Insect Biochemistry and Molecular Biology 41 (2011) 968–981

Contents lists available at SciVerse ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Evidence of a novel immune responsive protein in the Hymenoptera

Štefan Albert [a,b,*], Heike Gätschenberger [a], Klara Azzami [a], Olaf Gimple [a], Gudrun Grimmer [a,b], Seirian Sumner [c], Tomoko Fujiyuki [d], Jürgen Tautz [a], Martin J. Mueller [b]

[a] BEEgroup, Biozentrum, University of Würzburg, Am Hubland, 97074 Würzburg, Germany
[b] Pharmaceutical Biology, Julius-von-Sachs Institute, University of Würzburg, Julius-von-Sachs-Platz 2, 97082 Würzburg, Germany
[c] Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, United Kingdom
[d] Laboratory of Microbiology, Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1, Hongo, Bunkyo, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:
Received 15 July 2011
Received in revised form
20 September 2011
Accepted 27 September 2011

Keywords:
Honeybee
Hymenoptera
Innate immunity
Carboxylesterase
Haemolymph
Phylogeny

ABSTRACT

Honeybee populations are severely threatened by parasites and diseases. Recent outbreaks of Colony Collapse Disorder (CCD) has caused loss of more than 35% of bee colonies in the USA, and this is thought to at least in part be due to parasites and/or disease. Interestingly, the honeybee possesses of a limited set of immune genes compared to other insects. Non-canonical immune genes of honeybee are of interest because they may provide greater insights into the peculiar nature of the immune system of this social insect. Previous analyses of bee haemolymph upon bacterial challenge identified a novel leucine-rich repeat protein termed IRP30. Here we show that IRP30 behaves as a typical secreted immune protein. It is expressed simultaneously with carboxylesterase upon treatment with bacteria or other elicitors of immune response. Furthermore we characterize the gene and the mRNA encoding this protein and the IRP30 protein itself. Its regulation and evolution reveal that IRP30 belongs to a protein family, distributed broadly among Hymenoptera, suggesting its ancient function in immune response. We document an interesting case of a recent *IRP30* loss in the ant *Atta cephalotes* and hypothesize that a putative IRP30 homolog of *Nasonia* emerged by convergent evolution rather than diverged from a common ancestor.
© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Bacterial infection with pathogenic, facultative pathogenic, or non-pathogenic but proliferating bacteria is a serious life-threatening event for the affected organism. To protect themselves against infections and parasites, metazoans evolved an array of strategies ranging from passive barriers (skin, cuticle) to active responses (cellular and humoral), which may be constitutive or inducible in their nature. In contrast to higher metazoans, insects lack adaptive immunity, at least in the form as we define it, i.e. they do not express hypervariable immunoglobins. On the other hand insects evolved a unique strategy (encapsulation) based on deposition of melanin on the surface of invaders leading to their immobilization, isolation and subsequent killing. Insect immune defense is extremely effective, eradicating nearly all invaders within minutes (Haine et al., 2008). Insect immunity was optimized and proved successful through millions of years of evolution, resulting in the radiation of insects in virtually all environmental niches worldwide.

The honeybee *Apis mellifera*, a model eusocial insect, lives in densely populated, self air-conditioned colonies, consisting of about 50,000 of genetically closely related individuals in summer and about 20,000 long-living bees in winter (Tautz, 2008). A bee colony is characterized by overlapping generations, mutual feeding and long-term food storage. These conditions are known to promote proliferation of microorganisms resulting in rapid outbreaks of infection. Therefore the honeybee is predicted to possess massive immunity, with increased amounts and variability of immune-related genes. The sequence of the honeybee genome revealed that just the opposite is true: bees possess about half the amount of immune genes found in *Drosophila* (Evans et al., 2006). As a striking example, the genome of the mosquito *Anopheles gambiae* encodes nine prophenoloxidase genes, whereas the genome of the honeybee encodes just one (Consortium, 2006). However, another possibility would be that the honeybee evolved additional immune defense strategies relying on species-specific effectors.

In our previous work we characterized the humoral immune response of honeybees challenged with *Escherichia coli*. As expected, the infection induced hymenoptaecin, defensin and abaecin

Abbreviations: IRP30, immune responsive protein of 30 kDa; CE, carboxylesterase; PO, phenoloxidase; LPS, lipopolysaccharide; PGN, peptidoglycan.

