The EMBO Journal Vol.18 No.4 pp.959–967, 1999

Anopheles gambiae Ag-STAT, a new insect member of the STAT family, is activated in response to bacterial infection

Carolina Barillas-Mury[1,2], Yeon-Soo Han, Douglas Seeley and Fotis C.Kafatos

European Molecular Biology Laboratory, Meyerhofstrasse 1, Heidelberg, D-69117, Germany

[1]Present address: Department of Pathology, Colorado State University, Fort Collins, CO 80523-1671, USA

[2]Corresponding author

A new insect member of the STAT family of transcription factors (Ag-STAT) has been cloned from the human malaria vector *Anopheles gambiae*. The domain involved in DNA interaction and the SH2 domain are well conserved. Ag-STAT is most similar to *Drosophila* D-STAT and to vertebrate STATs 5 and 6, constituting a proposed ancient class A of the STAT family. The mRNA is expressed at all developmental stages, and the protein is present in hemocytes, pericardial cells, midgut, skeletal muscle and fat body cells. There is no evidence of transcriptional activation following bacterial challenge. However, bacterial challenge results in nuclear translocation of Ag-STAT protein in fat body cells and induction of DNA-binding activity that recognizes a STAT target site. *In vitro* treatment with pervanadate (vanadate and H_2O_2) translocates Ag-STAT to the nucleus in midgut epithelial cells. This is the first evidence of direct participation of the STAT pathway in immune responses in insects.

Keywords: immunity/infection/mosquito/STAT

Introduction

When the immune system is challenged by a pathogen, the organism needs to coordinate the activity of various organs participating in the response. In insects, the fat body is the main site of immune peptide biosynthesis (reviewed by Hoffmann and Reichhart, 1997), but other organs may be the first ones to come into contact with the invading organism: examples are the midgut in the case of malarial infection of mosquitoes, or the epidermis in the case of cuticle wounding in the presence of fungi or bacteria. Our understanding of cell–cell communication during the immune response in insects is very limited. Several regulatory pathways involved in the regulation of anti-microbial peptides have been identified in *Drosophila*, such as the *imd* pathway which is only defined genetically (Lemaitre *et al.*, 1995), and the pathways that involve related receptors such as Toll (Lemaitre *et al.*, 1996) and 18-Wheeler (Williams *et al.*, 1997), as well as downstream transcription factors of the Rel family, such as Dorsal, Dif and Relish.

The Toll pathway acts through the transcription factor Dorsal in dorso-ventral pattern formation in embryos, and is also involved in the regulation of the gene for the anti-fungal peptide Drosomycin. The *spätzle* gene activates the Toll pathway during embryonic development (Morisato and Anderson, 1994) and also functions in immunity: loss-of-function *spätzle* mutants fail to activate Drosomycin gene transcription in response to immune challenge (Lemaitre *et al.*, 1996). Furthermore, the Spätzle protein shows a cysteine knot motif also found in extracellular ligands such as nerve growth factor (NGF), transforming growth factor β2 (TGF-β2) and platelet-derived growth factor BB (PDGF-BB) (McDonald and Hendrickson, 1993). These two lines of evidence suggest that Spätzle is the ligand that activates the Toll receptor in fat body cells, and thus it probably represents the first example of an insect cytokine involved in the immune response. However, direct biochemical evidence of Spätzle-mediated activation of Toll is not available. Interleukin-1 (IL-1)-like activity has been detected in the hemolymph of the moth *Manduca sexta*, but the molecular nature of this activity remains to be defined (Beck *et al.*, 1996).

In vertebrates, cytokines such as interleukins and interferons play a central role in regulating and coordinating the immune response (reviewed by Leaman *et al.*, 1996). These cytokines interact with specific membrane receptors and result in activation of members of the JAK (Janus kinase) family, which in turn activate members of the STAT (signal transducers and activators of transcription) family of transcription factors (reviewed by Ihle, 1996; Darnell, 1997; O'Shea, 1997). The seven known members of the STAT family share several redundant functions, but gene disruption experiments in mice have shown that three of them have non-redundant functions regulating distinct aspects of the immune response: STAT1 participates in the innate immune response to viral and bacterial infections (Meraz *et al.*, 1996), while STAT4 (Kaplan *et al.*, 1996) and STAT6 (Takeda *et al.*, 1996) are involved in the regulation of T_H1 and T_H2 acquired immune responses, respectively.

In *Drosophila*, two components of the STAT pathway have been identified: a member of the JAK family encoded by the *hopscotch* (*hop*) gene (Perrimon and Mahowald, 1986; Binari and Perrimon, 1994) and the transcription factor encoded by *d-STAT* or *STAT 92E* (Hou *et al.*, 1996; Yan *et al.*, 1996). Epistatic analysis has revealed that *d-STAT* acts downstream of *hop*. These two genes participate in the regulation of pair-rule genes during embryonic development, through a maternal contribution (Binari and Perrimon, 1994). A second zygotic function involves regulation of cell proliferation. Loss-of-function mutations of *hop* result in reduced size of larval diploid structures and larval–pupal lethality, while a dominant gain-of-function mutation in *hop*, $hop^{Tum\text{-}1}$, is associated with hypertrophy of the larval lymph gland (the hemato-

poietic organ), a leukemia-like phenotype and the formation of melanotic tumours (Hanratty and Dearolf, 1993; Harrison *et al*., 1995). The Toll pathway has also been implicated in controlling hemocyte density (Qiu *et al*., 1998), leading to the suggestion that, in the lymph gland, the Rel family and JAK/STAT pathways are linked, in a manner analogous to the NF-κB-activated production of cytokines that activate the JAK/STAT pathway in mammals (Mathey-Prevot and Perrimon, 1998). However, no involvement of the *Drosophila* JAK/STAT pathway in immunity has been reported to date.

The mosquito *Anopheles gambiae* is the main vector involved in human malaria transmission in Africa. We are interested in the immune pathways of this mosquito and ultimately in their possible involvement in interactions with the malaria parasite. We have previously characterized Gambif1 (Gambae immune factor 1), a member of the Rel family of transcription factors, and shown that it is translocated to the nucleus in response to bacterial challenge (Barillas-Mury *et al*., 1996). We report here the cloning and functional characterization of Ag-STAT, a new member of the STAT family. Two independent functional assays indicate that this transcription factor is activated in mosquitoes in response to bacterial infection. To our knowledge, this is the first evidence that a member of the STAT family participates in immune response regulation in insects. Our results support the growing body of evidence that the pathways regulating gene expression in response to infection have remarkable evolutionary conservation, from insects to humans.

