40%; however, this deficiency was kept for further evaluation due to the low number of double heterozygote animals produced in the primary assay.

Df(2L)Exel7027

The first interacting region, 26F6-27B1, is uncovered by the deficiency *Df(2L)Exel7027*. This region contains 21 genes, including the GPCR gonadotropin-releasing hormone receptor (GRHR). Doubly heterozygous *Df(2L)Exel7027/+; red* Sb^{63b} *e/+* animals exhibit 57% leg malformation, suggesting that one or more of the 21 genes deleted on the deficiency

chromosome is involved in leg disc morphogenesis. Because a high frequency of malformation was obtained for the deficiency in combination with Sb^{63b}, we performed a series of follow-up experiments to verify the initial genetic interaction, and reduce the size of the interacting region carrying the interacting gene(s) (Table 12). To assay for consistent interaction between the deficiency allele and the *Stubble* locus, *Df(2L)Exel7027* animals were crossed to animals carrying Sb^{70}, a second mutant allele of *Stubble*. Animals of the doubly heterozygous *Df(2L)Exel7027/+;* Sb^{70}/+ class exhibit 61% leg malformation, consistent with the malformation frequency obtained with Sb^{63b}. Deficiency animals were also mated with animals carrying br^1, a mutant allele of *broad*. The *broad* locus encodes an ecdysone-inducible gene which interacts genetically with the *Stubble* locus (Beaton et al., 1988). The br^1/*Y*; *Df(2L)Exel7027/+* male progeny from this cross exhibit only 6% malformation indicating that *Df(2L)Exel7027* does not interact dominantly with br^1. This low malformation rate does not necessarily exclude the (unidentified) interacting gene from involvement in the ecdysone hierarchy. Previous experience in our laboratory has shown that genetic interactions can be very allele specific where loss-of-function alleles do not always show interactions while gain-of-function alleles of the same gene may. For example, null alleles of the non-muscle myosin heavy chain gene, *zipper*, are non-interactive in SSNC assays, whereas weak gain-of-function alleles of this gene, e.g. zip^{Ebr}, are interactive. A reasonable explanation for these observations is that one functional copy of the *zipper* gene product is sufficient for proper development. However, because the zipper protein functions as a dimer, the presence of a weak gain-of-function allele results in a high proportion of unproductive dimers and interferes with the function of the wild-type gene. Further assays with mutants of *Rho1* and *Notopleural* (*Np*), a genetic locus that potentially cooperates with the

Stubble locus in regulation of bristle formation and leg development (Ruggerio, 2006), are pending.

Table 12. Second-site noncomplementation analyses with Stubble interactors

Deficiency	Sb^{70}	$Rho1^{E3.10}$	$Rho1^{J3.8}$	Np^{2}	Np^{125A}	br^{1}/Y
Df(2L)Exel7027	61.3 (106)	ND	ND	ND	ND	6.3 (32)
Df(2R)Exel6068	58.4 (89)	ND	ND	ND	ND	38.6 (44)
Df(3L)Exel6084	95.6 (68)	2.3 (88)	5.6 (72)	6.2 (97)	5.0 (140)	33.3 (39)
Df(3R)red-P52	ND	ND	ND	79.4 (63)	80.0 (55)	ND

Animals are doubly heterozygous, *e.g.* $Rho1^{E3.10}$/+; *Df(3R)red-P52*/+, or + Sb^{70}/*Df(3R)red-P52* +. The numbers shown indicate the percentage of animals with malformed legs, with the total number of animals of the indicated genotype scored shown in parentheses. ND, not determined.

Df(2R)Exel6068

A second genomic region showing interaction with Sb^{63b} is 56A1-56B5. This region contains 16 genes, including the GPCRs *5-HT1A* and *5-HT1B*, and is uncovered by *Df(2R)Exel6068*. Doubly heterozygous animals, *i.e.*, *Df(2R)Exel6068/+; red* Sb^{63b} *e*/+, exhibit 35% malformation (Table 11). This malformation frequency is below the 40% threshold considered significant, but due to the low number of progeny produced in the primary cross, secondary SSNC assays were done to determine the reliability of the data (Table 12). Animals of the genotype *Df(2R)Exel6068/CyO* were crossed to Sb^{70}/*TM6B, Tb Hu e* animals. The resulting doubly heterozygous *Df(2R)Exel6068/+;* Sb^{70}/+ class, which contained three times as many animals as the class of interest in the original screen, exhibits 58% leg malformation. Further evidence that *Df(2R)Exel6068* uncovers an interacting gene comes from the moderate interaction observed for *Df(2R)Exel6068* in the presence of br^{1}. Thirty-nine percent of br^{1}/Y;

Df(2R)Exel6068/+ animals exhibit leg malformation, suggesting that an interacting gene(s) contained within the region uncovered by *Df(2R)Exel6068* is involved in ecdysone signaling. It will be interesting to determine the strength of genetic interactions between *Df(2R)Exel6068* and *Rho1* and *Np* alleles.

Df(3L)Exel6084

A third deficiency which shows very strong interaction with Sb^{63b} is *Df(3L)Exel6084*. This stock carries a deletion of genomic region 61B2-61C1, which completely removes 43 genes, including the GPCRs methuselah-like (mthl) 9, mthl10 and mthl14. Mutant *Df(3L)Exel6084* animals exhibit 80% leg malformation in *trans*-heterozygous condition with Sb^{63b}(Table 11). The high frequency of malformation observed implies a very strong interaction between *Stubble* and a gene(s) within the deleted region, thus a number of secondary experiments have been done to determine which deleted gene product interacts with Stubble and to further define the role of the deleted region in leg development (Table 12).

To verify the strength of the interaction observed between the deficiency chromosome and *Stubble*, *Df(3L)Exel6084* animals were crossed to mutant Sb^{70} animals. The *trans*-heterozygous *Df(3L)Exel6084/*Sb^{70} animals exhibit 96% leg malformation, consistent with the data obtained with Sb^{63b}. To test for further interaction with ecdysone-responsive genes, *Df(3L)Exel6084* animals were crossed to br^{1} mutants. From the br^{1} assay, a moderate malformation frequency of 33% was observed for progeny of the genotype br^{1}*/Y; Df(3L)Exel6084/+*. Interestingly, only very weak interactions were observed between *Df(3L)Exel6084* and $Rho1^{E3.10}$ or $Rho1^{J3.8}$ (2% and 6%, respectively) or between *Df(3L)Exel6084*

and the *zipper* allele, zip^{Ebr} (12%; data not shown). Similarly, weak interactions were observed for *Df(3L)Exel6084* in combination with the *Notopleural* alleles Np^{125A} and Np^{2} (5% and 6%, respectively). Collectively, these data raise the possibility that the interacting gene(s) uncovered by *Df(3L)Exel6084* may cooperate with *Stubble* to regulate leg development through a pathway independent of *Notopleural* or *Rho1*.

In an attempt to identify the interacting gene(s) uncovered by *Df(3L)Exel6084*, nine additional deficiencies which overlap the 61B2-61C1 genomic region were tested for SSNC with Sb^{63b} (Table 13; Figure 10). Of the nine deficiencies tested, two, *Df(3L)BSC128* and *Df(3L)Exel9057*, show significant interactions (79% and 51%, respectively) in combination with Sb^{63b}. This suggests that there are two interacting genes within the deleted region.

Several inconsistencies are apparent from the overlapping deficiency map shown in Figure 10. Based on the published breakpoints, *Df(3L)ED201*, which shows no interaction with Sb^{63b}, encompasses both strongly interacting deficiencies as well as the original deficiency line *Df(3L)Exel6084*. In addition, *Df(3L)BSC126*, which displays moderate interaction in combination with Sb^{63b} (Table 13), completely overlaps with the strong interactor *Df(3L)Exel 9057*. It has been our experience that the published breakpoints for many deficiencies are unreliable, thus we are currently defining the molecular breakpoints of the deficiencies shown in Figure 10 using molecularly defined lethal P-element insertions and mutations to assay for non-complementation when in combination with the various deficiencies. The deficiencies overlapping *Df(3L)Exel6084* are also being mated to one another to determine whether the genomic overlaps are real (Table 14). Though only a few of these experiments have been completed, several informative pieces of data have been obtained. For example,

Df(3L)BSC128/Df(3L)7C animals are viable, indicating that the left boundary of *Df(3L)7C* does not likely extend as far to the left as indicated by the dashed line (region of uncertainty) in Figure10. The *Df(3L)ED201/Df(3L)7C* combination, however, is lethal, indicating that, though the left breakpoint of *Df(3L)7C* may be much farther to the right than indicated, there is still some region of overlap between *Df(3L)7C* and *Df(3L)ED201*. *Df(3L)ED201* also exhibits lethality in combination with *Df(3L)BSC128*. This confirms the predicted overlap between these two deficiencies but does not explain how *Df(3L)BSC128* exhibits strong interaction with Sb^{63b} while *Df(3L)ED201* shows no interaction with the same allele. Additional assays are necessary to resolve this issue. As predicted by their molecular breakpoints, none of the deficiencies *Df(3L)BSC121*, *Df(3L)BSC128* or *Df(3L)BSC125* is lethal in combination with *Df(3L)Exel9057*. These data confirm that there is likely no region of overlap between the these BSC deficiency chromosomes and *Df(3L)Exel9057*. In contrast, *Df(3L)Exel9057* shows partial lethality in combination with *Df(3L)BSC126*. Interestingly, the few animals that survive to adulthood show no malformation. This is consistent with the semi-lethal phenotype of homozygous *Df(3L)Exel9057* animals.

Table 13. Deficiencies overlapping Df(3L)Exel6084

Deficiency	Genomic Region Uncovered	% mlf with Sb^{63b}
Df(3L)BSC126	61B3-61C1	26.6
Df(3L)BSC125	61B3	25.0
Df(3L)BSC128	61B2	79.4
Df(3L)Exel6083	61A6-61B2	3.2
Df(3L)Exel9057	61C1	51.2
Df(3L)ED201	61B1-61C1	0.0
Df(3L)7C	61B2-61C	5.3
Df(3L)BSC121	61B2	35.8
Df(3L)ED4079	61A5-61B1	0.0

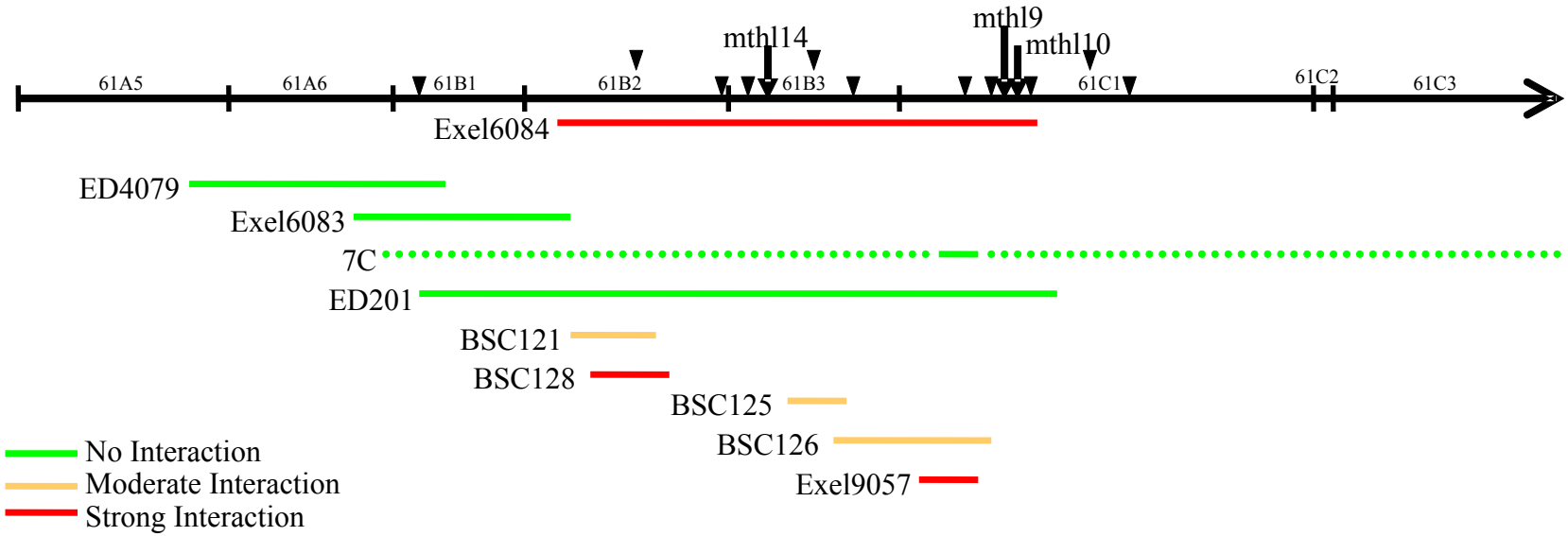

Figure 10. Deletion map showing breakpoints for Df(3L)Exel6084 and overlapping deficiencies.

The scale along the top shows the genomic region. Positions of the 3 GPCRs uncovered by *Df(3L)Exel6084* are denoted with arrows. Positions of molecularly mapped lethal P-element insertions and mutations are denoted by arrowheads. Molecular breakpoints have not been determined for 7C, thus the extent of the deleted region is unknown. The area of uncertainty is indicated with dashed lines.

Table 14. SSNC analyses of deletions overlapping Df(3L)Exel6084

	Df(3L)7C	Df(3L)Exel6084	Df(3L)ED201	Df(3L)BSC128	Df(3L)Exel9057	Df(3L)BSC125	Df(3L)BSC126
Df(3L)Exel6084	ND						
Df(3L)ED201	lethal	ND					
Df(3L)BSC128	0 (40)	ND	lethal				
Df(3L)Exel9057	ND	ND	ND	0 (141)			
Df(3L)BSC125	ND	ND	ND	ND	0 (19)		
Df(3L)BSC126	ND	ND	ND	ND	semi-lethal	ND	
Df(3L)BSC121	ND	ND	ND	ND	2.2 (45)	ND	ND

Animals are *trans*-heterozygous, *e.g.*, *Df(3L)Exel6084/ Df(3L)7C*. The numbers shown indicate the percentage of animals with malformed legs, with the total number of animals of the indicated genotype scored shown in parentheses. ND, not determined.