* Corresponding author. Pharmaceutical Biology, Julius-von-Sachs Institute, University of Würzburg, Julius-von-Sachs-Platz 2, 97082 Würzburg, Germany. Tel.: +49 931 3180189; fax: +49 931 3186182.
E-mail address: Stefan.Albert@mail.uni-wuerzburg.de (Š. Albert).

0965-1748/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2011.09.006

as antimicrobial peptides in bee haemolymph. In addition, an infection-specific induction of four proteins in adult bees has been observed. Three of these proteins were previously characterized and are known to play important roles in counteracting invading microorganisms in insects. The phenoloxidase (PO) is involved in wound healing and encapsulation of microbes and parasites (Cerenius and Söderhall, 2004). For peptidoglycan recognition proteins (PGRPs) it was shown that they bind to peptidoglycan, a bacterial cell wall component (Steiner, 2004), and some of these proteins even possess an amidase activity and are therefore capable of digesting their substrate (Mellroth, 2002). Insect carboxylesterases (CEs) were shown to be mutated or upregulated in insects, which evolved resistance against organophosphate insecticides (Pan et al., 2009), However, some CE isoforms are induced by *E. coli* infection in *Bombyx mori* (Shiotsuki and Kato, 1999). A fourth protein with a molecular mass of 37 kDa did not have homologs in other organisms (Randolt et al., 2008). At the time of its identification, the gene and encoded protein were completely unknown and annotated as "hypothetical protein of 30 kDa", which we abbreviated as HP30. As the function of the protein in innate immune system becomes more evident, we now term the protein as Immune Responsive Protein 30 (IRP30). The aim of this work is to characterize *IRP30* gene and encoded protein, demonstrate its involvement in immune response and determine the role of IRP30 in immune response system of *A. mellifera* and other Hymenoptera.

2. Experimental procedures

2.1. Chemicals and enzymes

Chemicals were purchased from Sigma (St Louis, USA) and Applichem, Darmstadt, Germany and Roth, Karlsruhe, Germany.

Restriction endonucleases, Taq polymerase, T4 DNA ligase, pJET1.2 cloning kit were purchased from Fermentas, Lithuania. Phusion DNA polymerase was from New England Biolabs, USA. Electrophoretic chambers were from PeqLab, Germany.

ToPo PCR cloning kit was purchased from Invitrogen (Carlsbad, USA), Absolute CYBR Green Quantitative Real-Time PCR kit was from Thermo Fisher Scientific (USA).

Primers were custom-synthesized by Metabion, Martinsried, Germany; rabbit polyclonal anti-IRP30 serum was made against purified heterologously expressed protein by Immunoglobe, Himmelstadt, Germany.

2.2. Sample collection

Honeybee workers (predominantly freshly emerged nurses), drones and queens were collected from healthy naturally mated colonies held near the bee station of the BEEgroup, University of Würzburg.

Individuals from the bumblebee *Bombus terrestris* were obtained from colonies held at the Department of Zoology II, University of Würzburg. Individuals from the wasp *Vespula germanica* were caught from wild colonies near the University of Würzburg. *Polistes canadensis* was obtained from Panama (ANAM permit to S. Sumner SE/A-33-99) and *Polistes dominulus* was collected from Italy by Rita Cervo (University of Florence). Ants were from colonies cultivated in the Department of Zoology II, University of Würzburg. Additional *Atta cephalotes* ants originating from a distant location (Trinidad and Tobago) were collected from a colony maintained at the Zoological Society of London, UK.

cDNA libraries from *Vespula squamosa* and *Bombus ignitus* were kindly donated by Mike Goodisman (Atlanta, USA) and Seok-Jo Hwang (Suwon, South Korea), respectively.

2.3. Bacteria and cell wall components for mock infection

Gram-negative bacteria *E. coli* (DSM 682) were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, Braunschweig, Germany), whereas the Gram-positive *Micrococcus flavus* were a gift from Dr. U. Rdest (Institute of Microbiology, University of Würzburg). The *E. coli* strain was cultivated in NB medium (5 g Nutrient broth, 5 g Bacto peptone, and 10 g NaCl per liter). *M. flavus* was grown in LB medium (5 g Bacto yeast extract, 10 g Bacto tryptone, and 10 g NaCl per liter). All ingredients were purchased from Becton Dickinson (Heidelberg, Germany). The bacteria were grown to an absorbance of $A_{550} = 0.5$. After centrifugation for 5 min at 3,000 *g*, the bacterial cells were washed two times with phosphate-buffered saline (PBS), resuspended and diluted to the desired concentration (10^4–10^5 cells/μl) with PBS.