Results

Isolation and characterization of Ag-STAT

The Ag-STAT gene was isolated using PCR with degenerate primers based on two protein regions well conserved between mammalian and *Drosophila* STATs (black arrows in Figure 1A). *Anopheles gambiae* genomic DNA was used as template. An 840 bp product was obtained, corresponding to an open reading frame (ORF) with high similarity to other STAT family members (see below). This fragment was used as a probe to isolate an Ag-STAT clone from a genomic library. Two 1.5 and 2.9 kb *Eco*RI fragments were subcloned and a 2918 bp sequence was obtained. It encompasses a 2166 bp coding region corresponding to a 722 amino acid continuous ORF (Figure 1A), plus 450 bp upstream and 302 bp downstream of this ORF (EMBL Data Library, accession No. AJ010299). Five potential methionine codons found upstream of the ORF are followed by stop codons after eight, five, nine, 22 and 48 amino acids, suggesting the possibility of translational control (Abastado *et al*., 1991). None of these short potential ORFs showed homology to sequences in the database. The overall organization and the deduced Ag-STAT amino acid sequence is shown in Figure 1A.

Ag-STAT belongs to an ancient class of the STAT family of transcription factors

Figure 1B displays the sequence alignment of Ag-STAT with six different human STATs and *Drosophila*, *Dictyostelium* and *Caenorhabditis elegans* STATs (the latter identified by searching the sequence database). The blue arrows in Figure 1A indicate the central conserved region which is included in the alignment; the invertebrate members of this family are strongly divergent at both the N- and C-termini of the proteins.

The degree of sequence similarity and conserved features leave no doubt that the newly isolated sequence is a member of the STAT family. The conserved features include the DNA interaction domain, the phosphotyrosine-binding SH2 domain and a tyrosine at the position where other family members are known to be phosphorylated by JAKs. A motif at the core of the SH2 domain, encompassing an arginine that is predicted to make direct contact with the membrane receptor phosphotyrosine residue, is conserved in all sequences. The conserved residues cluster around the predicted DNA interaction and SH2 domains.

The probable evolutionary relationships between the analysed STAT sequences are illustrated in the matrix of Figure 1C and the rooted dendrogram of Figure 1D. Ag-STAT, the *Drosophila* STAT and human STATs 5 and 6 are most closely related to each other (37.6–51.2% identity in the aligned central region); we refer to these sequences as the Class A homology group. Similarly, the human STATs 1–4 sequences form a separate homology group, class B (45.2–63.4% identity). The percentage identities between classes A and B are distinctly lower (25.8–32.2%) than within each class. The *C.elegans* and *Dictyostelium* STATs are even more distantly related to the insect and vertebrate STATs as well as to each other (15.2–21.4% identity). The phylogenetic analysis indicates that, most probably, an ancestral STAT gene had already duplicated prior to the split between insects and vertebrates, creating the founder genes of classes A and B (see Discussion).

Ag-STAT is expressed in all developmental stages and in several different adult tissues

The pattern of expression of Ag-STAT mRNA was analysed semi-quantitatively using radioactive RT–PCR at the embryonic, larval, pupal and adult stages (Figure 2). Ag-STAT was found to be expressed throughout development. After normalization against the ribosomal S7 product, the transcript level was lowest in the pupa and also low in the embryo; it was higher in larvae and adult females, and highest in adult males. The transcript level remained stable or decreased (+13% to –35% change) after bacterial infection, and decreased by 27% following blood feeding in adult females.

The tissue distribution of Ag-STAT protein was determined by whole-tissue mount immunofluorescence in adult females. Rabbit polyclonal antibodies were raised against recombinant Ag-STAT protein corresponding to the region initially amplified using degenerate oligonucleotides (indicated by the black arrows in Figure 1A). Hemocytes show strong cytoplasmic and nuclear staining; in anopheline mosquitoes, these cells can be observed conveniently, attached to body wall preparations on the surface of structures such as the alary muscles (Figure 3A) or tracheae (Figure 3B). The pericardial cells (Figure 3C) and the fat body cells (see below) also show cytoplasmic and nuclear staining. In skeletal flight muscle, Ag-STAT is present in the nuclei, which are evenly spaced along the fibre axis (Figure 3D), and in strong and weak transverse bands paralleling the sarcomeric pattern (Figure 3E). The cytoplasmic space surrounding spherical organ-