Df(3R)red-P52

The fourth and final genomic region showing strong interaction with *Stubble* is 88A4-88B2, which is uncovered by *Df(3R)red-P52* (Table 11). This region contains at least 43 genes, including the GPCRs CG9918 and Dopamine Receptor (DopR). Animals of the *Df(3R)red-P52/red* Sb^{63b} *e* class exhibit 88% leg malformation. Interestingly, the malformation observed in this class is not typical of the malformation phenotype normally observed (see Figure 6). All malformed *Df(3R)red-P52/red* Sb^{63b} *e* animals have at least one excessively bowed tibia on the third set of legs that twists slightly to the outside of the body such that it crosses the femur when the legs are folded. The femurs, however, are all of the wild-type length and shape. Due to its subtle nature, this phenotype would generally be scored as wild-type when seen in an individual animal. However, since almost 90% of animals of the genotype *Df(3R)red-P52/red* Sb^{63b} *e* exhibit this phenotype, the deficiency was considered worthy of additional experiments.

In an attempt to further define the nature of the interaction observed between *Df(3R)red-P52* and the *Stubble* locus, deficiency animals were crossed to Np^{125A} and Np^{2} (Table 12). As with Sb^{63b}, *Df(3R)red-P52* shows very strong interaction in combination with both Np^{125A} and Np^{2}, yielding malformation frequencies of 80% and 79%, respectively. The phenotype seen in combination with *Notopleural* mutants is the same as that observed in *Df(3R)red-P52/red* Sb^{63b} *e* animals. Further SSNC and dominant interaction analyses will be performed between *Df(3R)red-P52* and Sb^{70}, $Rho1^{E3.10}$, $Rho1^{J3.8}$ and br^{1}.

CHAPTER FIVE: DISCUSSION

To understand the role that TTSPs play in development, homeostasis and disease progression, it is necessary to understand their mechanism of action. The Stubble protease serves as an excellent model in which to study the mechanism by which TTSPs interact with intracellular signaling pathways to regulate developmental processes. In this study, I have attempted to develop a greater understanding of the mechanism of action of the Stubble TTSP using two distinct approaches. First, I sequenced 12 mutant alleles of *Stubble* to determine which of the conserved domains contained within the protease are necessary for proper function of Stubble. Most of the mutations identified are located within or directly effect the protease domain, indicating that this domain is necessary for function. One mutation was found outside of the protease domain, but the relevance of this mutation to protease function has not been established. Second, I performed a genetic interaction screen of the non-gustatory, non-odorant GPCRs within the *Drosophila* genome in an attempt to find a target of the Stubble protease. Though no individual targets have been identified to date, a few promising leads have been obtained which are currently being investigated further.

Sequencing of *Stubble* mutant alleles

i.) Identification of four *sbd* alleles as sbd^{201}

In this study I show that four of the putative sbd alleles (sbd^{46}, sbd^{173}, sbd^{241} and sbd^{277}) isolated by Ruggiero (2006) are actually sbd^{201} alleles. An overview of the strategy used by Ruggiero (2006) to identify enhancers of the recessive sbd^{201} leg and bristle phenotypes is shown

in Figure 11. In the strategy, a tester stock with the genotype *Pm/CyO; red* sbd^{201} *e/TM6B, Tb Hu e* was mated to mutagenized animals (G1 cross in Figure 11). It is possible that during the subsequent recovery and isolation of the enhancer chromosome, a recombined sbd^{201} chromosome lacking the recessive ebony marker could have been recovered rather than the enhancer of *sbd* chromosome. As a consequence, at one or more points during this screen, the sbd^{*}/sbd^{201} class may have been confused with the sbd^{201}/sbd^{201} class resulting in isolation of the sbd^{201} allele rather than the *sbd** allele. The sequence data support the possibility that the tester and enhancer chromosomes were exchanged: sbd^{201} was produced in an earlier EMS screen which used br^{1} as the progenitor stock (Beaton et al., 1988), and the polymorphisms found in sbd^{46}, sbd^{173}, sbd^{241}, and sbd^{277} are identical to those found in br^{1} and sbd^{201}. The data presented here differ from the findings of Hammonds and Fristrom (2006) in that both sbd^{201} and its br^{1} progenitor lack the del2090_2098 sequence variant in exon 7. This deletion was observed in the w^{1118}*; iso2; iso3* progenitor stock used in Ruggiero's EMS screen but was not found in the sbd^{201} or br^{1} stocks that I sequenced.

G1) Cross mutagenized males to females from the tester stock

((EMS))

$$\frac{w^{1118}}{y} ; \frac{+}{+} ; \frac{+}{+} \quad X \quad \frac{Pm}{Cy} ; \frac{red, sbd^{201}, e}{TM6B, Hu, e}$$

G2) Score all G2 progeny for leg and bristle defects

▶ Isolate all malformed animals, and cross each individually to tester stock to test if malformations are heritable.

$$X \quad \frac{Pm}{Cy} ; \frac{red, sbd^{201}, e}{TM6B, Hu, e}$$

G3) Score G3 progeny to identify line that breed true for mutant phenotypes

▶ Isolate single chromosomes carrying enhancer mutations by crossing to balancer lines

A) Cross malformed animals to second and third chromosome balancer lines

$$\frac{?}{Y} ; \frac{?}{Cy} ; \frac{?}{red, sbd^{201}, e} \quad X$$

i) on the second chromosme

$$\frac{CR2}{sco} ; \frac{+}{+}$$

ii) on the third chromosme

$$X \quad \frac{+}{+} ; \frac{Sb^{1}}{Tb}$$

B) Cross virgin females to males carrying first chromosome balancers

$$\frac{?}{?} ; \frac{?}{Cy} ; \frac{?}{red, sbd^{201}, e} \quad X \quad \frac{M5}{y}$$

C) Mate sibling malformed males and females

G4) Generate stock by crossing siblings carrying the isolated enhancer over a balancer chromosome

Figure 11. A schematic overview of the Ruggiero *sbd*^{201} enhancer screen strategy.
Taken from Ruggiero, 2006.

ii.) Effects of protease domain mutations on Stubble activity

Of the 12 sbd alleles sequenced, 8 carried specific mutations in the protein coding region. In every case, mutations that are very likely to affect Stubble proteolytic function (see below) were observed. In 2 cases additional mutations outside the proteolytic domain, were also identified.

Upon activation of trypsin-like serine proteases, approximately 85% of the structures of the protease domain remain organized identically and in the same conformation in the active peptide as they were in the inactive zymogen. The remaining 15% of the protease domain (the activation domain), which includes the substrate binding site and the oxyanion hole (see below), undergoes a series of conformational changes. The folding of the activation domain is facilitated by four conserved glycine residues which serve as hinges around which activation domain segments move to form the active enzyme. Upon folding, the activation domain becomes ordered while forming a closed domain that is stabilized by a salt bridge between the now-liberated N-terminus of the activation cleavage site isoleucine, Ile16 (Ile544 in Stubble), and aspartic acid 194 (Asp737 in Stubble; Gombos et al., 2008) (Note: When giving coordinates for amino acid residues in the protease domain, the chymotrypsinogen numbering system will be used unless otherwise noted). Formation of the Ile16-Asp194 salt bridge is also necessary to form a structural linkage between the catalytic triad and the oxyanion hole during substrate binding (Figure 12). The oxyanion hole is formed by the backbone amide NHs of Gly193 (Stubble Gly736) and catalytic Ser195 (Stubble Ser738) and forms a pocket of positive charge that works to stabilize the negatively charged oxyanion of an intermediate structure during substrate binding (Matthews

et al., 1975). Incomplete folding of the activation domain in the zymogen results in inactivity of the protease.

The main chain residues 214-216 (Stubble 758-760) make up the substrate binding site and form an antiparallel beta sheet with the backbone P1-P3 residues of the substrate (Figure 13). Hydrogen bonds form between the carbonyl oxygen of Ser214 (Stubble Ser758) and the amide NH of the P1 residue, the amide NH of Trp215 (Stubble Trp759) and the carbonyl oxygen of P2 and the carbonyl of Gly216 (Stubble Gly760) and the amide NH of P3. These interactions are essential for efficient proteolysis of the substrate. In addition to the primary substrate binding site, residues 214-216 form a portion of one wall of the S1 subsite (specificity pocket; Figure 14), and the carbonyl of Ser214 forms a hydrogen bond with the catalytic histidine. These structural interactions form a connection between the substrate binding site, the S1 site, and the catalytic triad (reviewed in Hedstrom, 2002).

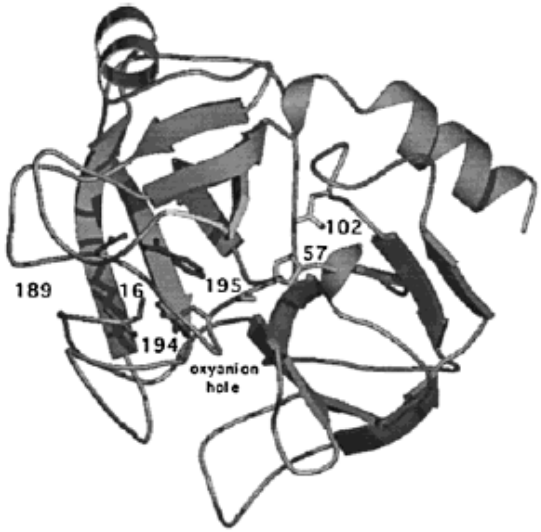

Figure 12. Structure of chymotrypsin showing positioning of the oxyanion hole.

The oxyanion hole is structurally linked to the Ile16-Asp194 (Stubble Ile544-Asp737) salt bridge and the catalytic triad residues His57, Asp102 and Ser195 (Stubble His590, Asp640 and Ser738) via catalytic Ser195 (Stubble Ser738). Figure taken from Hedstrom, 2002.

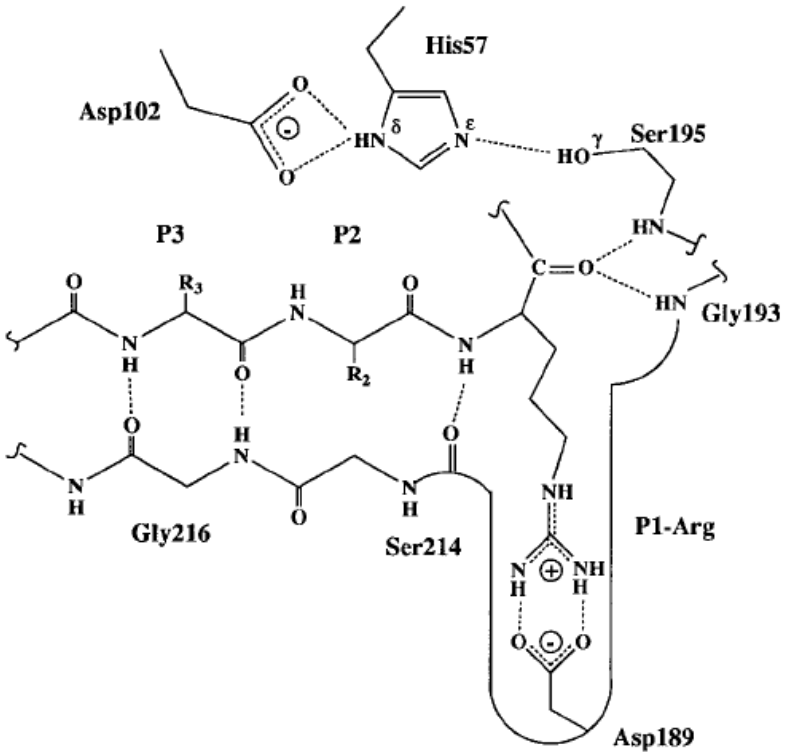

Figure 13. Mechanism of substrate binding for serine proteases.

Catalytic groups of chymotrypsin-fold serine protease are depicted interacting with a substrate at its P1 site (substrates are hydrolyzed at the P1-P1' bond). Five enzyme-substrate hydrogen bonds at positions P1 and P3, including those involving Ser214 and Gly216 (Ser758 and Gly760 in Stubble), are shown in addition to hydrogen bonds among members of the catalytic triad. The interaction of amide NH of Trp215 (Stubble Trp759) with the carbonyl oxygen of the P2 residue is not labeled. Figure and legend taken from Perona and Craik, 1997.

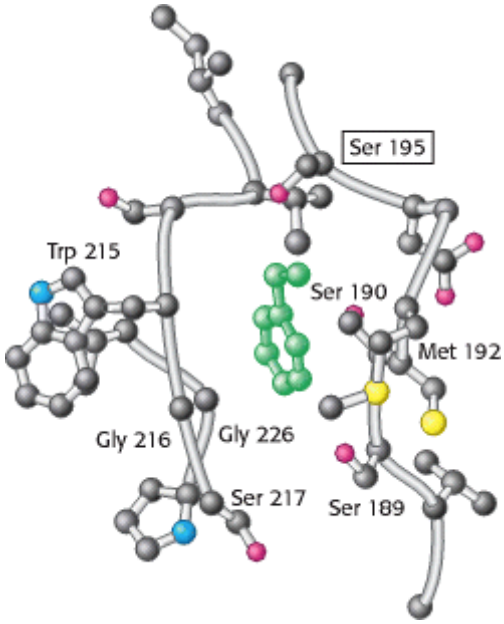

Figure 14. Schematic diagram of the S1 subsite of chymotrypsin interacting with a substrate. The S1 subsite is composed of three walls formed by residues 189-192, 214-216 and 224-228 (732-735, 758-760 and 768-772 in Stubble). Residues 214, 224 and 225 (Stubble 758, 768 and 769) are not labeled. Substrate specificity is usually conferred by amino acids at residues 189, 216 and 226 (Stubble residues 732, 760 and 770). Carbon atoms are shown in gray, oxygen in red, sulfur in yellow and nitrogen in blue. The catalytic serine is boxed. The substrate is shown in green. Taken from Berg, Tymoczko, and Stryer, 2007.