The tested cell wall components were all purchased from Sigma, Germany. For septic injuries Lipopolysaccharide (LPS) as cell wall component of Gram-negative bacteria (No. L2630), peptidoglycan (PGN) as cell wall component of Gram-positive bacteria (No. 53243) and laminarin as cell wall component of fungi (No. L9634) were used.

2.4. Bacterial infection and haemolymph collection

Bacterial infection and haemolymph collection were performed as described previously (Randolt et al., 2008). Briefly, animals were anesthetized on ice and fixed. Bacterial suspensions were injected intra-abdominally between ventral tergites of adult animals. Larvae were injected into their dorsal cuticle. Aseptic injuries were done in a similar way with sterile PBS. Haemolymphs have been collected at indicated time points (usually 24 h post infection) from the abdomens or thoraces using capillaries.

2.5. One-dimensional SDS polyacrylamide gel electrophoresis (SDS–PAGE)

The gel electrophoresis was carried out in vertical polyacrylamide gels (8.5 × 13 × 0.1 cm) containing 0.1% SDS with a 1.5-cm-long 5% stacking gel on top of the separating gel (Laemmli, 1970). Haemolymph samples were diluted with 2× concentrated Laemmli sample buffer (100 mM Tris–HCl, pH 6.8, 4% SDS, 17% glycerol and 0.8 M 2-mercaptoethanol), heated for 5 min at 95 °C and subjected to electrophoresis at constant voltage (120 V). 10% polyacrylamide/0.1% SDS gels were run for the separation of proteins in the range of 30–200 kDa and 15% polyacrylamide/0.1% SDS gels for the separation of proteins in the range of 5–30 kDa, respectively. For colloidal Coomassie staining, the gels were first fixed for 30 min in 0.85% o-phosphoric acid/20% methanol followed by staining overnight in a solution of Roti-Blue (Roth, Germany) and 20% methanol according to the manufacturer's instructions. Gels were destained in 25% methanol.

2.6. Two-dimensional gel electrophoresis

For 2D-gels, the pooled haemolymph of LPS-injected and control bees (10 individuals each) were centrifuged at 10,000 × *g* for 7 min at 4 °C and protein concentration has been determined.

2D-Electrophoresis was performed by combining isoelectrofocusing (IEF) on 18 cm IPG strips (GE Healthcare), pH range 3–10 NL, for the first dimension and 12.5% polyacrylamide SDS–PAGE for the second dimension. IPG strips were loaded with 340 μl of DeStreak rehydration solution (GE Healthcare) containing 120 μg of protein sample and 0.5% IPG buffer (GE Healthcare).

Rehydration of the Immobiline dry strips and IEF was carried out at 20 °C using the Ettan IPGphor II equipment (GE Healthcare) according to the following program: 12 h at 50 V; 500 V for 1 h; 1000 V (gradient) for 1 h; 8000 V (gradient) for 3 h and 8000 V for 1 h.

Prior to SDS–PAGE, the IPG strips were first equilibrated for 15 min in 15 ml equilibration buffer 1 (6 M urea, 75 mM Tris–HCl pH 8.8, 29.3% glycerol, 2% SDS, 0.002% bromophenol blue, 1% DTT) and later in 15 ml equilibration buffer 2 (6 M urea, 75 mM Tris–HCl pH 8.8, 29.3% glycerol, 2% SDS, 0.002% bromophenol blue, 2.5% iodoacetamide) for 15 min. After equilibration, the strip was placed on top of a lab-cast SDS–polyacrylamide gel (12.5%, 1.0 mm) and covered with agarose sealing solution (25 mM Tris base, 192 mM glycine, 0.1% SDS, 0.5% agarose, 0.002% bromophenol blue).

Electrophoresis was performed for 45 min at 60 V, followed by constant 160 V until the bromophenol blue front reached the bottom of the gel.