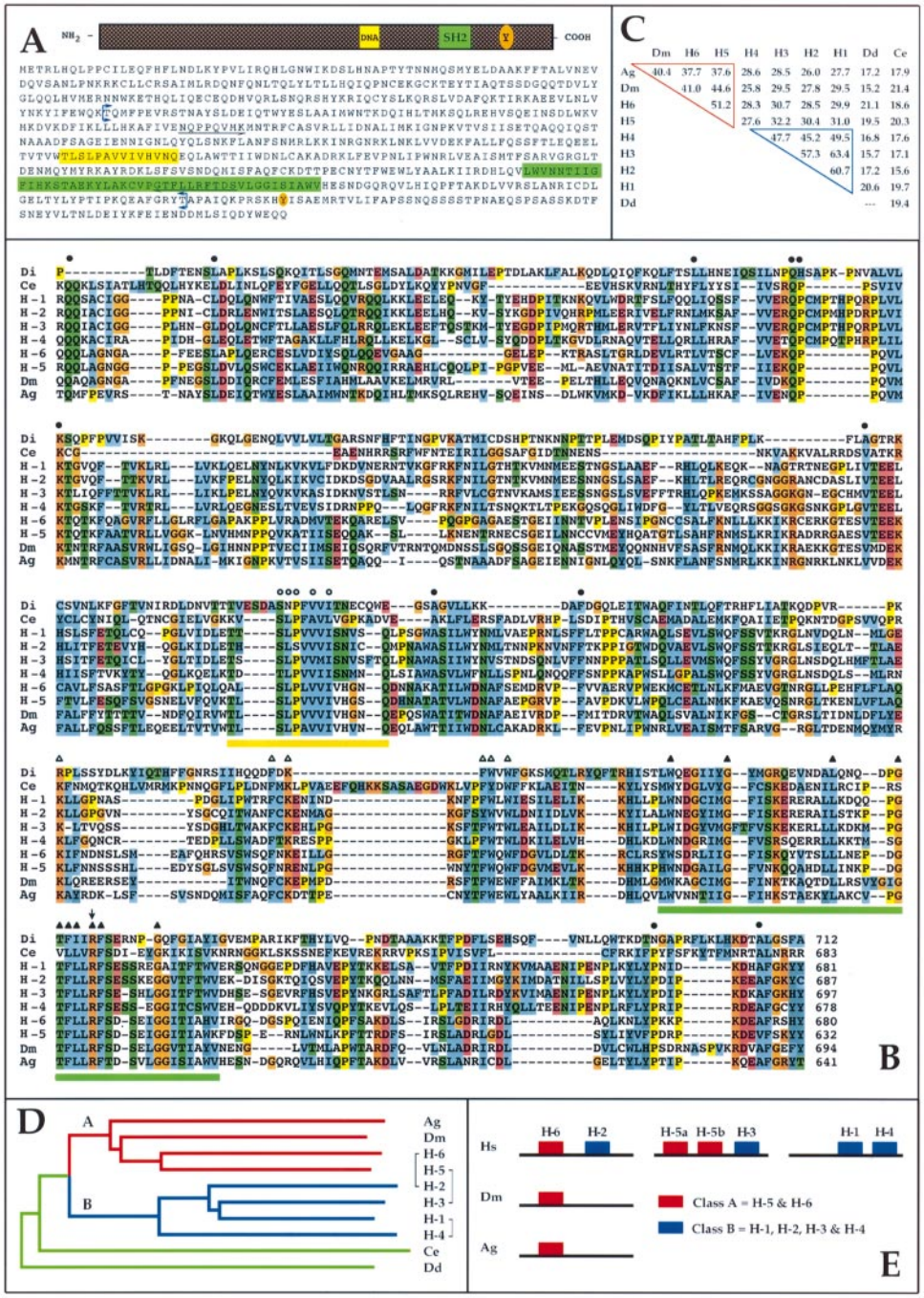

Fig. 1. (**A**) Schematic representation and deduced amino acid sequence of Ag-STAT (EMBL Data Library, accession No. AJ010299). Black arrows indicate the two conserved regions from which the degenerate oligonucleotides were designed, and the blue brackets mark the central conserved region included in the alignment. The DNA-binding domain is indicated in yellow and the SH2 domain in green; the orange oval indicates the putative phosphorylated tyrosine. (**B**) Aligned sequences of the conserved domains of Ag-STAT (Ag) and nine other STAT family members (from EMBL Data Library; accession Nos indicated). *Dictyostelium* (Dd) = Y13097, *C.elegans* (Ce) = Z70754, human STAT1 (H-1) = P42224, H-2 = P526301, H-3 = P40763, H-4 = Q14765, H-5 = P42229, H-6 = P42226, *Drosophila* (Dm) = Q24151. The DNA-binding and SH2 domains are underlined with coloured rectangles as in (A). The vertical arrow indicates the phosphotyrosine-binding arginine of the SH2 domain which is conserved in all sequences. Amino acids conserved in at least nine out of these 10 sequences are indicated by black dots, those corresponding to the region known to be directly involved in DNA interaction in vertebrates by open dots, those matching the SH2 domain consensus as black triangles, and those matching the SH3 consensus as open triangles. The recently solved structures of STAT3 (Becker *et al*., 1998) and STAT1 (Chen *et al*., 1998) revealed that matches with the SH3 domain consensus do not reflect the presence of such a domain at the structural level. Instead, structural similarity to calcium-binding domains, as found in troponin C, has been described (Becker *et al*., 1998). (**C**) Matrix of percentage amino acid identities in the aligned central region between STAT family members. Red and blue triangles outline the members of the A and B classes, respectively. (**D**) Dendrogram based on the alignment of the STAT sequences in (B), constructed using CLUSTAL X (Thompson *et al.*, 1994). (**E**) Schematic representation of the chromosomal locations of vertebrate and insect STATs. Genes in the class A homology group are indicated in red and those in class B in blue.

elles (most likely the large mitochondria characteristic of insect skeletal muscle) is also positive (Figure 3F). These tissues were negative when stained with pre-immune serum (data not shown). The anti-STAT antibodies are highly specific according to Western blot analysis. They and an anti-His-Tag monoclonal antibody recognize the same protein band in extracts of *Escherichia coli* bearing the Ag-STAT sequence in a His-Tag expression vector, but only if expression is induced by isopropyl-β-D-thiogalactopyranoside (IPTG) (data not shown). Moreover, the anti-STAT antibodies recognize a single band of ~78 kDa in homogenates of adult female mosquitoes, which is not observed with the pre-immune serum at the same dilution (Figure 3G).

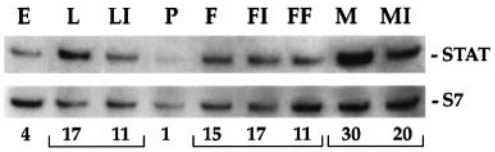

Fig. 2. RT–PCR analysis of Ag-STAT transcript levels at different developmental stages and physiological conditions. In all samples, RNA sequences encoding the ribosomal protein S7 (Salazar *et al*., 1993) were amplified simultaneously as internal control for the reaction. E = embryos, L = larvae, LI = larvae post-infection, P = pupae, F = adult females, FI = females post-infection, FF = blood-fed females, M = adult males and MI = males post-infection. The post-infection samples were obtained 4 h after pricking the mosquitoes with a tungsten needle dipped in a mixture of *E.coli* and *M.luteus*. The intensity of each PCR product was quantitated using a phosphoimager and normalized relative to the corresponding S7 band. Relative abundances of STAT transcript in samples are indicated by numbers at the bottom.

Ag-STAT is activated in response to bacterial challenge

Electrophoretic mobility shift assays (EMSAs) were used to monitor DNA-binding activity that recognizes the optimum D-STAT target site (Yan *et al*., 1996) in response to infection. Larval and adult *A.gambiae* were challenged by pricking with a tungsten needle that had been dipped in a bacterial pellet containing *E.coli* and *Micrococcus luteus* (Gram-negative and Gram-positive bacteria, respectively). The body walls (including the cuticle, epidermis and other associated cells as well as fat body tissue) were dissected 90 min post-challenge, frozen in liquid nitrogen and used to prepare nuclear extracts. DNA-binding activity was observed in nuclear extracts from both larval and adult female preparations, but only after bacterial challenge (Figure 4). A single band was observed in larval and a doublet in adult female extracts. This suggested the existence of more than one type of complex in adults, potentially a heterodimer or a homodimer of a still uncharacterized STAT family member. The binding was competed completely by pre-incubating the extract with a 100-fold excess of unlabelled oligonucleotide. Pre-incubation with anti-Ag-STAT antibodies consistently reduced binding by ~40–60% and, in the case of the adult female, resulted in a weak supershifted band.