Several of the mutations found in the 12 *sbd* alleles analyzed (see Table 8) affect structures involved in protease activation, folding and substrate binding as discussed above. Given the importance of proper maintenance of these structures on protease function, these mutations are likely to have substantial effects on Stubble proteolytic activity. For example, Glycine 760, which functions as one of the four glycine hinge residues necessary for proper activation domain folding, is substituted for a serine residue in $sbd^{E(br)228}$. It has been proposed that restriction of flexibility around these hinges by amino acid substitution may lead to a zymogen-like conformation in the active protease (Gombos et al., 2008). The necessity of

functional glycine hinges for enzymatic activity is exemplified by the finding that substitution of any of the hinge glycines in blood coagulation-related serine protease factors VII, IX and X leads to bleeding disorders (Bernardi et al., 1996; Giannelli et al., 1994; Uprichard et al., 2002; Wulff et al., 2000).

Both $sbd^{E(br)20}$ and $sbd^{E(br)228}$ contain mutations in highly conserved residues that directly interact with the substrate during binding (Huber and Bode, 1978). In the allele $sbd^{E(br)20}$, serine 758, which is involved in direct binding with the P1 residue of the substrate, is altered to a phenylalanine. In $sbd^{E(br)228}$, glycine 760 is altered to a serine residue. Gly760 is involved in direct binding with the P3 residue of the substrate. Both of these changes are nonconservative. In the case of S758F, a polar residue is replaced with a nonpolar residue, and a nonpolar residue is replaced with a polar residue as a result of the G760S mutation. A shift between nonpolar and polar amino acids in each case could result in misfolding of the substrate binding pocket. Nonpolar amino acids tend to face the interior of the folded molecule while polar residues tend to cluster to the surface where interactions with water molecules in the surrounding environment are maximized. Thus the substitution of Ser758 for phenylalanine in $sbd^{E(br)20}$ as well as the substitution of Gly760 for serine in $sbd^{E(br)228}$ potentially interfere with normal Stubble function be inhibiting substrate binding and disrupting intramolecular interactions between various regions of the catalytic domain.

sbd^{PNR11} contains a point mutation which results in the substitution of aspartic acid 737 for asparagine. Aspartic acid 737 is involved in the formation of the Ile544-Asp737 salt bridge (Ile16-Asp194 in chymotrypsinogen numbering). This salt bridge stabilizes the folded activation domain and is necessary for linkage of the catalytic triad with the oxyanion hole during substrate

binding (see Figure 12). Disruption of this salt bridge due to the substitution of asparagine for aspartic acid at amino acid 737 is predicted to inhibit proper stabilization of the active molecule and lead to a zymogen-like conformation with greatly diminished substrate binding capabilities.

The *sbd*^{258} mutant allele contains a single amino acid substitution, G755E, in the protease domain that alters the normal activity of the Stubble protein. Glycine 755 is a highly conserved residue in a highly conserved region of the protease domain. This glycine residue is located 3 residues amino-terminal to the polypeptide binding site but is not known to be directly involved in substrate binding. The substitution of a nonpolar, hydrophobic glycine residue for a polar, hydrophilic glutamic acid residue in such close proximity to the substrate binding pocket could inhibit proper folding of the substrate binding pocket and thus interfere with the ability of the Stubble protease to bind and cleave substrates properly.

In addition to the specific mutations found within the various structures of the protease domain, several *sbd* alleles contain truncations which remove a portion or the whole of the protease domain. *sbd*^{206} and *sbd*$^{E(br)448}$ both contain single point mutations which result in the change of a glutamine residue for a stop codon. The *sbd*^{206} allele contains a nonsense mutation at residue 445 in the stem of the Stubble protein (see Figure 9). This substitution leads to truncation of the protein within the stem, completely eliminating the protease domain. The *sbd*$^{E(br)448}$ allele contains a nonsense mutation at residue 714, twenty-four residues N-terminal to the catalytic triad serine residue in the protease domain (see Figure 9). This substitution results in a truncation which removes the catalytic serine as well as the substrate binding domain. A third allele, *sbd*^{266}, also causes a truncation of the protein, but it does not involve a nonsense mutation. This allele contains a threonine-to-asparagine amino acid substitution at residue 587 (see Figure 9) which is

directly followed by an out-of-frame 292-bp deletion. The frameshift results in the addition of 14 incorrect residues onto the modified asparagine before reading into a stop codon. This deletion/frameshift mutation effectively removes the majority of the protease domain, including all three residues of the catalytic triad and the substrate binding pocket.

In addition to an amino acid substitution and an out-of-frame deletion, sbd^{266} also contains an amino acid substitution within the Stubble stem region. The proline residue at position 137 is altered to a leucine. This proline residue directly precedes the disulfide-knotted domain, and a substitution could potentially cause improper folding and disruption of the flanking disulfide-knotted domain. As this allele also contains a deletion that removes the majority of the protease domain, it is unclear to what extent the P137L amino acid substitution contributes to the sbd^{266} phenotype.

The $sbd^{E(br)623}$ allele contains a 24-bp deletion in the intron separating exons 10 and 11 in the proteolytic domain (see Table 8 and Figure 9). The deleted portion of the intron contains the lariat branch point consensus sequence CTAAT, in which branch formation occurs at the underlined A (Mount et al., 1992; Figure 15). Splicing events begin with assembly of the spliceosome onto the pre-mRNA by binding sequences located at the 5' and 3' ends of the intron (reviewed in Graveley, 2004). This assembly is initiated by association of the U1 small nuclear ribonucleoprotein (snRNP) with the 5' splice site, the branch-point-binding-protein/SF1 with the branch point, and the U2 snRNP auxiliary factor (U2AF) with the pyrimidine tract. Stable association of the U2 snRNP at the branch point and splice-site pairing then occurs upon ATP hydrolysis. Few mutations of branch point sequences have been identified, but those that have been found have resulted in cryptic splicing (Rosenthal et al., 1992; Brand et al., 1996; Fujimaru

et al., 1998), intron retention (Kuivenhoven et al., 1996) and exon skipping (Putman et al., 1997; Burrows et al., 1998; Hamlington et al., 2000). Therefore, the splicing defect in *sbd*$^{E(br)623}$ is likely to alter the C-terminus of the protease domain. If the intron is retained as a result of branchpoint deletion, 31 nucleotides would be added between exons 10 and 11. This addition would cause a frameshift reading into exon 11, resulting in the addition of 51 new amino acids before reading into a stop codon. Exon 11, which is C-terminal to all members of the catalytic triad (see Figure 9), contains the substrate binding site as well as the S1 specificity pocket. The frameshift would disrupt both of these structures and thus result in disruption of normal protease activity.

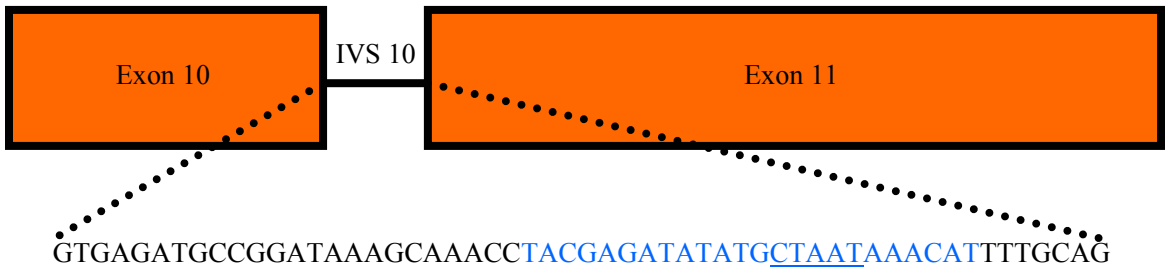

Figure 15. Schematic of the *sbd*$^{E(br)623}$ intronic mutation.

Exons 10 and 11 are shown as orange boxes with intron, or intervening sequence (IVS), 10 as a connecting line. The intronic sequence is expanded below: the deleted 24-bp segment is in blue and the lariat branch point consensus sequence is underlined.

The final mutation identified, P662L, is found in the *sbd*^{VE3} allele. This proline residue is very highly conserved across serine proteases and is found in a highly conserved region of the protease domain. Two residues N-terminal to this residue is cysteine 660, which forms a salt bridge with cysteine 532 within the stem region. This disulfide bridge tethers the protease domain to the stem region upon activation cleavage. Altering the proline residue to a lysine could result in disruption of this disulfide bridge by misfolding the protein such that the thiol groups of the cysteine residues are no longer able to form a covalent linkage. This would result in shedding of the protease domain from the protease stem region and the cell surface upon activation cleavage.

iii.) Apparent *sbd* regulatory mutations

Of the 12 sbd alleles sequenced, *sbd*^{1}, *sbd*$^{E(br)48}$ and *sbd*$^{E(br)536}$ have no changes within the coding region that would obviously affect protease function. These alleles are therefore presumed to be regulatory in nature, although this would have to be confirmed with further experimentation such as steady state RNA or protein level analysis, etc.

Two polymorphisms were found in *sbd*^{1} that are not present in the other *sbd* alleles analyzed. The first change, 1025C>T, is a synonymous change altering a leucine codon from CTA to TTA. Neither of these is a preferred leucine codon in *Drosophila melanogaster*, but the alteration from CTA to TTA results in a change to the rarest codon for leucine (Codon Usage Database; http://www.kazusa.or.jp/codon/). The second change, 2812G>A, results in alteration of the preferred CAG glutamine codon to the less frequently used CAA glutamine codon (Akashi, 1995). Thus these changes, while silent with respect to amino acid substitution, could potentially affect translation of the Stubble polypeptide through use of non-preferred codons.

Adzhubei et al. (1996) suggested that the presence of a rare non-preferred codon in the mRNA can be associated with mis-folding of the nascent polypeptide. A recent finding by Kimchi-Sarfaty et al. (2007) supports this hypothesis. They showed that the human *Mutlidrug Resistance* 1 (*MDR*1) gene containing a single nucleotide silent polymorphism is associated with altered conformation and substrate specificity of the *MDR1* gene product P-glycoprotein. The polymorphism, which was previously linked to altered function of P-glycoprotein, results in usage of a rare synonymous codon.

iv.) Future investigations of *Stubble* mutations identified in these studies

To date, the majority of, if not all, published mutational studies of type II transmembrane serine proteases have been done in human patients affected by a number of diseases. Such studies allow researchers to draw associations between mutations in TTSPs and disease, but they do not allow for direct functional analysis of proteases harboring disease-associated mutations in an *in vivo* system. However, taking advantage of the multitude of resources that are available in *Drosophila*, we can express a protease containing a mutation of choice and subsequently analyze the direct effect of that mutation on protease function *in vivo*.

We plan to use two assays, a leg malformation assay and a bristle defect rescue assay, to determine the direct effects of a set of specific mutations on Stubble function during epithelial morphogenesis. Using site-directed mutagenesis, I generated 12 individual mutations (Table 15) in a full-length wild-type *Stubble* cDNA and placed them under the control of a heat-shock promoter in the vector pCaSpeR-hs (Thummel and Pirrotta, 1991). These mutant constructs will be used to make transgenic animals which, upon exposure to high temperatures, will upregulate expression of the mutant protein.

Previously, Bayer et al. (2003) showed that a 1-hr induction of a wild-type *hs-Stubble* transgene at 0 hr and 3 hr after pupariation (AP) results in a high percentage of severe leg malformation in adults and determined that the leg malformation associated with overexpression of Stubble is specific to a critical period of leg development between 0 and 10 hours AP. These data provide us with an assay which allows for analysis of protease function in animals expressing the mutant *hs-Stubble* transgenes. Staged 0 hr or 3 hr prepupae homozygous for the mutant *hs-Stubble* constructs will be exposed to a constant temperature of 37° for 1 hour. If Stubble activity is retained at significant levels upon introduction of the mutant protein, adult animals are expected to have an elevated frequency of leg malformation compared to control animals in which the transgene is not induced. If, on the other hand, the mutation severely reduces Stubble activity, few malformed adults will be observed. One caveat to the interpretation of these experiments is that a mutation that causes an increase in Stubble activity will also result in a high frequency of malformed adults and would therefore be difficult to distinguish from mutations that do not significantly affect Stubble activity.

Table 15. Mutant constructs for leg malformation and bristle defect rescue assays

Mutant Constructs	Description
H573R	Alters histidine 573 to arginine (CAC to CGC)
P137L	Alters proline 137 to leucine (CCA to CTA)
G760S	Alters glycine 760 to serine (GGC to AGC)
G760R	Alters glycine 760 to arginine (GGC to CGC)
S549I	Alters serine 549 to isoleucine (AGC to ATC)
S549N	Alters serine 549 to asparagine (AGC to AAC)
Cys2→Ser	Alters cysteine 147 to serine (TGC to TCC)
Cys3→Ser	Alters cysteine 153 to serine (TGC to TCC)
Cys5→Ser	Alters cysteine 173 to serine (TGC to TCC)
Met1→Leu	Alters methionine 1 to leucine (ATG to TTG)
Met2→Leu	Alters methionine 24 to leucine (ATG to TTG)
S49N	Alters serine 49 to asparagine (AGC to AAC)

To confirm interpretations based on the leg malformation assay, a second assay, which looks for rescue of the *sbd* bristle phenotype, will be employed. Hammonds and Fristrom (2006) showed that the recessive bristle phenotype of the mild hypomorphic mutant, sbd^2, could be rescued completely with a wild-type *hs-Stubble* transgene upon a 1-hr induction at 37° at 27.5 hr AP. This assay can also be used for analysis of protease function in animals expressing the mutant *hs-Stubble* transgenes. For this assay, mutant transgenes will be placed in a homozygous sbd^2 background. The resulting *hs-Stubble*/hs-Stubble*;* sbd^2/sbd^2 animals will be staged to 27.5 hr AP, at which point they will be subjected to a 1-hr heat-shock induction at 37°. If the introduced mutation causes no significant change from the normal level of Stubble activity, 100% of adult animals should exhibit normal bristles (i.e. the short bristle phenotype of sbd^2 homozygotes will be rescued). If the mutation reduces Stubble function, incomplete, or no rescue, of the bristle defect is expected. As with the leg malformation assay, mutations that

increase Stubble activity cannot be easily distinguished from those with close to normal activity in this assay. Nevertheless, these assays should provide a useful tool to determine the effect of a given mutation on proteolytic activity *in vivo*.