Proteins were visualized by colloidal Coomassie staining as described above and evaluated using Delta2D software from Decodon.

2.7. IRP30 gene characterization and genomic organization

Genomic DNAs were isolated by phenol/chloroform extraction and subsequent ethanol precipitation from the animal heads and thoraces. Sequences of primers used in this work are listed in Supplementary Table I.

IRP30 gene and cDNA cloning was performed *via* polymerase chain reaction (PCR) using high fidelity Phusion polymerase to minimize errors introduced by DNA polymerase. Where necessary (such as for subsequent gene manipulations), the PCR primers were modified to contain 5′ overhangs with recognition sites for restriction endonucleases.

IRP30-related sequences from other species were obtained either *via* 5′ and 3′ rapid amplification of cDNA ends (5′ and 3′ RACE) or by low stringency PCR using primers priming in the conserved regions (Albertova et al., 2005). For further characterization, PCR products were usually inserted into pCR2.1 or pJET1.2 vectors and sequenced.

2.8. DNA, RNA isolation, cDNA synthesis, quantitative RT-PCR

Total RNA was extracted by TRIzol, (Invitrogen, USA) from the whole animals (>5 individuals) or dissected tissues stored in RNAlater (Invitrogen, USA). RNA concentration was determined and its quality tested by gel electrophoresis. RNAs that passed quality control (1 μg) were used for cDNA synthesis using Superscript III cDNA synthesis kit (Invitrogen, USA) and oligo(dT)17 primer. Obtained cDNAs were diluted (1:20 or 1:50) and used as a template for qRT-PCR (Absolute SYBR Green qRT-PCR, Thermo Fisher Scientific) on a PCR cycler with an online detection of fluorescence (CFX96, BioRad, USA). $\Delta\Delta C_T$ method was used to calculate the relative gene expression. *Actin* mRNA has been initially used as an internal standard. However we observed that amount of *actin* mRNA decreases significantly during pupal stages. For this reason we tested additional standards such as ribosomal protein S18, *rps18*, and glyceraldehyde-phosphate dehydrogenase, *GAPDH* (Scharlaken et al., 2008). Both mRNAs exhibited stable expression, which was not affected by bacterial infection. All qRT-PCR experiments in this work were normalized using the *rps18* mRNA.

Homologs of *rps18* were also used for normalization in expression studies of the bumblebee *B. terrestris* and the ant *A. cephalotes*. Also in these species *rps18* transcription did not differ between infected and non-infected states.

2.9. Generation of antibodies, immunoblotting

E. coli-expressed GST-IRP30 and GST-CE-D from the honeybee (*A. mellifera*) were purified from insoluble inclusion bodies by the modified method of (Nagai and Thogersen, 1987). The final purification step employed preparative gel electrophoresis and subsequent elution of the GST-IRP30 protein.

Female rabbits *Chinchilla bastard* were immunized by a purified GST-IRP30 on a commercial basis using standard protocols. Obtained antiserum was tested and further affinity-purified on an affinity column with immobilized IRP30 protein. This step eliminated GST-reactive antibodies from the serum and concentrated the IRP30-reactive antibodies.

For immunoblots, size-separated proteins were electrophoretically transferred (Western blotting) onto nitrocellulose membranes (Schleicher and Schuell, Germany). Membranes were blocked by incubation with 5% skim milk in Tris-buffered saline, 0.05% Tween/20 (TBST), incubated for 1 h with α-IRP30 or antiserum or affinity-purified antibodies diluted 1:1000 or 1:5000 respectively. After extensive washing with TBST, the membranes were incubated with secondary α-rabbit, horseradish peroxidase-conjugated antibodies (Sigma, USA). Detection of immunoreactive proteins was done with enhanced chemiluminescence detection kit and autoradiography (GE Healthcare, USA).

2.10. Far-Western blotting, dot blotting, bacterial sedimentation assay

For detection of interaction between IRP30 and cell wall components, 10-fold dilutions of LPS, laminarin and peptidoglycan (20 mg–20 ng) were dropped onto a nitrocellulose filter and let to dry. The filters were first incubated for 1 h with 5% skim milk in phosphate-buffered saline (PBS), then overnight with infected bee haemolymph (containing large amounts of IRP30) diluted 1:100 in PBS. Later the membranes were washed 3× with TBST, incubated with anti-IRP30 antiserum and processed as described for immunoblotting (see above).