Ag-STAT is distributed throughout the cytoplasm and

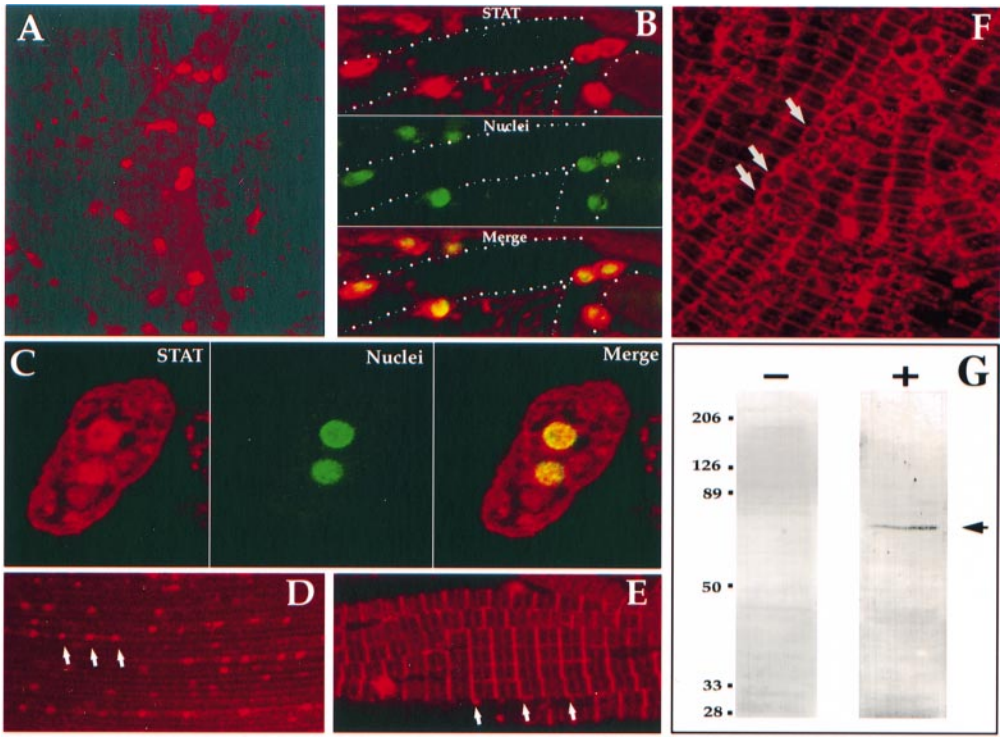

Fig. 3. Immunofluorescence analysis of Ag-STAT expression in adult females (red). Note that (B), (C), (E) and (F) are at higher magnification than (A) and (D). (**A**) Hemocytes attached to the alary muscles in a body wall preparation. (**B**) Hemocytes on the surface of tracheal branches doubly stained to indicate the nuclei in green. White dotted lines outline the tracheae. The merged image confirms that Ag-STAT is present in the nuclei (yellow) as well as in the cytoplasm (red). (**C**) STAT presence in both the nuclei and cytoplasm of binucleated pericardial cells (nuclei in green). (**D**–**F**) Ag-STAT staining in skeletal muscle. (D) The arrows indicate the nuclei located in rows along the fibre axis. (E) The arrows indicate strong STAT bands that alternate with weaker bands, in sarcomere-like periodicity. (F) The cytoplasm of the muscle surrounding round organelles, probably representing mitochondria (arrows), is also stained. (**G**) Western blots of adult female homogenates using anti-Ag-STAT rabbit antiserum (+) or pre-immune serum (–). The molecular weight markers are indicated in kDa.

nucleus of fat body cells of unchallenged females (Figure 5A). The subcellular distribution of immunostainable Ag-STAT changes dramatically between 30 and 60 min following bacterial challenge, as the cytoplasm is depleted and the nucleus enriched in Ag-STAT (Figure 5A). The detailed features of this phenomenon (see Discussion) are in agreement with the known mechanism of activation of STAT proteins by membrane recruitment and phosphorylation, followed by nuclear translocation (reviewed by Ihle, 1996; Darnell, 1997; O'Shea, 1997).

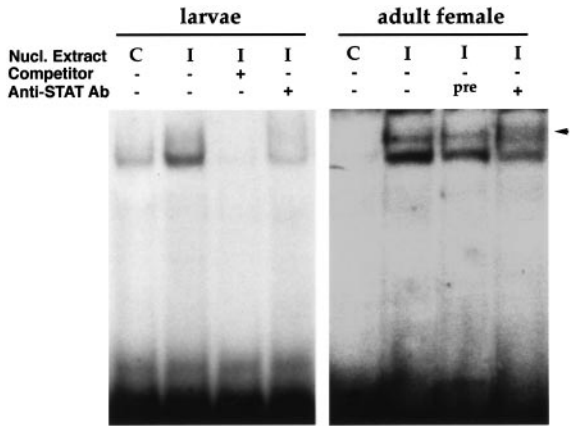

Fig. 4. Electrophoretic mobility shift assay (EMSA) of *A.gambiae* nuclear extracts from larvae and adult females before (C = control) or 90 min post-bacterial challenge (I = infected). The radioactive oligonucleotide encompassed the optimal binding site of *Drosophila* D-STAT (Yan *et al.*, 1996). Binding is completely abolished by cold competitor oligonucleotide in a 100-fold molar excess. Anti-Ag-STAT polyclonal antibodies partially decrease the intensity of the binding and, in the adult female, create a supershifted band (arrowhead).

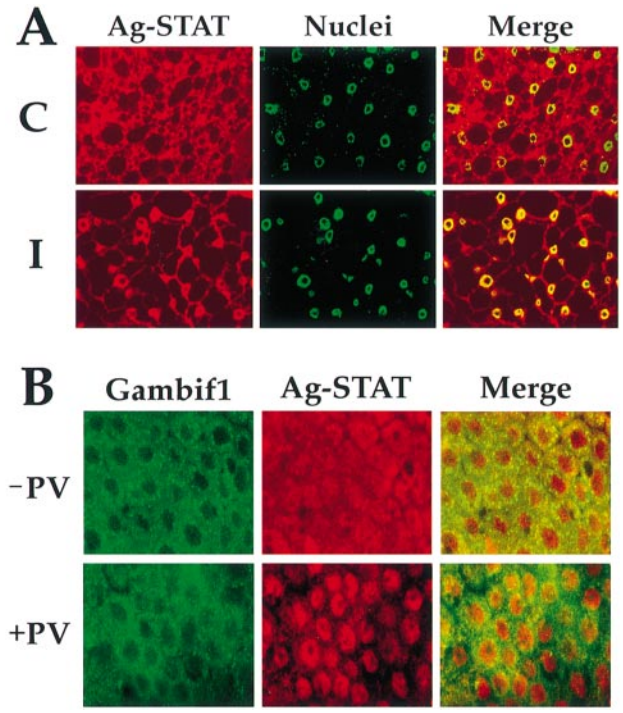

Fig. 5. (**A**) Immunofluorescence staining of fat body with anti-Ag-STAT antibodies (red). In unchallenged females (C), Ag-STAT is present in the nuclei (green) and cytoplasm of fat body cells. Round unstained inclusions are storage granules and lipid-filled vacuoles. Following bacterial challenge (I), the cytoplasmic staining decreases while the nuclear staining intensifies, indicating nuclear translocation in response to infection. (**B**) In untreated control midgut cells (–PV), Gambif1 is located in the cytoplasm (green) and STAT in the cytoplasm and nucleus (red). Following pervanadate treatment (+PV), the distribution of Gambif1 remains unchanged, while Ag-STAT is depleted from the cytoplasm and accumulates in the nucleus.