Twelve individual site-directed mutant constructs of the Stubble protease will be tested for function in the leg malformation and bristle rescue assay systems (Table 15). These constructs have been designed to answer a variety of questions. For example, the H573R construct mimics the mutation found in the sbd^{201} allele that has been attributed by Hammonds and Fristrom (2006) to the *sbd* phenotype observed. This construct will allow us to determine whether the H573R mutation reduces function of the protease *in vivo*.

The P137L mutation, which affects the proline residue directly preceding the disulfide knotted domain, was found in the sbd^{266} allele. This allele also contains a large deletion that effectively removes the majority of the protease domain. Thus, it is unclear whether the P137L mutation plays any role in altering the function of the Stubble protease in the sbd^{266} mutant. Using a P137L construct will allow us to test whether this mutation, which could potentially disrupt folding of the disulfide knotted domain, is capable of reducing Stubble function.

Mutation of the human matriptase equivalent of Stubble glycine 760 has been implicated in autosomal recessive skin disease. Basel-Vanagaite et al. (2007) identified a missense mutation which alters this glycine residue to an arginine residue (G827R in matriptase) in patients exhibiting ichthyosis with hypotrichosis (ARIH) syndrome. This mutation has been shown to result in inactivation of matriptase in an *in vitro* system (Désilets et al., 2008). The Stubble equivalent of G827 in matriptase is also mutated in the *Stubble* mutant allele $sbd^{E(br)228}$. In contrast to the G827R mutation in matriptase, Stubble Gly760 is substituted for a serine residue.

Because the resulting amino acids differ, I have made two mutant Stubble constructs affecting Stubble Gly760. The first, G760S, mimics the mutation found in *sbd*$^{E(br)228}$. The second, G760R, mimics the mutation found in human *ST14*, the gene encoding matriptase. These constructs will allow us to determine whether different types of substitutions at this position affect protease function differently and, more importantly, to determine in an *in vivo* system whether the Stubble G760R equivalent of the human matriptase G827R mutation yields an inactive protease.

The mutation N29I in human trypsinogen, the inactive form of trypsin, has been linked to pancreatitis (Gorry et al., 1997). In the Stubble protein, the equivalent residue is Ser549. Because mutation at this non-conserved amino acid position has been linked to disease, I decided to alter Stubble Ser549 to asparagine and isoleucine. These changes mimic the wild-type residue in human trypsinogen (N) and the mutant residue in human trypsinogen linked to pancreatitis (I). The S549N and S549I constructs will allow us to determine whether asparagine can functionally substitute for serine at position 549 in the Stubble protease, and to determine if isoleucine at this position reduces *in vivo* activity, thus lending support to the data linking the N29I mutation to pancreatitis.

To directly test whether the disulfide knotted domain is necessary for Stubble activity, I have made three mutant constructs which alter various cysteine residues within the disulfide knotted domain to serine residues. These changes are based on a previous study in the *Drosophila* snake protease in which mutations were introduced into the disulfide knotted domain of this protease (Smith et al., 1994). The three cysteine residues altered were chosen based on the fact that each is involved in formation of a separate salt bridge necessary for proper folding of the disulfide knotted domain. By altering each salt bridge individually, we can assay the

requirement for the salt bridge on Stubble protease activity. If the disulfide knotted domain is necessary for Stubble function, disrupting any of the salt bridges is predicted to lead to inactivity.

Two putative start codons are present in the 5' region of the Stubble protease. Though the second has a slightly stronger Kozak consensus, the first is generally given as the start site. Because it is unknown which of these methionine residues is used for the start of translation, I have made two mutant constructs in which one of the two residues is changed to a leucine residue. Alteration of the start codon, whether it be the ATG encoding Met1 or the ATG encoding Met2 (amino acid position 24), should eliminate activity of the Stubble protease unless the alternate ATG codon can be recruited as the start codon. In the former case, a reduction of activity would be expected. If both Met codons can be used to initiate translation, no significant effect on Stubble protease activity is expected.

Finally, the S49N construct targets a putative PKC phosphorylation site in the cytoplasmic domain. The PKC phosphorylation site is one of three motifs found within the cytoplasmic domain of the Stubble protease. All three motif types, PKC phosphorylation, cAMP phosphorylation, and N-myristoylation, have a high probability of occurrence based on the consensus patterns used for identification. Although mutations affecting the cytoplasmic domain were not observed in the sequenced alleles, the possibility that this domain is a substrate for a kinase should be resolved. The introduced mutation alters the serine at position 49 to an asparagine, thus eliminating the putative PKC phosphorylation site. Therefore, if this domain is necessary for function, it is likely that alteration of this site will inhibit function of the protease.

v.) Summary of molecular and genetic analysis of *sbd* alleles investigated in this study

I was interested to know if there is a correlation between the type of molecular lesion and the genetic characteristics of the *sbd* alleles tested. The data obtained from sequencing experiments, complementation assays and lethal phase analyses are summarized for each *sbd* allele in Table 16. At least one trend is apparent within individual datasets. All mutations resulting in truncation of the protease domain or direct disruption of substrate binding exhibit predominantly larval lethality with no animals reaching the pupal stage. However, when considering mutation type, genetic interaction with Rho1 pathway members and lethal phase together, no obvious correlations emerge.

Table 16. Summary of genetic data for the mutant *sbd* alleles

Allele	Molecular Lesion	Strength of interaction with					Lethal Phase
		Rho1^{720}	*Rho1*$^{J3.8}$	*Rho1*$^{E3.10}$	*Rho1*^{k02107}	*zip*^{Ebr}	
sbd^{1}	Putative Regulatory	None	None	None	None	Weak	Broad Spectrum
sbd^{206}	Q446X	None	None	Weak	Weak	Moderate	Larval
sbd^{258}	G755E	None	None	None	None	Moderate	Embryonic
sbd^{266}	P137L; T587N followed by 292-bpΔ and frameshift + 14aa then stop codon	None	None	None	None	Weak	Larval
sbd$^{E(br)20}$	S758F	None	None	None	Weak	Weak	Larval
sbd$^{E(br)48}$	Putative Regulatory	None	None	None	None	Weak	Larval w/9% pupariating
sbd$^{E(br)228}$	G760S	None	None	None	None	Weak	Pupal w/34% larval
sbd$^{E(br)448}$	Q714X	None	None	None	None	Weak	Larval
sbd$^{E(br)536}$	Putative Regulatory	None	Weak	Weak	Weak	Weak	Larval
sbd$^{E(br)623}$	IVS10:Δlariat consensus	None	Weak	None	None	Weak	Larval w/15% pupariating
sbd^{PNRII}	D737N	None	None	None	None	Moderate	Larval
sbd^{VE3}	P662L	None	None	None	None	None	Larval w/29% embryonic

Strength of interaction is defined for the doubly heterozygous class, *i.e. Rho1*^{720}*/+; sbd*^{1}*/+*, as none for a malformation rate of less than 5%, weak for a malformation rate of 5%-24%, moderate for a malformation rate of 25%-49% and strong for a malformation rate greater than 50%.

GPCRs and other candidate regulators of epithelial morphogenesis

Four genomic regions exhibiting genetic interaction with the Stubble protease in control of epithelial morphogenesis have been identified through the primary screen of the non-odorant, non-gustatory G protein-coupled receptors in the *Drosophila* genome. Combined, these regions uncover at least 123 genes including 8 GPCRs, as well as a number of other genes that could potentially play a cooperative role with Stubble and the Rho1 pathway in regulation of actin cytoskeletal dynamics.

A total of 21 annotated genes are contained within the region uncovered by *Df(2L)Exel7027* (see Appendix B), including the GPCR gonadotropin-releasing hormone receptor (GRHR). The *Drosophila* GRHR is evolutionarily related to the mammalian GRHR, which is involved in various aspects of reproduction (Hauser et al., 1998). *Drosophila* microarray data from whole animals suggests that GRHR is most highly expressed throughout larval development and during the first 8 hours of metamorphosis (Arbeitman et al., 2002). This expression pattern overlaps that of Stubble and corresponds to the developmental period during which the leg discs undergo morphogenesis.

The genomic region uncovered by *Df(2R)Exel6068* contains 16 annotated genes (see Appendix B). Two of these, *5-HT1A* and *5-HT1B*, are G protein-coupled serotonin receptors known to be involved in regulation of sleep (Yuan et al., 2006) and circadian rhythms (Yuan et al., 2005), respectively. Both receptors are expressed during late embryogenesis in the central nervous system (Saudou et al., 1992), but complete expression profiles have not been determined.

A third deficiency which shows very strong interaction with Sb^{63b} is *Df(3L)Exel6084*. This deficiency deletes genomic region 61B2-61C1, which completely removes 43 annotated genes (see Appendix B), including methuselah-like (mthl) 9, mthl10 and mthl14. These are three structurally related GPCRs potentially involved in aging and stress response (Brody and Cravchik, 2000). Other interesting genes uncovered by *Df(3L)Exel6084* are RhoGEF3 and the serine-type endopeptidase inhibitors Kaz1-ORFA and Kaz1-ORFB. Also deleted is Enhancer of bithorax (E(bx)), an ATP-dependent chromatin remodeling complex that must bind the Ecdysone Receptor (EcR) in order for ecdysone-responsive genes to be expressed (Badenhorst et al.,2005).

A secondary screen using overlapping deletions to further identify the gene(s) within *Df(3L)Exel6084* which interact with the Stubble protease was also performed (see Table 13 and Figure 10). This screen identified two much smaller regions uncovered by *Df(3L)BSC128* and *Df(3L)Exel9057* which split the larger *Df(3L)Exel6084* into two separate interacting regions. *Df(3L)BSC128* has molecular breakpoints 184799-212410 on chromosome 3L. This region uncovers the entire coding sequence of 7 genes and partially uncovers 2 additional genes (Table 17). The GPCR mthl14 is just outside the right breakpoint of the *BSC128* deficiency and could well be within the deleted region if the right breakpoint of *Df(3L)BSC128* is inaccurate. *Df(3L)Exel9057* has molecular breakpoints 306168-319594 on chromosome 3L, a region which contains only 2 genes, four wheel drive (fwd) and CG34264, and fully deletes the coding domain of each (Table 17). The GPCRs mthl9 and mthl10 are not deleted by this deficiency.

None of the genes lying within the breakpoints of *Df(3L)BSC128* or *Df(3L)Exel9057* stand out as obvious candidates to interact with Stubble in control of epithelial morphogenesis. It is possible that one of these deficiencies removes the regulatory region of a more likely

candidate located outside of the deficiency region. Alternatively, the given breakpoints for these deficiencies may be inaccurate. In the latter case, the list of genes uncovered by *Df(3L)BSC128* and *Df(3L)Exel9057*, as presented in Table 17, may not be comprehensive.

Table 17. Genes uncovered by *Df(3L)BSC128* and *Df(3L)Exel9057*

Df(3L)BSC128		
Gene	Molecular Position	Molecular Function
Mitochondrial carrier homolog 1 (Mtch)	184,635-186,317	transmembrane transporter activity; binding
rhinoceros (rno)	188,549-198,587	DNA binding; protein binding; zinc ion binding
mrityu (mri)	200,180-202,165	glycerol kinase activity; protein binding; voltage-gated potassium channel activity
Glycerol kinase (Gyk)	202,094-205,819	glycerol kinase activity
Nitrilase and fragile histidine triad fusion protein (NitFhit)	203,699-205,311	bis(5'-adenosyl)-triphosphatase activity; hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds; nitrilase activity
CG13876	206,192-207,112	amine oxidase activity; quinone binding; copper ion binding
CG7028	207,884-211,462	protein kinase activity; protein serine/threonine kinase activity; ATP binding
thoc7	211,636..212,687	unknown
CG16940	212,275..216,483	exoribonuclease II activity; RNA binding
methuselah-like 14 (mthl14)	247,495-250,314	G protein-coupled receptor activity
Df(3L)Exel9057		
four wheel drive (fwd)	306,207-319,880	1-phosphatidylinositol 5-kinase activity
CG34264	317,208-317,621	unknown
methuselah-like 9 (mthl9)	329,820-332,306	G protein-coupled receptor activity
methuselah-like 10 (mthl10)	333,475-336,682	G protein-coupled receptor activity

GPCRs uncovered by *Df(3L)Exel6084* are highlighted in blue. None of these GPCRs are uncovered by *Df(3L)BSC128* or *Df(3L)Exel9057*, but they are listed to show their proximity to these deficiencies.

Finally, the genomic region uncovered by *Df(3R)red-P52* contains at least 43 annotated genes (see Appendix B). These include two GPCRs, CG9918 and Dopamine Receptor (DopR). Interestingly, Srivastava et al. (2005) have recently shown that DopR functions as a cell-surface GPCR which may be responsible for some of the rapid actions of ecdysteroids during development. It has become apparent that the rapid action of ecdysteroids, such as ecdysone, is mediated by mechanisms that alter the levels of secondary messengers which in turn may be mediated by the activation of GPCRs. The *Drosophila* DopR protein is activated by dopamine to increase cAMP levels and activate the phosphoinositide 3-kinase pathway. However, the effects of dopamine can be inhibited by the binding of ecdysone to the receptor. Once bound, ecdysone couples the receptor to rapid activation of the mitogen-activated protein kinase pathway.

Additional genes uncovered by *Df(3R)red-P52* are CG31326, CG9649, Sp212 and CG9631, a cluster of four serine-type endopeptidases. It has been suggested that TTSPs may play a general role as initiators of protease cascades. This is supported by the fact that certain TTSPs, such as matriptase, have been shown to undergo autoactivation *in vitro* (Oberst et al., 2003; Takeuchi et al., 1999). Additionally, all TTSP family members exhibit a strong preference for substrates containing an arginine or lysine at their cleavage sites, and have these residues at their own cleavage sites suggesting that they may be capable of autoactivation. More support for this hypothesis is provided by the discovery that matriptase can activate pro-urokinase plasminogen activator (pro-uPA) bound to urokinase plasminogen activator receptor (uPAR) thus activating the plasminogen cascade (Kilpatrick et al., 2006). Interestingly, mutant alleles of *Stubble* interact with *Notopleural* deficiencies which also uncover a cluster of serine proteases (Ruggiero, 2006).