To assess the IRP30 interaction with living bacteria, haemolymph from bees expressing IRP30 was incubated for 15 min with living Gram-negative *E. coli* or Gram-positive *M. flavus* and centrifuged. Supernatant and pellet were separated and electrophoresed (SDS–PAGE). Obtained gel was blotted on nitrocellulose membrane and incubated with antiserum against IRP30 as described for immunoblotting.

2.11. Immunoprecipitation of IRP30

Haemolymph of infected and non-infected young adult bees (50 μl each group) was collected and diluted 1:1 with fresh sterile TBS. After clearing with Protein-A-Sepharose (without immobilized antibodies) the haemolymphs were incubated with 15 μl of IRP30-antibody-coated Protein-A-Sepharose at 4 °C for 1 h. The beads were washed (4 × 1 ml TBS). For elution, Laemmli sample loading buffer was added to the beads and the samples were boiled for 5 min and loaded on a SDS–PAGE. The gel was stained with Coomassie Brilliant Blue G250.

2.12. Bioinformatic analyses

Gene sequences were assembled using GCG program package. Homology searches were performed using BLAST (Altschul et al., 1997). Iterated BLAST searches (psiBLAST) recovered thousands of leucine-rich repeat containing domains. Therefore we set a criterion that the presence of LRR domain in combination with flanking regions and leucine zipper at the C-terminus are required for IRP30 identity.

Sequence alignment and phylogenetic analyses were conducted using CLUSTAL-X (Jeanmougin et al., 1998). Promoter and protein sequence analyses were performed using the Network Protein Sequence Analysis server (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl%3fpage=/NPSA/npsa_server.html).

3. Results

3.1. Genetic and biochemical characterization of IRP30

3.1.1. IRP30 gene characterization

A protein of 37 kDa induced by bacterial infection with *E. coli* was identified by MS–MS analysis of tryptic peptides (Randolt et al., 2008). The determined peptide sequences matched the peptides encoded by a gene locus LOC 408807 coding for a hypothetical protein of 30 kDa (formerly HP30, now termed IRP30, GenBank accession number JN104327).

Differently sized fragments (~800 bp and ~1200 bp) were obtained through PCR from the cDNA library and genomic DNA, respectively (Fig. 1). Comparison of the sequences of the genomic and cDNA fragments revealed a single intron of 400 bp that interrupts the *IRP30* gene between the codons 39 and 40. The sequenced *IRP30* DNA was identical with the predicted cDNA derived from the locus LOC 408807, except for one A–C transversion at the position of 631 of LOC 408807. Despite occurring in the first base of the coding triplet, this substitution does not change the encoded amino acid, arginine (Fig. 1). The *IRP30* locus maps to the central part of the chromosome 4 in an area comparatively poor in genes; there are no genes 6 kbp upstream and 18 kbp downstream of the *IRP30* locus.

We searched the region 2 kbp upstream of *IRP30* for the presence of regulatory motifs, which could serve as binding sites of different transcription factors. We found several motifs conserved among metazoans including c-Myc, Adf1 and, importantly, six motifs recognized by NFκB. Relish, the insect homolog of NFκB, was shown to participate in the immune response of honeybees (Schluns and Crozier, 2007). For the time being it is not possible to determine individual contributions of these regulatory motifs to the unique expression pattern of *IRP30* experimentally, but for example, both honeybee *defensin* genes, upregulated upon infection, were shown to contain two NFκB binding motifs (Klaudiny et al., 2005).

3.1.2. Biochemical properties of IRP30 protein

The obtained α-IRP30 antiserum recognized a single protein of ~37 kDa in the haemolymph as well as in the whole body extracts of infected honeybees. Non-infected bees generated either none or a very weak signal (Fig. 2B). We concluded that the obtained antiserum recognized the IRP30 polypeptide.