Ag-STAT can be activated by pervanadate treatment in vitro

Gambif1is a previously identified Rel transcription factor (Barillas-Mury *et al.*, 1996) which is also activated by bacterial challenge, as monitored by DNA binding and by nuclear translocation in fat body cells. Both Ag-STAT and Gambif1 are expressed in the midgut (Figure 5B). Gambif1 protein is largely restricted to the cytoplasmic compartment of midgut epithelial cells, while Ag-STAT is evenly distributed in the cytoplasm and nucleus. Pervanadate treatment (hydrogen peroxide and vanadate) is known to activate the STAT pathway in vertebrates and *Drosophila* through a ligand-independent mechanism (Zhong *et al.*, 1994; Yan *et al.*, 1996). *In vitro* treatment of dissected midguts with pervanadate has no noticeable effect on Gambif1 subcellular distribution, but results in prominent Ag-STAT translocation from the cytoplasm to the nucleus (Figure 5B).

Discussion

Evolutionary origins and classes A and B of the STAT family

The appearance of multicellular organisms created the need for cellular differentiation and for efficient cell–cell communication to coordinate suitable responses to environmental changes. The STAT family of transcription factors is involved in both differentiation and signalling. It is absent from the yeast genome but present in lower eukaryotes as far back as *Dictyostelium*, a facultative metazoan in which it mediates the selective expression of an extracellular matrix protein in pre-stalk cells, in response to a differentiation-inducing factor (Kawata *et al.*, 1997). The D-STAT of *Drosophila* has been implicated in embryonic pattern formation, hemocyte differentiation and the regulation of proliferation in imaginal discs. Mammalian STAT genes play an essential role in embryonic development and are also involved in innate and acquired immunity (reviewed by Hou *et al.*, 1996; Yan *et al.*, 1996; Darnell, 1997; O'Shea, 1997). All mammalian STAT proteins are widely expressed and are activated by several cytokines *in vitro*. However, gene disruption experiments reveal a smaller subset of evidently non-redundant functions. Thus, STATs 2 and 3 are essential for embryonic development, STATs 5A and 5B are required for breast development and lactation, STAT1 is required for the innate response of macrophages to viruses and bacteria (mediated by interferon-α and interferon-γ), while STAT4 and STAT6 are required for distinct aspects of acquired immunity (mediated by IL-12 and IL-4, respectively).

As summarized in Figure 1E, the seven known mammalian STAT genes are clustered at three different chromosomal locations. Interestingly, two of these clusters encompass members of both the A and B classes, and are linked to ErbB-related genes. These features (Copeland *et al.*, 1995) led to the postulate that a gene duplication

formed an ancestral STAT gene pair (A + B in our terminology), and that additional (possibly whole-genome) duplications then led to the formation of the multiple gene clusters; a very recent local reduplication apparently led to formation of the tandem STAT5a and 5b genes, which are 96% identical at the amino acid level. One inconsistency of the model of Copeland *et al.* (1995) is that one of the clusters only contains B class genes (STATs 1 and 4; 49.5% identity); a possible explanation is that soon after the appearance of an A + B pair at this chromosomal site, the B class gene reduplicated while the A class gene was deleted.

Whether the *Dictyostelium* or *C.elegans* STAT is used as the out group to root the dendogram of Figure 1D, the topology is maintained: the two insect STAT sequences cluster with the class A mammalian sequences (STATs 5 and 6). Alignment of the conserved central region of STAT sequences (Figure 1B) shows that the insect and class A mammalian STATs share five deletions/insertions, and have a proportion of shared residues much higher than with class B mammalian STATs. Interestingly, the insect sequences are as divergent from each other (40.4% identity) as from the mammalian class A sequences (37.6–44.6% identity). It should be noted that mosquitoes and *Drosophila* are quite distant in evolution (>200 million years since their last common ancester; Hennig, 1981). In the STAT family, the N- and C-terminal sequences are too divergent for reliable phylogenetic conclusions.

The simplest interpretation of the sequence comparisons in Figure 1B–D is that the A–B gene duplication occurred prior to the last common ancestor between insects and mammals. If this interpretation is correct, it begs the question of whether B family STAT genes were deleted in the insect lineage, or remain to be discovered. The predominant pattern of A–B pairing in mammals raises the possibility that B family STAT genes may reside in the vicinity of the presently known class A genes of *Drosophila* and *Anopheles.*

Ag-STAT expression and tissue distribution in adults

Ag-STAT mRNA has a similar developmental expression pattern to that of D-STAT (Hou *et al.*, 1996), being expressed in embryos, larvae and adults, with lower levels of expression in late pupal stages. The spatial expression patterns of D-STAT mRNA in the embryo are consistent with the demonstrated functional importance of STAT-binding sites in the promoters of pattern formation genes (Yan *et al*., 1996). The tissue distribution of D-STAT mRNA or protein in post-embryonic stages is unknown, but the leukemia-like overproliferation of hemocytes upon constitutive activation of the STAT pathway suggests that regulated expression of STAT is important for these cells (Hanratty and Dearolf, 1993; Harrison *et al*., 1995). Whether D-STAT also participates in immune responses presently is unknown.

We have observed Ag-STAT staining in mosquito hemocytes, which have multiple functions including bacterial phagocytosis and melanization, as well as pericardial cells which function in detoxification. In both of these cell types, the protein is constitutively present in the nucleus and cytoplasm, and no subcellular translocation is associated with infection (data not shown). Ag-STAT is also expressed in skeletal muscle. Interestingly, the class A STAT5a and STAT5b mRNAs are also expressed in mammalian skeletal muscle (Liu *et al*., 1995); no subcellular localization studies or muscle phenotypes in STAT5a or STAT5b mutants have been reported. We have also observed that Ag-STAT is present in all fat body cells, both in the cytoplasm and in the nucleus. This tissue is involved in storage of food resources and innate immunity, serving as the insect analogue of the vertebrate liver. Disruption of STAT5b in mice results in loss of sexual dimorphism of body growth rates and liver gene expression, a dimorphism which is known to be mediated by pulsatile growth hormone secretion (Udy *et al*., 1997). *In vivo* injection of growth hormone in rats results in phosphorylation of JAK2 and STAT5 (but not other STAT proteins) in liver and skeletal muscle (Chow *et al.*, 1996). In conclusion, the tissue expression patterns of Ag-STAT suggest regulatory and functional conservation in this family across a wide evolutionary distance.

Ag-STAT activation in response to bacterial challenge and pervanadate

In view of the central role of STAT genes in the cytokine-driven regulation of innate and acquired immunity in vertebrates, we have investigated whether Ag-STAT also participates in the anti-bacterial innate immune response. At the RNA level, Ag-STAT is not induced by bacterial challenge; in fact, the levels decrease slightly in larvae and adult males 4 h post-infection. However, protein level activation has been documented by two independent functional assays: immunolocalization and inducible-DNA binding activity to the optimal D-STAT DNA target sequence.

DNA-binding activity is induced rapidly and strongly in larvae and adults after bacterial challenge (Figure 4). A partial but consistent reduction of DNA binding by addition of Ag-STAT antibodies, and the observation of a supershifted band in adult extracts, indicate that Ag-STAT is part of the protein–DNA complex, as well as suggesting the existence of more than one type of inducible STAT component. Stable DNA binding was only achieved when all buffers contained vanadate (a phosphatase inhibitor), indicating that the DNA-binding activity is phosphorylation dependent, as is expected for STAT family members.