Broader impacts

Misexpression of TTSPs has been implicated in a number of pathologies, thus understanding how they function is of great importance. Using Stubble as a model in which to study the functional mechanism of TTSPs presents the possibility of defining how extracellular hormonal hierarchies, TTSPs and intracellular signaling pathways interact *in vivo*. Such a model of interaction could then be applied toward the understanding of similar interactions in other systems. The results of the studies presented here help to bring us a step closer to understanding the mechanistic role of the Stubble protease in epithelial development, and thus to understanding the mechanism of action of the TTSP family. Through sequence analysis of twelve mutant alleles of *Stubble*, the protease domain was identified as essential for proper protease function. Though only the protease domain has been clearly identified as essential, the remaining conserved domains cannot be ruled out as necessary components of the functional protein. Further studies, such as the leg malformation and bristle defect assays using transgenic animals carrying mutations in the remaining domains, are necessary to determine the contribution of these domains to Stubble function. In addition to providing a method with which to test functional necessity of conserved domains, the leg malformation and bristle rescue assays provide us with a system in which to test the *in vivo* effects of various mutations, including those found in human TTSPs implicated in disease, on protease function. Thus, this assay system could prove to be an extremely important tool for providing direct evidence that disease-associated mutations are capable of reducing or eliminating protease function.

Another important step in defining the mechanism of action of the TTSP family is the identification of TTSP substrates. Using a genetic interaction screen focusing on a subset of

GPCRs, I have identified four genomic regions which show significant interactions with *Stubble* mutants. Contained within these regions are a number of genes, including 8 GPCRs, which could serve as a link between Stubble and the Rho1 pathway or alternatively as Stubble targets in an independent pathway which regulates epithelial development. To establish a clearer picture of the types of targets utilized by TTSPs to function in development and homeostasis, further studies are necessary to identify their direct substrates *in vivo*.

APPENDIX A: GPCR MUTATIONS AND DEFICIENCIES TESTED FOR INTERACTION WITH SB^{63B}

Stock #	Deficiency	GPCRs Affected	% malformation with *Sb*^{63b}
BL7700	Df(1)Exel6223	fz3	8.3
BL13583	P{y^{+}w^{+}}fz3^{KG04032}	fz3	0.5
BL9299	Df(1)ED6565	CG4313	not enough data
BL13875	P{y^{+}w^{+}}KG05130	CG4313	0
BL17007	P{EP}moody^{EP1529}	CG4313	3.8
BL5974	Df(1)X12	CG13758	0
BL12523	P{w^{+mGT}=GT1}BG00979	CG13758	0
BL9169	Df(1)ED6712	AlstR	0
BL8956	Df(1)ED6727	CG6986	0
BL7708	Df(1)Exel6234	CG16752	0
BL7710	Df(1)Exel6236	Tre1	4.4
BL945	Df(1)C149	Tre1	0
BL10089	P{w^{+mC}=EP}Tre1^{EP496}	Tre1	3.2
BL9625	Df(1)ED6878	CG12796	0
BL947	Df(1)HA32	fz4	10.3
BL13934	P{y^{+}w^{+}}KG02367	fz4	1.6
BL903	Df(1)V-L3	CG2061	0
BL11291	P{w^{+mC}}CG2061^{EP1537}	CG2061	2.1
BL21192	P{w^{+mC}y^{+}}mthl1^{Ey16157}	mthl1	13.8
BL15318	P{w^{+mC}y^{+}}mthl1^{EY00861}	mthl1	0
BL8954	Df(1)ED7374	CG4875	2.4
BL5271	Df(1)D15	CG4875	5.9
BL18685	PBac{w^{+mC}}CG7536^{f03720}	CG7536	10.5
BL6217	Df(1)RR79	CG7536	0
BL22030	P{Mae-UAS.6.11}CG7536^{GE01344}	CG7536	5.6
BL7762	Df(1)Exel9051	CCKLR-17D1	0.5
BL7761	Df(1)Exel7463	CCKLR-17D1	0
BL7764	Df(1)Exel7464	CCKLR-17D1 and 17D3	11.4
BL13429	P{y^{+}w^{+}}KG01373	CCKLR-17D3	1.1
BL14884	P{y^{+}w^{+}}KG01171	CCKLR-17D3	4.4
BL9351	Df(1)ED7635	D2R	4.5
BL3277	smo^{3}	smo	0
BL16815	P{w^{+mC}y^{+}}CG9643^{EY07345}	CG9643	1.4
BL1567	Df(2L)JS17	CG9643	4.4
BL7500	Df(2L)Exel6014	CG31645	27.5
BL9185	Df(2L)ED343	CG13995	0
BL20304	P{w^{+mC}y^{+}}GRHR^{EY11371}	GRHR	4.8
BL11320	PBac{PB}GRHR^{c03956}	GRHR	8.6
BL7801	Df(2L)Exel7027	GRHR	57.1
BL7817	Df(2L)Exel8024	CG31720	10.8
BL7514	Df(2L)Exel6031	Rh5	16.7

Stock #	Deficiency	GPCRs Affected	% malformation with *Sb*^{63b}
BL6080	Df(2L)noc11	rk	2.8
BL3589	rk^{1}	rk	4.4
BL3590	rk^{4}	rk	not enough data
BL6965	rk^{81f2}	rk	4.7
BL10919	$P\{w^{+tAR}ry^{+t7.2AR}=wA^{R}\}rk^{11P}$	rk	0
BL3212	Df(2L)do1	CG4168	0
BL7831	Df(2L)Exel7063	kek3	21.6
BL7527	Df(2L)Exel6045	CG10481	13.8
BL9683	Df(2R)ED1484	CG14593	2.4
BL21572	$P\{y^{+t7.7}w^{+mC}\}mxr^{DG17503}$	mXr	1.5
BL7862	Df(2R)Exel7096	mXr	4.1
BL7863	Df(2R)Exel8047	Cirl	18.8
BL21398	$P\{w^{+mC}y^{+mDint}\}Cirl^{EY12930}$	Cirl	0.7
BL9410	Df(2R)BSC132	CG30340	22.9
BL6072	$stan^{81f1}$	stan	0.8
BL6969	$stan^{192}$	stan	11.1
BL6967	$stan^{fz3}$	stan	0
BL6969	$stan^{192}$	stan	8.1
BL7541	Df(2R)Exel6059	CG13229	3.8
BL9596	Df(2R)BSC161	mthl4	9.3
BL16906	$P\{EPgy2\}mthl3^{EY08706}$	mthl3	1.2
BL5574	$P\{w^{+mC}=lacW\}mthl3^{K10408}$	mthl3	1.9
BL10836	$P\{w^{+mC}\}l(2)k01209^{K08901a}$	mthl3	2.4
BL7550	Df(2R)Exel6068	5-HT1A, 5-HT1B	34.5
BL7998	Df(2R)Exel7166	CG3216	7.4
BL7912	Df(2R)Exel7184	CG13575	7.7
BL1473	$In(2LR)Px^{4}$	mAcR-60C	2.7
BL7914	Df(2R)Exel7185	CG13579	10.5
BL18327	$PBac\{w^{+mC}\}CG13579^{f00432}$	CG13579	24.1
BL8046	Df(3L)ED4079	mthl8	0
BL2577	Df(3L)emc-E12	mthl8,14, 9,10, mth	0
BL7563	Df(3L)Exel6048	mthl14, mthl9, mthl10	98.0
BL17137	$P\{w^{+mC}=EP\}EP3542$	mthl9	11.0
BL18405	$PBac\{w^{+mC}\}mthl10^{f01047}$	mthl10	17.1
BL8048	Df(3L)ED4177	mth	0
BL2400	Df(3L)R-G7	DmsR-2	9.1
BL8096	Df(3L)ED4287	DmsR-2, DmsR-1	2.7
BL7571	Df(3L)Exel6092	FR	0
BL7584	Df(3L)Exel6105	mthl2	26.0
BL17948	$PBac\{w^{+mC}=RB\}mthl2^{e01373}$	mthl2	30.0
BL6940	Df(3L)XAS96	CG10483	4.2

Stock #	Deficiency	GPCRs Affected	% malformation with *Sb*^{63b}
BL19366	PBac{}CG10483^{PL00048}	CG10483	3.5
BL1420	Df(3L)pbl-X1	mthl6	0
BL7929	Df(3L)Exel8104	mthl6	5.0
BL7591	Df(3L)Exel6112	mthl7	4.5
BL997	Df(3L)AC1	CG6749	7.5
BL7593	Df(3L)Exel6114	CG33696	3.7
BL19575	PBac{}CG33696^{PL00631}	CG33696	0
BL2611	Df(3L)vin5	Rh7	7.4
BL7595	Df(3L)Exel6116	Rh7	21.4
BL2612	Df(3L)vin^{7}	GRHRII	0
BL8070	Df(3L)ED4483	GRHRII	0
BL5408	l(3)SG11^{1}fz^{1}	fz	41.7
BL3124	Df(3L)fz-Gf3P	fz	6.7
BL1676	fz^{1}	fz	0
BL7601	Df(3R)Exel6174	fz	7.7
BL12512	P{w^{+mGT}=GT1}fz^{BG00824}	fz	0
BL15624	P{w^{+mC}y^{+}}fz^{EY03114}	fz	3.8
BL7937	Df(3L)Exel9004	Rh4	0
BL18260	PBac{w^{+mC}=RB}Rh4^{e04456}	Rh4	16.4
BL8100	Df(3L)ED4710	CG7497	0
BL7613	Df(3L)Exel6134	AlCR2	5.9
BL18622	PBac{w^{+mC}}AlCR2^{f0316}	AlCR2	15.6
BL6754	Df(3L)fz2	fz2	0
BL5126	Df(3L)XS533	NPFR76F	8.1
BL8088	Df(3L)ED4858	NPFR76F	0
BL8101	Df(3L)ED4978	capaR	0
BL9700	Df(3L)BSC223	TyrR	5.3
BL10268	P{hsneo}TyrR^{neo30}	TyrR	3.7
BL1559	Df(3R)110	5-HT2	3.5
BL19367	PBac{}5-HT2^{PL00052}	5-HT2	0
BL7628	Df(3R)Exel6149	CG8007	1.7
BL9215	Df(3R)ED5495	mthl11	0
BL8095	Df(3R)ED5558	mthl5	0
BL18486	PBac{w^{+mC}}CG8784^{f01901}	CG8784	19.4
BL7974	Df(3R)Exel8158	CG8784, CG8795, mthl12	5.5
BL18563	PBac{w^{+mC}}CG8795^{f02573}	CG8795	30.9
BL18464	PBac{w^{+mC}}CG9918^{f01726}	CG9918	17.1
BL1900	Df(3R)red-P52	CG9918, DopR	87.5
BL19491	PBac{}DopR^{PL00420}	DopR	15.3
BL7977	Df(3R)Exel7321	DopR	11.9
BL2057	Rh6^{1}	Rh6	0

Stock #	Deficiency	GPCRs Affected	% malformation with *Sb*^{63b}
BL7653	Df(3R)Exel6174	Rh6	1.6
BL7981	Df(3R)Exel8162	CG31183	5.0
BL9207	Df(3R)ED5785	Fsh	0
BL18659	PBac{WH}CG18208^{f03483}	CG18208	24.6
BL7660	Df(3R)Exel6181	CG18208, Rh2	23.6
BL19331	P{w^{+mC}=XP}cry^{d10630}	cry	0
BL3012	Df(3R)D1-BX12	cry, ninaE, Rh3	5.4
BL1861	ninaE^{8}	ninaE	2.7
BL2103	ninaE^{7}	ninaE	2.4
BL5701	ninaE^{17}	ninaE	7.7
BL9479	Df(3R)ED6027	Rh3	0
BL9485	Df(3R)ED10838	ETHR, SIFR	3.1
BL16202	PBac{w^{+mC}=5HPw^{+}}B322	SIFR	2.3
BL7681	Df(3R)Exel6202	CG11910	6.9
BL5601	Df(3R)Espl3	boss	4.5
BL9500	Df(3R)BSC140	boss	6.1
BL7683	Df(3R)Exel6204	CG12290	4.8
BL10400	PBac{w^{+mC}=PB}NepYr^{c01146}	NepYr	0
BL7684	Df(3R)Exel6205	NepYr	20.0
BL7688	Df(3R)Exel6210	AR-2	lethal - artifact
BL7692	Df(3R)Exel6214	AdoR	10.4
BL18848	PBac{w^{+mC} }5-HT7^{f05214}	5-HT7	18.0
BL5415	Df(3R)t11-e	5-HT7	3.3
BL18199	PBac{w^{+mC}}CG11318^{e03837}	CG11318	4.1
BL7919	Df(3R)Exel7379	CG11318, CG15556	2.4
BL18120	PBac{w^{+mC}}CG15556^{e03142}	CG15556	8.2
BL7084	Df(4)O2	Glu-RA	8.6
BL9579	Df(4)ED6380	Glu-RA	0

APPENDIX B: GENES UNCOVERED BY *DF(2L)EXEL7027*, *DF(2R)EXEL6068*, *DF(3L)EXEL6084* AND *DF(3R)RED-P52*

Df(2L)Exel7027	
Gene	**Molecular Function**
CG9232	UDP-glucose:hexose-1-phosphate uridylyltransferase activity; metal ion binding
Serpin-27A	serine-type endopeptidase inhibitor activity; enzyme inhibitor activity
CG34310	unknown
cortex	unknown
CG11329	unknown
Nhe3	sodium:hydrogen antiporter activity; solute:hydrogen antiporter activity
CG11327	unknown
Thrombospondin	calcium ion binding
Gonadotropin-releasing hormone receptor	neuropeptide receptor activity; G-protein coupled receptor activity; peptide receptor activity; gonadotropin-releasing hormone receptor activity; rhodopsin-like receptor activity
CG11188	unknown
homer	unknown
Liprin-α	unknown
CG11201	tubulin-tyrosine ligase activity
CG11323	tubulin-tyrosine ligase activity
CG11221	protein kinase activity; protein serine/threonine kinase activity; ATP binding
snoRNA:Me28S-G2743	unknown
CG31910	unknown
CG11322	unknown
CG11321	protein binding; zinc ion binding
CG17378	unknown
nervana 1	sodium:potassium-exchanging ATPase activity; hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances; cation transmembrane transporter activity