The calculated molecular mass of full-length IRP30 is 30.5 kDa. The protein contains an N-terminal hydrophobic region, with features typical for a signal peptide with a predicted cleavage site between amino acids 17/18. This means that IRP30 is secreted into the haemolymph. The molecular mass of processed IRP30 should be 28.7 kDa. Secreted proteins are often modified by glycosylation on asparagine residues (N-glycosylation). IRP30 contains two potential N-glycosylation sites at glutamines 56 and 76 (Fig. 1A, boxed). To test the N-glycosylation of IRP30, the haemolymph of infected bees was treated with peptide N-glycosidase F (PNGaseF), which removes carbohydrate residues attached to asparagine(s) (Fig. 2A). Faster electrophoretic mobility of IRP30 in PNGaseF-treated sample confirms that IRP30 is a glycoprotein. The apparent mobility of the deglycosylated IRP30 in SDS–PAGE is still significantly slower than its calculated mass. Therefore additional modifications cannot be ruled out. Modifications changing the net charge of the protein (such as phosphorylation) are improbable, since the calculated isoelectric point of IRP30 (pI = 6.1), coincides well with its mobility in the pH gradient (see Fig. S1).

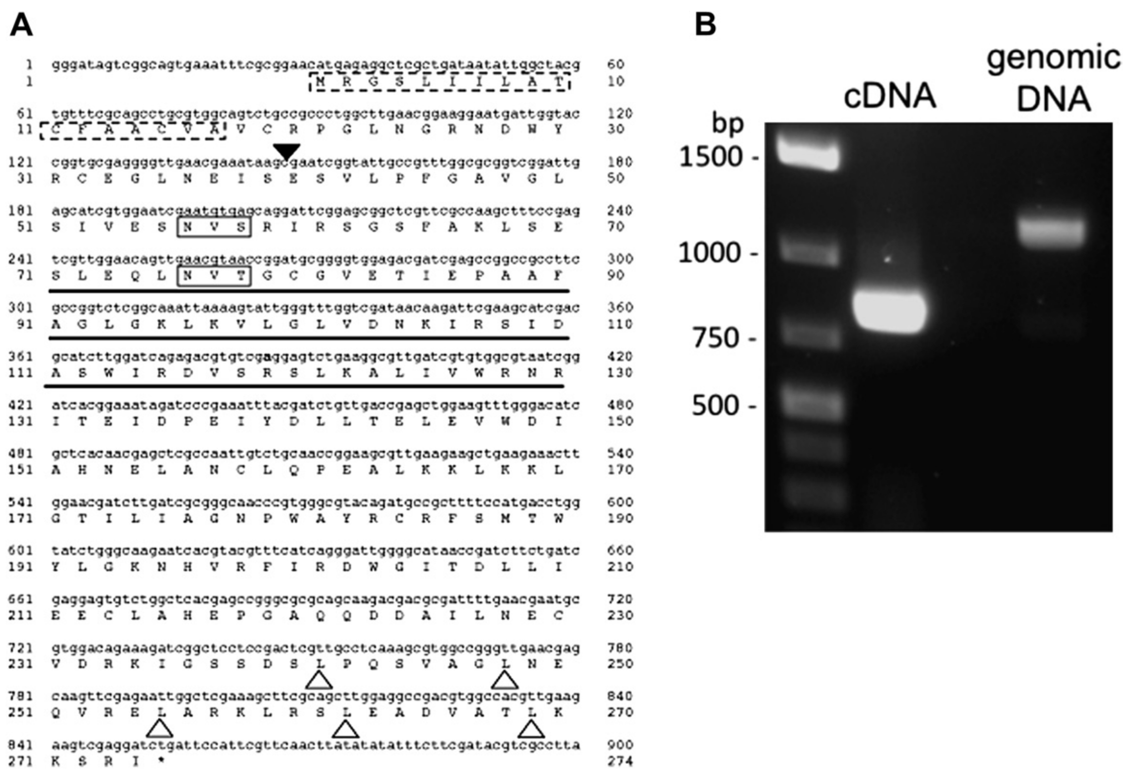

Fig. 1. IRP30 gene and encoded protein. A. cDNA sequence of *IRP30* gene and its translation. Dark triangle indicates the position of an intervening sequence (intron) removed in the process of RNA maturation. Signal peptide is boxed, leucine-rich repeat is underlined, leucines belonging to leucine zipper are marked with empty triangles. B. PCR amplifications of *IRP30* cDNA and gene. Primers designed for PCR amplification of *IRP30* were used with honeybee cDNA (left) and genomic DNA (right) as a template. DNA size marker is on the very left.