Although Ag-STAT is constitutively localized in both cytoplasm and nuclei in multiple tissues, in fat body cells it can be translocated rapidly to the nuclei within 45–60 min following bacterial challenge. The abdominal wall preparations allowed us to evaluate the entire abdominal fat body of each challenged insect. Interestingly, in a large number of experiments, translocation was observed in only half of the infected animals, but the positive animals showed translocation in all of their fat body cells. In a typical experiment, eight females were challenged; five or six of them showed systemic bacterial infection, and two or three of the latter showed clear nuclear translocation. Translocation was not observed in any control unchallenged individuals. The all-or-none behaviour suggests that the response is systemic, while its variability among animals may reflect the transient nature of STAT activation. A total activation–inactivation cycle time of ~20 min for any given STAT molecule has been estimated (reviewed by Darnell, 1997), based on

kinetic studies of interferon-γ activation of STAT1 (Haspel *et al*., 1996). *In vitro* treatment with pervanadate allowed us to activate the STAT pathway directly and synchronously in the midgut epithelium, a known immune organ (Dimopoulos *et al.*, 1997). Following a 20 min incubation, nuclear translocation of Ag-STAT was observed in all individuals treated (8–10 individuals per experiment) and in none of the untreated controls.

We have also explored preliminarily the possibility that Ag-STAT might be involved in the immune response of the mosquito to the malaria parasite (data not shown). EMSA experiments consistently reveal induction of a STAT-like DNA-binding activity after malaria infection. Immunolocalization experiments show nuclear translocation of STAT in malaria-infected midguts, but only sporadically; this irreproducibility may reflect the transient nature of STAT activation compounded by the known variability of the malaria infection process.

Concluding remarks

We present the first evidence of activation of the STAT pathway in response to bacterial infection in insects. The upstream components as well as the effector genes activated by this pathway remain to be determined. The existence of JAK kinase (Perrimon and Mahowald, 1986; Binari and Perrimon, 1994) and D-STAT in *Drosophila* (Hou *et al*., 1996; Yan *et al*., 1996) and the activation of Ag-STAT following bacterial challenge and pervanadate treatment indicate that the JAK/STAT signalling pathway and its regulation have been conserved in evolution. STAT genes from early metazoans need to be analysed and functionally characterized to understand more fully the structural and functional evolution of this family of transcription factors. It is noteworthy that Ag-STAT in the mosquito and STAT1 in mammals both participate in the innate immune response to bacteria, although they belong to different classes of the STAT family (A and B, respectively); moreover, mammalian members of both the A and B classes (STAT6 and STAT4, respectively) control distinct aspects of acquired immunity. Detailed phylogenetic studies of the STAT family may contribute new insights into how the vertebrate acquired immune system evolved from the more widespread and ancient innate immune system.

Materials and methods

Mosquitoes

Anopheles gambiae strain G3 was reared at 28°C, at 75% humidity on a 12 h light–dark cycle. Larvae were grown in distilled water with 0.01% table salt, and fed on cat food. Adults were fed *ad libidum* on 10% sucrose and females were blood-fed on anaesthetized rats. Larvae were challenged with *E.coli* 1160 and *M.luteus* A270 as described (Richman *et al*., 1996).

PCR-based cloning of the Ag-STAT gene

Two well conserved regions were chosen to synthesize degenerate oligonucleotides (these regions are indicated by arrows in Figure 1A). A 5′ extension which introduces an *Xho*I site (bold) was also included in both the forward primer 5′-GCCG**CTCGAG**AARCARCCICCICARGTIHTIAARAC-3′ and the reverse primer 5′-GCCG**CTCGAG**RTCIGARAAYCTIARIARRAAIGTNCC-3′. The PCR reactions were performed using 20 pmol of each primer, 50 ng of genomic DNA template in 50 µl reactions and AmpliTaq (Perkin Elmer) with standard buffer conditions (1.5 mM $MgCl_2$). Two cycles (1 min steps at 95, 55 and 72°C, and 95, 42 and 72°C) were followed by 30 cycles at moderate stringency (1 min steps at 95, 52 and 72°C) and a final 7 min extension at 72°C. A major band of ~900 bp was obtained and cloned using the TA Cloning Kit (Invitrogen). This product was sequenced and used to generate a probe for screening an *A.gambiae* genomic library. After secondary screening, the insert of one clone was analysed by Southern blotting and subcloned into the Bluescript vector (Stratagene). The Ag-STAT-coding region was contained within 1.5 and 2.9 kb *Eco*RI fragments. A 2918 bp region, encompassing a 722 amino acid continuous ORF (2166 bp), as well as 453 bp upstream and 299 bp downstream from the ORF, was sequenced by dideoxy double-stranded sequencing (Sequenase, US Biochemical Corp.).

Sequence comparisons

Sequence comparisons with the DDBJ/EMBL/GenBank and SWISSPROT databases were performed using FastA with the GCG software (Devereux *et al*., 1984). Sequence alignments and dendrograms were constructed using CLUSTAL X (Thompson *et al*., 1994). The Ag-STAT sequence has been submitted to the DDBJ/EMBL/GenBank Data Library, with accession No. AJ010299.

Ag-STAT expression in E.coli and antibody preparation

Part of the conserved Ag-STAT domain was recovered by PCR (nucleotides 1369–2157) and cloned in the His-Tag expression vector Pet-15b (Novagen). The proteins were expressed in *E.coli* and purified using an Ni^{2+}-NTA–agarose metal chelation chromatography column (Qiagen), under denaturing conditions (6 M guanidine–HCl), with step washings (pH 8.0 to 6.3 to 5.9) and elution at pH 4.0 following the manufacturer's protocol.

Initial immunizations were performed by injecting two rabbits with RIBI adjuvant containing denatured protein (300 ng in a band cut from an SDS–polyacrylamide gel). Boosters consisted of 50 ng of the same protein mixture in RIBI adjuvant and were injected three times, beginning 4 weeks after injection and at 3 week intervals thereafter. Immune response was followed by testing several serum dilutions in Western blots. The immunostaining patterns reported were specific, in that they were not observed with pre-immune sera.

Western blot analysis

Adult female homogenates were prepared, submitted to SDS–PAGE gel electrophoresis on a 10% gel and transferred to an Immobilon-P membrane (Millipore). The membrane was treated with 1 mM levamisol (Sigma) in Tris-buffered saline (TBS), pH 7.5 for 20 min to inhibit endogenous phosphatase activity. The specificity of the rabbit anti-Ag-STAT antibody was tested using Western blot analysis with serum before and after immunization with recombinant Ag-STAT at a 1/10 000 dilution and detected with goat anti-rabbit alkaline phosphatase-conjugated secondary antibody at a 1/10 000 dilution. All antibody incubations were done in a 6% bovine serum albumin (BSA) solution in TBS pH 7.