Df(2R)Exel6068	
Gene	**Molecular Function**
serrano	unknown
proliferation disrupter	chromatin binding
CG15107	unknown
Topoisomerase I-interacting protein	ubiquitin-protein ligase activity; DNA topoisomerase I binding; zinc ion binding
CG18605	unknown
CG15105	protein binding; zinc ion binding
Serotonin receptor 1B	5-HT1 receptor activity; serotonin receptor activity; G-protein coupled serotonin receptor activity; amine receptor activity
CG30126	unknown
CG15115	unknown
CG15116	glutathione peroxidase activity
CG15109	unknown
Serotonin receptor 1A	serotonin receptor activity; 5-HT1 receptor activity; G-protein coupled serotonin receptor activity; amine receptor activity
CG30125	unknown
CG15117	beta-glucuronidase activity; cation binding
brother of tout-velu	heparan sulfate N-acetylglucosaminyltransferase activity; acetylglucosaminyltransferase activity; glucuronyl-galactosyl-proteoglycan 4-alpha-N-acetylglucosaminyltransferase activity
CG15111	unknown

Df(3L)Exel6084	
Gene	**Molecular Function**
CG13875	calcium ion binding
CG34140	unknown
MAP kinase-specific phosphatase	MAP kinase tyrosine/serine/threonine phosphatase activity; phosphoprotein phosphatase activity; protein tyrosine phosphatase activity
Mitochondrial carrier homolog 1	transmembrane transporter activity; binding
rhinoceros	DNA binding; protein binding; zinc ion binding
mrityu	glycerol kinase activity; protein binding; voltage-gated potassium channel activity
Glycerol kinase	glycerol kinase activity
Nitrilase and fragile histidine triad fusion protein	bis(5'-adenosyl)-triphosphatase activity; hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds; nitrilase activity
CG13876	amine oxidase activity; quinone binding; copper ion binding

Gene	Molecular Function
CG7028	protein kinase activity; protein serine/threonine kinase activity; ATP binding
thoc7	unknown
CG16940	exoribonuclease II activity; RNA binding
Kaz1-ORFA	serine-type endopeptidase inhibitor activity
Kaz1-ORFB	serine-type endopeptidase inhibitor activity
pyrexia	calcium channel activity; cation channel activity
CG13877	unknown
CG16971	unknown
CG33229	unknown
CR32477	unknown
CG34454	serine-type endopeptidase inhibitor activity
CG34453	unknown
Enhancer of bithorax	ligand-dependent nuclear receptor binding; zinc ion binding
methuselah-like 14	G-protein coupled receptor activity
DISCO Interacting Protein 2	protein binding; catalytic activity; transcription factor binding
CG13879	unknown
Tudor-SN	transcription coactivator activity; nuclease activity; nucleic acid binding
snoRNA:Ψ18S-996	unknown
mitochondrial ribosomal protein L17	structural constituent of ribosome
CG34263	unknown
miple	growth factor activity
miple2	growth factor activity
CG32845	isomerase activity; protein binding
tiny tim 2	unknown
CG13881	unknown
CG13882	unknown
CG13883	unknown
RhoGEF3	Rho guanyl-nucleotide exchange factor activity; signal transducer activity
four wheel drive	1-phosphatidylinositol 4-kinase activity
CG34264	unknown
CG32344	helicase activity; ATP-dependent RNA helicase activity; ATP binding; RNA binding
CG32343	transcription factor activity
methuselah-like 9	G-protein coupled receptor activity
methuselah-like 10	G-protein coupled receptor activity

Df(3R)red-P52	
Gene	**Molecular Function**
Origin recognition complex subunit 2	DNA binding
roadkill	protein binding
Cyp6d5	electron carrier activity; heme binding; iron ion binding; monooxygenase activity
CG3061	unfolded protein binding; heat shock protein binding
CG9922	unknown
transfer RNA:ser2b:88A	triplet codon-amino acid adaptor activity
forkhead box, sub-group O	transcription factor activity; sequence-specific DNA binding
CG3153	unknown
CG14358	unknown
CG9920	unfolded protein binding; ATP binding
CG9918	G-protein coupled receptor activity; growth hormone-releasing hormone receptor activity; neuropeptide receptor activity
CG12402	protein binding
Odorant receptor 88a	olfactory receptor activity; odorant binding
CG14357	unknown
Kif19A	microtubule motor activity; ATP binding
Abelson Interacting Protein	signal transducer activity; protein binding
twinfilin	actin binding; protein-tyrosine kinase activity
upstream of RpII140	protein transporter activity
RNA polymerase II 140kD subunit	DNA-directed RNA polymerase activity; DNA binding
CG14356	unknown
CG14355	unknown
CG14359	unknown
CG9722	receptor binding
CG31533	unknown
CG31327	unknown
CG3199	unknown
Dopamine receptor	dopamine receptor activity; dopamine D1 receptor-like receptor activity; G-protein coupled receptor activity
CG9649	elastase activity; serine-type endopeptidase activity
CG31326	serine-type endopeptidase activity
CG33109	unknown
CG33329	serine-type endopeptidase activity
CG9631	elastase activity; serine-type endopeptidase activity
CG9624	unknown

Gene	Molecular Function
CG9616	unknown
Methionyl-tRNA synthetase	methionine-tRNA ligase activity; ATP binding
CG14840	unknown
CG14841	unknown
CG14839	unknown
trithorax	DNA binding; histone lysine N-methyltransferase activity (H3-K4 specific); sequence-specific DNA binding; protein binding; zinc ion binding; transcription factor activity
CG12207	unknown
CG3259	unknown
suppressor of Hairy wing	DNA binding; transcription regulator activity; protein binding; zinc ion binding
RNA polymerase II 15kD subunit	DNA-directed RNA polymerase activity; DNA binding; zinc ion binding; transcription regulator activity

REFERENCES

Adzhubei, A.A., I.A. Adzhubei, I.A. Krasheninnikov, and S. Neidle. 1996. Non-random usage of 'degenerate' codons is related to protein three-dimensional structure. *FEBS Letters*. 399:78-82.

Ahmed, Z.M., X.C. Li, S.D. Powell, S. Riazuddin, T.L. Young, K. Ramzan, Z. Ahmad, S. Luscombe, K. Dhillon, L. MacLaren, B. Ploplis, L.I. Shotland, E. Ives, S. Riazuddin, T.B. Friedman, R.J. Morell, and E.R. Wilcox. 2004. Characterization of a new full length *TMPRSS3* isoform and identification of mutant alleles responsible for nonsyndromic recessive deafness in Newfoundland and Pakistan. *BMC Med. Genet*. 5:24.

Akashi, H. 1995. Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in Drosophila DNA. *Genetics*. 139:1067-1076.

Amano M., M. Ito, K. Kimura, Y. Kukata, K. Chihara, T. Natano, Y. Matsuura, and K. Kaibuchi. 1996. Phosphorylation and activation of myosin by Rho-associated kinase (Rho-kinase). *J. Biol. Chem*. 271:20246-20249.

Appel, L.F. 1992. The cloning and characterization of *Stubble-stubbloid*, a morphogenetic locus in *Drosophila*. Ph.D. Thesis, University of California, Berkeley, CA.

Appel, L.F., M. Prout, R. Abu-Shumays, A. Hammonds, J.C. Garbe, D. Fristrom, and J. Fristrom. 1993. The Drosophila Stubble-stubbloid gene encodes an apparent transmembrane serine protease required for epithelial morphogenesis. *Proc. Natl. Acad. Sci*. 90:4937-4941.

Arbeitman, M.N., E.M. Furlong, F. Imam, E. Johnson, B.H. Null, B.S. Baker, M.A. Krasnow, M.P. Scott, R.W. Davis, and K.P. White. 2002. Gene expression during the life cycle of Drosophila melanogaster. *Science*. 297:2270-2275.

Badenhorst, P., H. Xiao, L. Cherbas, S.Y. Kwon, M. Voas, I. Rebay, P. Cherbas, and C. Wu. 2005. The *Drosophila* nucleosome remodeling factor NURF is required for Ecdysteroid signaling and metamorphosis. *Genes Dev*. 19:2540-2545.

Basel-Vanagaite, L., R. Attia, A. Ishida-Yamamoto, L. Rainshtein, D.B. Amitai, R. Lurie, M. Pasmanik-Chor, M. Indelman, A. Zvulunov, S. Saban, N. Magal, E. Sprecher, and M. Shohat. 2007. Autosomal recessive ichthyosis with hypotrichosis caused by a mutation in *ST14*, encoding type II transmembrane serine protease matriptase. *Am. J. Hum. Genet*. 80:467-477.

Bayer C.A., S.R. Halsell, J.W. Fristrom, D.P. Kiehart, and L. von Kalm. 2003. Genetic interactions between the *RhoA* and *Stubble-stubbloid* loci suggest a role for a type II transmembrane serine protease in intracellular signaling during Drosophila imaginal disc morphogenesis. *Genetics*. 165:1417-1432.

Benitah, S.A., P.F. Valerón, L. van Aelst, C.J. Marshall, and J.C. Lacal. 2004. Rho GTPases in human cancer: an unresolved link to upstream and downstream transcriptional regulation. *Biochimica et Biophysica Acta*. 1705:121-132.

Ben-Yosef, T., M. Wattenhofer, S. Riazuddin, Z.M. Ahmed, H.S. Scott, J. Kudoh, K. Shibuya, S.E. Antonarakis, B. Boone-Tamir, U. Radhakrishna, S. Naz, Z. Ahmed, S. Riazuddin, A. Pandya, W.E. Nance, E.R. Wilcox, T.B. Friedman, and R.J. Morell. 2001. Novel mutations of *TMPRSS3* in four DFNB8/B10 families segregating congenital autosomal recessive deafness. *J. Med. Genet*. 38:396-400.

Berg, J.M., J.L Tymoczko, and L. Stryer. 2007. *Biochemistry*. 6th edition. W.H. Freeman and Company. New York.

Bernardi, F., G. Castaman, M. Pinotti, P. Ferraresi, M.G. Di Iasio, B. Lunghi, F. Rodeghiero, and G. Marchetti. 1996. Mutation pattern in clinically asymptomatic coagulation factor VII deficiency. *Hum. Mutat*. 8:108-115.

Bilen, J, and N.M. Bonini. 2005. Drosophila as a model for human neurodegenerative disease. *Annu. Rev. Genet*. 39:153-171.

Bode, W., and R. Huber. 1992. Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem*. 204:433-451.

Booden, M.A., D.P. Siderovski, and C.J. Der. 2002. Leukemia-associated Rho guanine nucleotide exchange factor promotes Gαq-coupled activation of RhoA. *Mol. Cell. Biol*. 22:4053-4061.

Boudreault, A., D. Gauthier, and C. Lazure. 1998. Proprotein convertase PC1/3-related peptides are potent slow tight-binding inhibitors of murine PC1/3 and Hfurin. *J. Biol. Chem*. 273:31574-31580.

Brand, K., K.A. Dugi, J.D. Bruznell, D.N. Nevin, and S. Santamarina-Fojo. 1996. A novel A→G mutation in intron I of the hepatic lipase gene leads to alternative splicing resulting in enzyme deficiency. *J. Lipid Res*. 37:1213-1223.

Brody, T., and A. Cravchik. 2000. *Drosophila melanogaster* G protein-coupled receptors. *J. Cell Biol*. 150:F83-F88.

Bugge, T.H., K. List, and R. Szabo. 2007. Matriptase-dependent cell surface proteolysis in epithelial development and pathogenesis. *Front. Biosci*. 12:5060-5070.

Burrows, N.P., A.C. Nicholls, A.J. Richards, C. Luccarini, J.B. Harrison, J.R. Yates, and F.M. Pope. 1998. A point mutation in an intronic branch site results in aberrant splicing of COL5A1 and in Ehlers-Danlos syndrome type II in two British families. *Am. J. Hum. Genet*. 63:390-398.

Cadigan, K.M., and R. Nusse. 1997. Wnt signaling: a common theme in animal development. *Genes Dev*. 11:3286-3305.

Carmona, E., É. Dufour, C. Plouffe, S. Takebe, P. Mason, J.S. Mort, and R. Ménard. 1996. Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry*. 35:8149-8157.

Chokki, M., H. Eguchi, I. Hamamura, H. Mitsuhashi, and T. Kamimura. 2005. Human airway trypsin-like protease induces amphiregulin release through a mechanism involving protease-activated receptor-2-mediated ERK activation and TNF alpha-converting enzyme activity in airway epithelial cells. *FEBS J*. 272:6387-6399.

Cohen, S.M. 1993. Imaginal disc development. In *The Development of* Drosophila melanogaster (ed. M. Bate and A. Martinez-Arias), pp. 747-841. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Condic M.L., D. Fristrom, J.W. and Fristrom. 1990. Apical cell shape changes during Drosophila imaginal leg disc elongation: a novel morphogenetic mechanism. *Development*. 111:23-33.

Cottrell, G.S., S. Amadesi, E.F. Grady, and N.W. Bunnett. 2004. Trypsin IV, a novel agonist of protease-activated receptors 2 and 4. *J. Biol. Chem*. 279:13532-13539.

Désilets, A., F. Béliveau, G. Vandal, F.O. McDuff, P. Lavigne, and R. Leduc. 2008. Mutation G827R in matriptase causing autosomal recessive ichthyosis with hypotrichosis yields an inactive protease. *J. Biol. Chem*. 283:10535-10542.

Dhanasekaran, S.M., T.R. Barrette, D. Ghosh, R. Shah, S. Varambally, K. Kurachi, K.J. Pienta, M.A. Rubin, and A.M. Chinnaiyan. 2001. Delineation of prognostic biomarkers in prostate cancer. *Nature*. 412:822-826.

Dobshansky, T. 1930. The manifold effects of the genes Stubble and stubbloid in *Drosophila melanogaster*. *Z. Indukt. Abstamm. Vererbungsl.* 54:427-457.

Dries, D.L., R.G. Victor, J.E. Rame, R.S. Cooper, X. Wu, X. Zhu, D. Leonard, S.I. Ho, Q. Wu, W. Post, and M.H. Drazner. Corin gene minor allele defines by 2 missense mutations is common in blacks and associated with high blood pressure and hypertension. *Circulation*. 112:2403-2410.

Elbracht, M., J. Senderek, T. Eggermann, C. Thürmer, J. Park., M. Westhofen, and K. Zerres. 2007. Autosomal recessive postlingual hearing loss (DFNB8): compound heterozygosity for two novel *TMPRSS3* mutations in German siblings. *J. Med. Genet*. 44:e81.

Esser, V., L.E. Limbird, M.S. Brown, J.L. Goldstein, and D.W. Russell. 1988. Mutational analysis of the Ligand binding domain of the low density lipoprotein receptor. *J. Biol. Chem*. 104:765-772.

Fan, B., J. Brennan, D. Grant, F. Peale, L. Rangell, and D. Kirchhofer. 2007. Hepatocyte growth factor activator inhibitor-1 (HAI-1) is essential for the integrity of basement membranes in the developing placental labyrinth. *Dev. Biol*. 303:222-230.

Fristrom, D. and J.W. Fristrom. 1993. The metamorphic development of the adult epidermis. *In* The Development of *Drosophila melanogaster*. M. Bate and A. Martinez-Arias, editors. Cold Spring Harbor Laboratory Press, Plainview, NY.

Fujimaru, T., A. Tanaka, K. Choeh, N. Wakamatsu, H. Sakuraba, and G. Isshiki. 1998. Two mutations remote from an exon/intron junction in the beta-hexosaminidase beta-subunit gene affect 3'-splice site selection and cause Sandhoff disease. *Hum. Genet*. 103:462-469.

Fukuhara, S., H. Chikumi, and S. Gutkind. 2000. Leukemia-associated Rho guanine nucleotide exchange factor (LARG) links heterotrimeric G proteins of the G_{12} family to Rho. *FEBS Letters*. 485:183-188.

Fukuhara, S., C. Murga, M. Zohar, T. Igishi, and J.S. Gutkind. 1999. A novel PDZ domain containing guanine nucleotide exchange factor links heterotrimeric G proteins to Rho. *J. Biol. Chem*. 274:5868-5879.

Geneste, O., J.W. Copeland, and R. Treisman. 2002. LIM kinase and Diaphanous cooperate to regulate serum response factor and actin dynamics. *J. Cell. Biol*. 157:831-838.

Giannelli F., P.M. Green, S.S. Sommer, D.P. Lillicrap, M. Ludwig, R. Schwaab, P.H. Reitsma, M. Goossens, A. Yoshioka, and G.G. Brownlee. 1994. Haemophilia B: database of point mutations and short additions and deletions , fifth edition, 1994. *Nucleic Acids Res*. 22:3534-3546.

Gombos, L., J. Kardos, A. Patthy, P. Medveczky, L. Szilágyi, A. Málnási-Csizmadia, andL. Gráf. 2008. Probing conformational plasticity of the activation domain of trypsin: The role of glycine hinges. *Biochem*. 47:1675-1684.

Gorry, M.C., D. Gabbaizedeh, W. Furey, L.K. Gates, Jr., R.A. Preston, C.E. Aston, Y. Zhang, C. Ulrich, G.D. Erlich, and D.C. Whitcomb. 1997. Mutations in the cationic trypsinogen gene are associated with recurrent acute and chronic pancreatitis. *Gastroenterology*. 113:1063-1068.

Gotwals, P.J., and J.W. Fristrom. 1991. Three neighboring genes interact with the Broad-Complex and the Stubble-stubbloid locus to affect imaginal disc morphogenesis in Drosophila. *Genetics*. 127:747-759.

Graveley, B.R. 2004. A protein interaction domain contacts RNA in the prespliceosome. *Mol. Cell*. 13:302-304.

Greenwald, I. 1998. LIN-12/Notch signaling: lessons from worms and flies. *Genes Dev*. 12:1751-1762.

Guipponi, M., S.E. Antonarakis, and H.S. Scott. 2008. *TMPRSS3*, a type II transmembrane serine protease mutated in non-syndromic autosomal recessive deafness. *Front. Biosci*. 13:1557-1567.

Guipponi, M., J. Tan, P.Z.F. Cannon, L. Donley, P. Crewther, M. Clarke, Q. Wu, R.K. Shepherd, and H.S. Scott. 2007. Mice deficient for the type II transmembrane serine protease, *TMPRSS1/hepsin*, exhibit profound hearing loss. *Am. J. Pathol*. 171:608-616.

Guipponi, M., G. Vuagniaux, M. Wattenhofer, K. Shibuya, M. Vazquez, L. Dougherty, N. Scamuffa, E. Guida, M. Okui, C. Rossier, M. Hancock, K. Buchet, A. Reymond, E. Hummler, P.L. Marzella, J. Kudoh, N. Shimizu, H.S. Scott, S.E. Antonarkis, and B.C. Rossier. 2002. The transmembrane serine protease (TMPRSS3) mutated in deafness DFNB8/10 activates the epithelial sodium channel (ENaC) in vitro. *Hum. Mol. Genet*. 11:2829-2836.

Halsell, S.R., B.I. Chu, and D.P. Kiehart. 2000. Genetic analysis demonstrates a direct link between rho signaling and nonmuscle myosin function during Drosophila morphogenesis. *Genetics*. 155:1253-1265.

Hamlington, J.D., M.V. Clough, J.A. Dunston, and I. McIntosh. 2000. Deletion of a branch-point consensus sequence in the *LMX1B* gene causes exon skipping in a family with nail patella syndrome. *Eur. J. Hum. Genet.* 8:311-314.

Hariharan, I.K., and D. Bilder. 2006. Regulation of imaginal disc growth by tumor-suppressor genes in Drosophila. *Annu. Rev. Genet*. 40:335-361.

Hariharan, I.K., K.Q. Hu, H. Asha, A. Quintanilla, R.M. Ezzell, and J. Settleman. 1995. Characterization of rho GTPase family homologues in Drosophila melanogaster: overexpressing Rho1 in retinal cells causes a late developmental defect. *EMBO J*. 14:292-302.

Hart, M.J., X. Jiang, T. Kozasa, W. Roscoe, W.D. Singer, A.G. Gilman, P.C. Sternweis, and G. Bollag. 1998. Direct stimulation of the guanine nucleotide exchange activity of p115RhoGEF by $G\alpha_{13}$. *Science*. 280:2112-2114.

Hauser, F., L. Søndergaard, and C.J.P. Grimmelikhuijzen. 1998. Molecular cloning, genomic organization and developmental regulation of a novel receptor from *Drosophila melanogaster* structurally related to gonadotropin-releasing hormone receptors from vertebrates. *Biochem. Biophys. Res. Commun*. 249:822-828.

Hedstrom, L. 2002. Serine protease mechanism and specificity. *Chem. Rev.* 102:4501-4523.

Heitzler, P. and P. Simpson. 1991. The choice of cell fate in the epidermis of Drosophila. *Cell*. 64:1083-1092.

Hobson, J.P., S. Netzel-Arnett, R. Szabo, S.M. Réhault, F.C. Church, D.K. Strickland, D.A. Lawrence, T.M. Antalis, and T.H. Bugge. 2004. Mouse *DESC1* is located within a cluster of seven *DESC1*-like genes and encodes a type II transmembrane serine protease that forms serpin inhibitory complexes. *J. Biol. Chem*. 279:46981-46994.

Hooper, J.D., J.A. Clements, J.P. Quigley, and T.M. Antalis. 2001. Type II transmembrane serine proteases. *J. Biol. Chem*. 276:857-860.

Hooper, J.D., A.L. Scarman, B.E. Clarke, J.F. Normyle, and T.M. Antalis. 2000. Localization of the mosaic transmembrane serine protease corin to heart myocytes. *Eur. J. Biochem*. 267:6931-6937.

Hutchin, T., N.N. Coy, H. Conlon, E. Telford, K. Bromelow, D. Blaydon, G. Taylor, E. Coghill S. Brown, R. Trembath, X.Z. Liu, M. Bitner-Glindzicz, and R. Mueller. 2005. Assessment of the genetic causes of recessive childhood non-syndromic deafness in the UK – implications for genetic testing. *Clin. Genet*. 68:506-512.

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

Iwanaga, S., S. Kawabata, and T. Muta. 1998. New types of clotting factors and defense molecules in horseshoe crab hemolymph: their structures and functions. *J. Biochem*. 123:1-15.

Jackson, G.R., I. Salecker, X. Dong, X. Yao, N. Arnheim, P.W. Faber, M.E. MacDonald, and S.L. Zipursky. 1998. Polyglutamine-expanded human Huntington transgenes induce degeneration of Drosophila photoreceptor neurons. *Neuron*. 21:633-642.

Jacquinet, E., N.V. Rao, G.V. Rao, Z.M. wang, K.H. Albertine, and J.R. Hoidal. 2001. Cloning, genomic organization, chromosomal assignment and expression of a noval mosaic serine proteinase: Epitheliasin. *FEBS Letters*. 468:93-100.

Johnson, R.L., and M.P. Scott. 1998. New players and puzzles in the Hedgehog signaling pathway. *Curr*. *Opin*. *Genet*. *Dev*. 8:450-456.

Kilpatrick, L.M., R.L. Harris, K.A. Owen, R. Bass, C. Ghorayeb, A. Bar-Or, and V. Ellis. 2006. Initiation of plasminogen activation on the surface of monocytes expressing the type II transmembrane serine protease matriptase. *Blood*. 108:2616-2623.

Kimchi-Sarfaty, C., J.M. Oh, I.W. Kim, Z.E. Sauna, A.M. Calcagno, S.V. Ambudkar, and M.M. Gottesman. 2007. A "silent" polymorphism in the *MDR*1 gene changes substrate specificity. *Science*. 315:525-528.

Kimura, K., M. Ito, M. Amano, K. Chihara, Y. Fukata, M. Nakafuku, B. Yamamori, J. Feng, T. Nakano, K. Okawa, A. Iwamatsu, and K. Kaibuchi. 1996. Regulation of myosin phosphatase by Rho and Rho-associated kinase (Rho-Kinase). *Science*. 273:245-248.

Klezovitch, O., J. Chevillet, J. Mirosevich, R.L. Roberts, R.J. Matusik, and V. Vasioukhin. 2004. Hepsin promotes prostate cancer progression and metastasis. *Cancer Cell*. 6:185-195.

Kozasa, T., X. Jiang, M.J. Hart. P.M. Sternweis, W.D. Singer, A.G. Gilman, G. Bollag, and P.C. Sternweis. 1998. p115RhoGEF, a GTPase activating protein for $G\alpha_{12}$ and $G\alpha_{13}$. *Science*. 280:2109-2111.

Kuivenhoven, J.A., H. Weibusch, P.H. Pritchard, H. Funke, R. Benne, G. Assmann, and J.J. Kastelein. 1996. An intronic mutation in a lariat branchpoint sequence is a direct cause of an inherited human disorder (fish-eye disease). *J*. *Clin*. *Invest*. 98:358-364.

Lin, B.Y., C. Ferguson, J.T. White, S.Y. Wang, R. Vessella, L.D. True, L. Hood, and P.S. Nelson. 1999. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer Res*. 59:4180-4184.

Lin, C.Y. J.K. Wang, J. Torri, L. Dou, Q.X.A. Sang, and R.B. Dickson. 1997. Characterization of a novel membrane-bound, 80-kDa matrix-degrading protease from human breast cancer cells – Monoclonal antibody production, isolation, and localization. *J. Biol. Chem*. 272:9147-9152.

List, K., C.C. Haudenschild, R. Szabo, W. Chen, S.M. Wahl, W. Swaim, L.H. Engelholm, N. Behrendt, and T.H. Bugge. 2002. Matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis. *Oncogene*. 21:3765-3779.

List, K., R. Szabo, A. Molinolo, B.S. Nielsen, and T.H. Bugge. 2006. Delineation of matriptase protein expression by enzymatic gene trapping suggests diverging roles in barrier function, hair formation, and squamous cell carcinogenesis. *Am. J. Pathol*. 168:1513-1525.

List, K., R. Szabo, A. Molinolo, V. Sriuranpong, V. Redeye, T. Murdock, B. Burke., B.S. Nielsen, J.S. Gutkind, and T.H. Bugge. 2005. Deregulated matriptase causes ras-independent multistage carcinogenesis and promotes ras-mediated malignant transformation. *Genes Dev*. 19:1934-1950.

Luo, J., D.J. Duggan, Y. Chen, J. Sauvageot, C.M. Ewing, M.L. Bittner, J.M. Trent, and W.B. Isaacs. 2001. Human prostate cancer and benign prostatic hyperplasia: Molecular dissection by gene expression profiling. *Cancer Res*. 61:4683-4688.

Lutz, S., A. Freichel-Blomquist, Y. Yang, U. Rümenapp, K.H. Jakobs, M. Schmidt, and T. Wieland. 2005. The guanine nucleotide exchange factor p63RhoGEF, a specific link between $G_{q/11}$-coupled receptor signaling and RhoA. *J. Biol. Chem*. 280:11134-11139.

Lutz, S., A. Shankaranarayanan, C. Coco, M. Ridilla, M.R. Nance, C. Vettel, D. Baltus, C.R. Evelyn, R.R. Neubig, T. Wieland, and J.J.G. Tesmer. 2007. Structure of $G\alpha_q$-p63RhoGEF-RhoA complex reveals a pathway for the activation of RhoA by GPCRs. *Science*. 318:1923-1927.

MacFarlane S.R., M.J. Seatter, T. Kanke, G.D. Hunter, and R. Plevin. 2001. Proteinase-activated Receptors. *Pharmacol. Rev*. 53:245-282.

Maekawa M., T. Ishizaki, S. Boku, N. Watanabe, A. Fujita, A. Iwamatsu, T. Obinata, K. Ohashi, K. Mizuno, and S. Narumiya. 1999. Signaling from Rho to the actin cytoskeleton through protein kinases ROCK and LIM-kinase. *Science*. 285:895-898.