RT–PCR

Total RNA was isolated from groups of 50 mosquitoes using RNaid glass beads, and poly$(A)^+$ RNA was selected using Oligotex-dT beads (Qiagen). First strand cDNA was synthesized using random hexamers and Superscript II (Gibco-BRL) for 1 h at 37°C. For expression studies, PCRs were performed using 20 pmol of each primer in 50 µl reactions and AmpliTaq (Perkin Elmer) with standard buffer conditions (1.5 mM $MgCl_2$) and [γ-^{32}P]ATP dCTP. The primer set 5′-TTCGAGGTACCGAATCTAATACCC-3′ and 5′-GATAATGGTATTGTCAACTAG-3′ was used to amplify a 290 bp region from Ag-STAT (1 min steps at 94, 55 and 72°C) for 10 cycles. A pause of 4 min at 72°C was then used to add a second pair of primers, complementary to the *A.gambiae* ribosomal protein gene S7 (Salazar *et al.*, 1993), 5′-GGCGATCATCATCTACGTGC-3′ and 5′-GTAGCTGCTGCAAACTTCGG-3′, to provide an internal control. The reaction was then allowed to proceed for 18 additional cycles with both sets of primers present. Staggered addition of primers was necessary due to the extremely low abundance of Ag-STAT mRNA relative to RNA for the S7 ribosomal protein. The PCR products were analysed by electrophoresis in a 6% non-denaturing polyacrylamide gel, followed by autoradiography and phosphoimager quantification. The cDNA templates used for this RT–PCR experiment were tested to ensure that there was no genomic DNA contamination. With the same cDNA templates, primers to a region of Gambif1 that has a small intron amplified a single band corresponding in size to the spliced mRNA, indicating the absence of genomic DNA templates (data not shown).

Fat body and midgut immunostainings

The abdominal wall of adult females was dissected with the fat body attached. The tissue was fixed for 2 h in 4% paraformaldehyde,

100 mM PIPES buffer pH 7.4, 2 mM $MgSO_4$, 1 mM EGTA, and was then treated with 80% ethanol for 3 min and blocked in 1% BSA, 0.1% SDS in phosphate-buffered saline (PBS) pH 7.5 (PBT) for 2 h at room temperature. In between each of these steps, the tissue was washed twice for 5 min in PBS pH 7.5. The rabbit antibody was used at 1:200 dilution in overnight incubations in PBT at 4°C and the nuclei were visualized with an anti-histone mouse monoclonal antibody (1:200 dilution). Secondary Cy3-conjugated goat anti-rabbit was used at 1:500 dilution and fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse at a 1:300 dilution in PBT for 3–4 h at room temperature. After each antibody incubation, the tissue was washed twice for 20 min with PBT. The nuclei were counterstained with a 4′,6-diamidino-2-phenylindole (DAPI) solution (1:15 000 dilution of stock solution, Boehringer Mannheim) for 2 min. Immunostaining was analysed by fluorescence microscopy, and nuclear DAPI staining by UV light illumination. The final images were obtained using confocal microscopy.

Midguts were dissected in PBS (130 mM NaCl, 7 mM Na_2HPO_4, 3 mM $NaH_2PO_4{\cdot}H_2O$ pH 7.2) and incubated for 20 min at room temperature, either in PBS alone (–PV) or in BPS containing 1 mM sodium orthovanadate and 2 mM H_2O_2 (+PV). The midguts were fixed and stained following the same protocol as described above for fat body, except that the tissue was fixed for 1 h and was not treated with ethanol.

Nuclear extracts and electrophoretic mobility shift assays (EMSA)

Nuclear extracts were prepared using a minipreparation protocol (Deryckere and Gannon, 1994) and the protein concentration was determined by the Bradford method (Bio-Rad). All adult females were blood-fed 48 h previously, because extracts from unfed females gave extremely low protein concentrations and no detectable binding (although Ag-STAT RNA was present; Figure 2). It was necessary to include vanadate at a final concentration of 100 μM, as a phosphatase inhibitor, in all buffers used to prepare the nuclear extracts to obtain stable binding. In the absence of vanadate, multiple inconsistent bands and smears were observed. For EMSA reactions, a 20 μl sample volume was used containing 5% glycerol, 20 mM HEPES pH 7.9, 1 mM EDTA, 100 mM NaCl, 1 mM dithiothreitol (DTT), 3 μg of poly(dI–dC)–poly(dI–dC) (Pharmacia), 6 μg of nuclear extract protein, 0.05% Triton X-100, 5 ng/μl of BSA and 50 000 c.p.m. of radiolabelled DNA probe (~0.5 ng). The upper strand sequence of the double-stranded oligonucleotide used was: 5′-GGATTTTTCCCGGAAATGGTC-3′ and represents the optimal binding site for *Drosophila* STAT (Yan *et al.*, 1996). After addition of the DNA and protein, the samples were incubated for 40 min at room temperature and were then analysed by electrophoresis in a non-denaturing 6% polyacrylamide gel with 5% glycerol and 0.25× TBE, dried and submitted to autoradiography. For the competition experiments, 50 ng of cold oligonucleotide was mixed with the labelled probe before the protein was added. For the supershift experiments, IgG was enriched from the rabbit serum using the caprylic acid purification method, followed by ammonium sulfate precipitation and dialysis with PBS (Harlow and Lane, 1998); for each reaction, 10.5 μg of IgG protein were added and incubated with the nuclear extract for 30 min on ice before adding the labelled DNA probe.

Acknowledgements

This work was funded by grants from the John D. and Catherine T.MacArthur Foundation, the Human Frontiers Science Programme and the Training and Mobility of Researchers programme of the European Union (F.C.K.), and by fellowships from the European Molecular Biology Organization and the Swiss National Science Foundation (C.B.-M.).

References

Abastado,J.P., Miller,P.F. and Hinnebusch,A.G. (1991) A quantitative model for translational control of the *GCN4* gene of *Saccharomyces cerevisiae*. *New Biol.*, **3**, 511–524.

Barillas-Mury,C., Charlesworth,A., Gross,I., Richman,A., Hoffmann,J.A. and Kafatos,F.C. (1996) Immune factor Gambif1, a new rel family member from the human malaria vector, *Anopheles gambiae*. *EMBO J.*, **15**, 4691–4701.

Beck,G., Cardinale,S., Wang,L., Reiner,M. and Sugumaran,M. (1996) Characterization of a defense complex consisting of interleukin 1 and phenol oxidase from the hemolymph of the tobacco hornworm, *Manduca sexta*. *J. Biol. Chem.*, **271**, 11035–11038.

Becker,S., Groner,B. and Muller,C.W. (1998) Three-dimensional structure of the Stat3beta homodimer bound to DNA. *Nature*, **394**, 145–151.

Binari,R. and Perrimon,N. (1994) Stripe-specific regulation of pair-rule genes by hopscotch, a putative Jak family tyrosine kinase in *Drosophila*. *Genes Dev.*, **8**, 300–312.

Chen,X., Vinkemeier,U., Zhao,Y., Jeruzalmi,D., Darnell,J.E.,Jr and Kuriyan,J. (1998) Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell*, **93**, 827–839.

Chow,J.C., Ling,P.R., Qu,Z., Laviola,L., Ciccarone,A., Bistrian,B.R. and Smith,R.J. (1996) Growth hormone stimulates tyrosine phosphorylation of JAK2 and STAT5, but not insulin receptor substrate-1 or SHC proteins in liver and skeletal muscle of normal rats *in vivo*. *Endocrinology*, **137**, 2880–2886.

Copeland,N.G. *et al*. (1995) Distribution of the mammalian STAT gene family in mouse chromosomes. *Genomics*, **29**, 225–228.

Darnell,J.E.,Jr (1997) STATs and gene regulation. *Science*, **277**, 1630–1635.

Deryckere,F. and Gannon,F. (1994) A one-hour minipreparation technique for extraction of DNA-binding proteins from animal tissues. *BioTechniques*, **16**, 405.

Devereux,J., Haeberli,P. and Smithies,O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.*, **12**, 387–395.

Dimopoulos,G., Richman,A., Muller,H.M. and Kafatos,F.C. (1997) Molecular immune responses of the mosquito *Anopheles gambiae* to bacteria and malaria parasites. *Proc. Natl Acad. Sci. USA*, **94**, 11508–11513.

Hanratty,W.P. and Dearolf,C.R. (1993) The *Drosophila* tumorous-lethal hematopoietic oncogene is a dominant mutation in the hopscotch locus. *Mol. Gen. Genet.*, **238**, 33–37.

Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Harrison,D.A., Binari,R., Nahreini,T.S., Gilman,M. and Perrimon,N. (1995) Activation of a *Drosophila* Janus kinase (JAK) causes hematopoietic neoplasia and developmental defects. *EMBO J.*, **14**, 2857–2865.

Haspel,R.L., Salditt-Georgieff,M. and Darnell,J.E.,Jr (1996) The rapid inactivation of nuclear tyrosine phosphorylated Stat1 depends upon a protein tyrosine phosphatase. *EMBO J.*, **15**, 6262–6268.

Hennig,W. (1981) *Insect Phylogeny*. Wiley, New York, NY.

Hoffmann,J.A. and Reichhart,J.-M. (1997) *Drosophila* immunity. *Trends Cell Biol.*, **7**, 309–316.

Hou,X.S., Melnick,M.B. and Perrimon,N. (1996) Marelle acts downstream of the *Drosophila* HOP/JAK kinase and encodes a protein similar to the mammalian STATs [published erratum appears in *Cell* 1996, **85**, following 290]. *Cell*, **84**, 411–419.

Ihle,J.N. (1996) STATs: signal transducers and activators of transcription. *Cell*, **84**, 331–334.

Kaplan,M.H., Sun,Y.L., Hoey,T. and Grusby,M.J. (1996) Impaired IL-12 responses and enhanced development of Th2 cells in STAT4-deficient mice. *Nature*, **382**, 174–177.

Kawata,T., Shevchenko,A., Fukuzawa,M., Jermyn,K.A., Totty,N.F., Zhukovskaya,N.V., Sterling,A.E., Mann,M. and Williams,J.G. (1997) SH2 signaling in a lower eukaryote: a STAT protein that regulates stalk cell differentiation in *Dictyostelium*. *Cell*, **89**, 909–916.

Leaman,D.W., Leung,S., Li,X. and Stark,G.R. (1996) Regulation of STAT-dependent pathways by growth factors and cytokines. *FASEB J.*, **10**, 1578–1588.

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

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

Liu,X., Robinson,G.W., Gouilleux,F., Groner,B. and Hennighausen,L. (1995) Cloning and expression of STAT5 and an additional homologue (STAT5b) involved in prolactin signal transduction in mouse mammary tissue. *Proc. Natl Acad. Sci. USA*, **92**, 8831–8835.

Mathey-Prevot,B. and Perrimon,N. (1998) Mammalian and *Drosophila* blood: JAK of all trades? *Cell*, **92**, 697–700.

McDonald,N.Q. and Hendrickson,W.A. (1993) A structural superfamily of growth factors containing a cystine knot motif. *Cell*, **73**, 421–424.

Meraz,M.A. *et al*. (1996) Targeted disruption of the STAT1 gene in mice

reveals unexpected physiologic specificity in the JAK–STAT signaling pathway. *Cell*, **84**, 431–442.
Morisato,D. and Anderson,K.V. (1994) The *spätzle* gene encodes a component of the extracellular signaling pathway establishing the dorsal–ventral pattern of the *Drosophila* embryo. *Cell*, **76**, 677–688.
O'Shea,J.J. (1997) Jaks, STATs, cytokine signal transduction and immunoregulation: are we there yet? *Immunity*, **7**, 1–11.
Perrimon,N. and Mahowald,A.P. (1986) l(1)hopscotch, a larval–pupal zygotic lethal with a specific maternal effect on segmentation in *Drosophila*. *Dev. Biol.*, **118**, 28–41.
Qiu,P., Pan,P.C. and Govind,S. (1998) A role for the *Drosophila* Toll/Cactus pathway in larval hematopoiesis. *Development*, **125**, 1909–1920.
Richman,A.M., Bulet,P., Hetru,C., Barillas-Mury,C., Hoffmann,J.A. and Kafalos,F.C. (1996) Inducible immune factors of the vector mosquito *Anopheles gambiae*: biochemical purification of a defensin antibacterial peptide and molecular cloning of preprodefensin cDNA. *Insect Mol. Biol*., **5**, 203–210.
Salazar,C.E., Mills-Hamm,D., Kumar,V. and Collins,F.H. (1993) Sequence of a cDNA from the mosquito *Anopheles gambiae* encoding a homologue of human ribosomal protein S7. *Nucleic Acids Res*., **21**, 4147.
Takeda,K. *et al*. (1996) Essential role of STAT6 in IL-4 signalling. *Nature*, **380**, 627–630.
Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
Udy,G.B., Towers,R.P., Snell,R.G., Wilkins,R.J., Park,S.H., Ram,P.A., Waxman,D.J. and Davey,H.W. (1997) Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. *Proc. Natl Acad Sci. USA*, **94**, 7239–7244.
Williams,M.J., Rodriguez,A., Kimbrell,D.A. and Eldon,E.D. (1997) The 18-wheeler mutation reveals complex antibacterial gene regulation in *Drosophila* host defense. *EMBO J.*, **16**, 6120–6130.
Yan,R., Small,S., Desplan,C., Dearolf,C.R. and Darnell,J.E.,Jr (1996) Identification of a STAT gene that functions in *Drosophila* development. *Cell*, **84**, 421–430.
Zhong,Z., Wen,Z. and Darnell,J.E.,Jr (1994) Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science*, **264**, 95–98.

Received August 25, 1998; revised and accepted December 18, 1998