Magee, J.A., T. Araki, S. Patil, T. Ehrig, L. True, P.A. Humphrey, W.J. Catalona, M.A. Watson, and J. Milbrandt. 2001. Expression profiling reveals hepsin overexpression in prostate cancer. *Cancer Res*. 61:5692-5696.

Magie, C.R., M.R. Meyer, M.S. Gorsuch, and S.M. Parkhurst. 1999. Mutations in the Rho1 small GTPase disrupt morphogenesis and segmentation during early Drosophila development. *Development*. 126:5353-5364.

Masmoudi, S., S.E. Antonarakis, T. Schwede, A.M. Ghorbel, M. Gratri, M.P. Papasavvas, M. Drira, A. Elgaied-Boulila, M. Wattenhofer, C. Rossier, H.S. Scott, H. Ayadi, and M. Guiponni. 2001. Novel missense mutations of TMPRSS3 in two consanguineous Tunisian families with non-syndromic autosomal recessive deafness. *Hum. Mutat*. 18:101-108.

Matsui T., M. Maeda, Y. Doi, S. Yonemura, M. Amano, K. Kaibuchi, S. Tsukita, and S. Tsukita. 1998. Rho-kinase phosphorylates COOH-terminal threonines of ezrin/radixin/moesin (ERM) proteins and regulates their head-to-tail association. *J. Cell Biol*. 140:647-657.

Matthews, D.A., R.A. Alden, J.J. Birktoft, S.T. Freer, and J. Kraut. 1975. X-ray crystallographic study of boronic acid adducts with subtilisin BPN' (Novo). *J. Biol. Chem*. 250:7120-7126.

Mount, S.M., C. Burks, G. Hertz, G.D. Stormo, O. White, and C. Fields. 1992. Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. *Nucleic Acids Res*. 20:4255-4262.

Muller, H.J. 1932. Further studies on the nature and causes of gene mutation. *Proc. 6th Int. Congr. Genet*. 1:213-255.

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

Muta, T., T. Oda, and S. Iwanaga. 1993. Horseshoe crab coagulation factor B: a unique serine proteinase zymogen activated by cleavage of an Ile-Ile bond. *J. Biol. Chem*. 268:21384-3-21388.

Netzel-Arnett, S., B.M. Currie, R. Szabo, C.Y. Lin, L.M. Chen, K.X. Chai, T.M. Antalis, T.H. Bugge, and K. List. 2006. Evidence for a matriptase-prostasin proteolytic cascade regulating terminal epidermal differentiation. *J. Biol. Chem*. 281:32941-32945.

Oberst, M.D., C.A. Williams, R.B. Dickson, M.D. Johnson, and C.Y. Lin. 2003. The activation of matriptase requires its noncatalytic domains, serine protease domain, and its cognate inhibitor. *J. Biol. Chem*. 278:26773-26779.

Perona, J.J, and C.S. Craik. 1997. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem*. 272:29987-29990.

Pruyne, D., M. Evangelista, C. Yang, E. Bi, S. Zigmond, A. Bretscher, and C. Boone. 2002. Role of formins in actin assembly: Nucleation and barbed-end association. *Science*. 297:612-615.

Putman, E.A., E.S. Park, C.M. Aalfs, R.C.M. Hennekam, and D.M. Milewicz. 1997. Parental somatic an germ-line mosaicism for a FBN2 mutation and analysis of FBN2 transcript levels in dermal fibroblasts. *Am. J. Hum. Genet*. 60:818-827.

Ruggiero, R.P., 2006. Notopleural mutations enhance defects in imaginal disc epithelial morphogenesis and marcochete elongation associated with mutations in the *Stubble-stubbloid* locus. Master's Thesis, University of Central Florida, Orlando, FL.

Sagot, I., A.A. Rodal, J. Moseley, B.L. Goode, and D. Pellman. 2002. An actin nucleation mechanism mediated by Bni1 and Profilin. *Nature Cell Biol*. 4:626-631.

Saudou, F., U. Boschert, N. Amlaiky, J.L. Plassat, and R. Hen. 1992. A family of Drosophila serotonin receptors with distinct intracellular signaling properties and expression patterns. *EMBO J*. 11:7-17.

Scott, H.S., J. Kudoh, M. Wattenhofer, K. Shibuya, A. Berry, R. Chrast, M. Guipponi, J. Wang, K. Kawasaki, S. Asakawa, S. Minoshima, F. Younus, S.Q. Mehdi, U. Radhakrishna, M.P. Papasavvas, C. Gehrig, C. Rossier, M. Korostishevsky, A. Gal, N. Shimizu, B. Boone-Tamir, and S.E. Antonarakis. 2001. Insertion of beta-satellite repeats identifies a transmembrane protease causing both congenital and childhood onset autosomal recessive deafness. *Nat. Genet*. 27:59-63.

Shi, Y.E., J. Torri, L. Yieh, A. Wellstein, M.E. Lippman, and R.B. Dickson. 1993. Identification and characterization of a novel matrix-degrading protease from hormone dependent human breast-cancer cells. *Cancer Res*. 53:1409-1415.

Singer S.J. 1990. The structure and insertion of integral membrane proteins. *Ann. Rev. Cell Biol*. 6:247-296.

Smith, C., H. Giordano, and R. DeLotto. 1994. Mutational analysis of the Drosophila *snake* protease: An essential role for domains within the proenzyme polypeptide chain. *Genetics*. 136:1355-1365.

Soller, M.J., P. Elfving, R. Lundgren, and I. Panagopols. 2006. Confirmation of the high frequency of the TMPRSS2/ERG fusion gene in prostate cancer. *Genes Chromosomes and Cancer*. 45:717-719.

Speck, O., S.C. Hughes, N.K. Noren, R.M. Kulikauskas, and R.G. Fehon. 2003. Moesin functions antagonistically to the Rho pathway to maintain epithelial integrity. *Nature*. 421:83-87.

Srivastava, D.P., E.J. Yu, K. Kennedy, H. Chatwin, V. Reale, M. Hamon, T Smith, and P.D. Evans. 2005. Rapid, nongenomic responses to ecdysteroids and catecholamines mediated by a novel *Drosophila* G-protein-coupled receptor. *J. Neurosci*. 25:6145-6155.

Stamey, T.A., J.A. Warrington, M.C. Caldwell, Z. Chen, Z. Fan, M. Mahadevappa, J.E. McNeal, R. Nolley, and Z. Zhang. 2001. Molecular genetic profiling of gleason grade 4/5 prostate cancers compared to benign prostatic hyperplasia. *J. Urol*. 166:2171-2177.

Strutt, D.I., U. Weber, and M. Mlodzik. 1997. The role of RhoA in tissue polarity and Frizzled signalling. *Nature*. 387:292-295.

Szabo, R, A. Molinolo, K. List, and T.H. Bugge. 2007. Matriptase inhibition by hepatocyte growth factor activator inhibitor-1 is essential for placental development. *Oncogene*. 26:1546-1556.

Szabo, R., S. Netzel-Arnett, J.P. Hobson, T.M. Antalis, and T.H. Bugge. 2005. Matriptase-3 is a novel phylogenetically preserved membrane-anchored serine protease with broad serpin reactivity. *Biochem. J*. 390:231-242.

Szabo, R., Q. Wu, R.B. Dickson, S. Netzel-Arnett, T.M. Antalis, and T.H. Bugge. 2003. Type II transmembrane serine proteases. *Thromb. Haemost*. 90:185-193.

Takeuchi T., J.L. Harris, W. Huang, K.W. Yan, S.R. Coughin, and C.S. Craik. 2000. Cellular localization of membrane-type serine Protease 1 and identification of Protease-activated Recptor-2 and single-chain urokinase-type plasminogen activator as substrates. *J. Biol. Chem*. 275:26333-26342.

Takeuchi, T.,M.A., Shuman, and C.S. Craik. 1999. Reverse biochemistry: Use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrane-type serine protease in epithelial cancer and normal tissue. *Proc. Natl. Acad. Sci. U.S.A*. 96:11054-11061.

Thummel, C.S., and V. Pirrotta. 1991. Technical Notes: New pCaSpeR P-element vectors. *D.I.N.* 2.

Tomlins, S.A., D.R. Rhodes, S. Perner, S.M. Dhanasekaran, R. Mehra, X.W. Sun, S. Varambally, X. Cao, J. Tchinda, R. Kuefer, C. Lee, J.E. Montie, R.B. Shah, K.J. Pienta, M.A. Rubin, and A.M. Chinnaiyan. 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*. 310:644-648.

Tsukita S., S. Yonemura, and S. Tsukita. 1997. ERM proteins: head-to-tail regulation of actin-plasma membrane interaction. *Trends Biochem. Sci*. 22:53-58.

Uprichard, J., and D.J. Perry. 2002. Factor X deficiency. *Blood Rev*. 16:97-110.

Vaarala, M.H., K. Porvari, A. Kyllonen, O. Lukkarinen, and P. Vihko. 2001. The TMPRSS2 gene encoding transmembrane serine protease is overexpressed in a majority of prostate cancer patients.: Detection of mutated TMPRSS2 form in a case of aggressive disease. *Int. J. Cancer*. 94:705-710.

Van Aelst L., M. and Symons. 2002. Role of Rho family GTPases in epithelial morphogenesis. *Genes Dev*. 16:1032-1054.

Velasco, G., S. Cal, V. Quesada, L.M. Sánchez, and C. Lopez-Otín. 2002. Matriptase-2, a membrane-bound mosaic serine proteinase predominantly expressed in human liver and showing degrading activity against extracellular matrix proteins. *J. Biol. Chem*. 277:37637-37646.

Vogt, S., R. Grosse, G. Schultz, and S. Offermanns. 2003. Receptor-dependent RhoA activation in G_{12}/G_{13}-deficient cells. *J. Biol. Chem*. 278:28743-28749.

von Kalm, L., D.J. Fristrom, and J.W. Fristrom. 1995. The making of a fly leg: a model for epithelial morphogenesis. *BioEssays*. 17:693-702.

Vu, T.K.H., D.T. Hung, V.I. Wheaton, and S.R. Coughlin. 1991. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell*. 64:1057-1068.

Ward R.E., J. Evans, and C.S. Thummel. 2003. Genetic modifier screens in Drosophila demonstrate a role for Rho1 signaling in ecdysone-triggered imaginal disc morphogenesis. *Genetics*. 165:1397-1415.

Warrick, J.M., H.L. Paulson, G.L. Gary-Board, Q.T. Bui, K.H. Fischbeck, R.N. Pittman, and N.M. Bonini. 1998. Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in Drosophila. *Cell*. 93:939-949.

Wassarman, P.M., M. Therrien, and G.M. Rubin. 1995. The Ras signaling pathway in Drosophila. *Curr. Opin. Genet. Dev*. 5:44-50.

Watanabe N., T. Kato, A. Fujita, T. Ishizaki, and S. Narumiya. 1999. Cooperation between mDia1 and ROCK in Rho-induced actin reorganization. *Nature Cell Biol*. 1:136-143.

Wattenhofer, M., M.V. Di Iorio, R. Rabionet, L. Dougherty, A. Pampanos, T. Schwede, B. Montserrat-Sentis, M.L. Arbones, T. Iliades, A. Pasquadibisceglie, M. D'Amelio, S. Alwan, C. Rossier, H.H.M. Dahl, M.B. Petersen, X. Estivill, P. Gasparini, H.S. Scott, and S.E. Antonarakis. 2002. Mutations in the TMPRSS3 gene are a rare cause of childhood nonsydromic deafness in Caucasian patients. *J. Mol. Med*. 80:124-131.

Wattenhoffer, M., N. Sahin-Calapoglu, D. Andreasen, E. Kalay, R. Caylan, B. Braillard, N. Fowler-Jaeger, A. Reymond, B.C. Rossier, A. Karaguzel, S.E. Antonarakis. 2005. A novel TMPRSS# missense mutation in a DFNB8/10 family prevents proteolytic activation of the protein. *Hum.Genet*. 117:528-535.

Welsh, J.B., L.M Sapinoso, A.I. Su, S.G. Kern, J. Wang-Rodriguez, C.A. Moskaluk, H.F. Frierson, Jr., and G.M. Hampton. 2001. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res*. 61:5974-5978.

Wilkie, T.M., D.J. Gilbert, A.S. Olsen, X.N. Chen, T.T. Amatruda, J.R. Korenberg. B.J. Trask, P. de Jong, R.R. Reed, M.I. Simon, N.A. Jenkins, and N.G. Copeland. 1992. Evolution of the mammalian G protein alpha subunit multigene family. *Nature Genet*. 1:85-91.

Wilson S., B. Greer, J. Hooper, A. Zijlstra, B. Walker, J. Quigley, and S. Hawthorne. 2005. The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. *Biochem. J*. 388:967-972.

Wulff, K., and F.H. Hermann. 2000. Twenty two novel mutations of the factor VII deficiency. *Hum. Mutat*. 15:489-496.

Yan, W., N. Sheng, M. Seto, J. Morser, and Q. Wu. 1999. Corin, a mosaic transmembrane serine protease encoded by a novel cDNA from human heart. *J. Biol. Chem*. 274:14926-14935.

Yuan, Q., W.J. Joiner, and A. Sehgal. 2006. A sleep-promoting role for the Drosophila serotonin receptor 1A. *Curr. Biol*. 16:1051-1062.

Yuan, Q., F. Lin, X. Zheng, and A. Sehgal. 2005. Serotonin modulates circadian entrainment in Drosophila. *Neuron*. 47:115-127.

Zheng Y., D. Zangrilli, R.A. Cerione, and A. Eva. 1996. The pleckstrin homologydomain mediates transformation by oncogenic Dbl through specific intracellular targeting. *J. Biol. Chem*. 271:19017-19020.

Zou, Z., D.L. Lopez, M.R. Kanost, J.D. Evans, and H.B. Jiang. 2006. Comparative analysis of serine protease-related genes in the honey bee genome: Possible involvement in embryonic development and innate immunity. *Insect Mol. Biol*. 15:603-614.