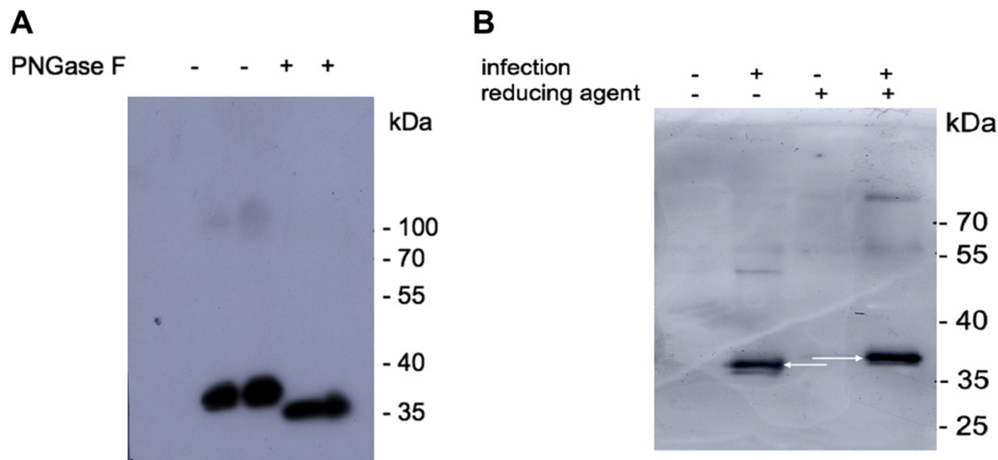

Fig. 2. Biochemical properties of IRP30 protein. A. IRP30 is modified by N-glycosylation. Proteins in two infected haemolymph samples were depleted of polysaccharides by incubation with Protein N-glycosidase F (PNGaseF). Untreated controls and samples were analyzed by immunoblotting. Faster mobility of PNGaseF-treated samples implicates presence of glycosylated asparagine residue(s). B. IRP30 structure is stabilized by disulfide bridges. Non-infected and infected haemolymphs were electrophoresed on a gel in the absence or presence of reducing agent (β-mercaptoethanol). Slight shift of IRP30 mobility in the absence of reducing agent indicates presence of disulfide bridge(s).

Next, we subjected the infected haemolymph to SDS–PAGE under reducing and non-reducing conditions (Fig. 2B). Slight shift in the mobility of IRP30 protein indicates that its structure is stabilized by one or more disulfide bridges.

In summary IRP30 is a slightly acidic glycoprotein, whose expression is induced by components of cell walls of different microorganisms.

3.1.3. *IRP30 is not a strong interactor*

The primary sequence of IRP30 was sought for structural motifs. Two motifs involved in protein interaction were identified: a leucine-rich repeat (LRR) motif located in the N-terminal half of IRP30 and a leucine zipper located at the very C-terminus. These motifs in the IRP30 structure predispose it to interact with other proteins or compounds. Several insect LRR-containing proteins were shown to interact with pathogen's surface molecules to elicit the immune reaction (Osta et al., 2004; Zhu et al., 2010).

Interaction of IRP30 with living bacteria was tested in a bacterial sedimentation assay with Gram-negative *E. coli* and Gram-positive *M. flavus*. In both cases, IRP30 protein did not co-sediment with bacteria (Fig. S2A). Interaction with common elicitors of insect immunity was tested by dot blot method with Far-Western detection using immobilized LPS, peptidoglycan and laminarin in a concentration range 20 ng–2 mg/ml. IRP30 did not bind to any of these compounds *in vitro* (not shown).

Interaction of IRP30 with components of bee haemolymph was tested by immunoprecipitation. No interacting proteins were identified by this method (Fig. S2B).

Taken together, IRP30 does not interact, at least under conditions used, with tested living bacteria, isolated cell wall components of Gram-negative, Gram-positive bacteria and fungi or any components of bee haemolymph.

3.2. *Immune response of IRP30 in the honeybee*

3.2.1. *Role of IRP30 in an immune response of the honeybee*

IRP30 expression can be elicited effectively by viable *E. coli*, but not by aseptic injury (Randolt et al., 2008). We further tested its induction with different components of microbial cell walls such as lipopolysaccharides (LPS, Gram-negative bacteria) or laminarin (fungi). Both substances induced IRP30 expression effectively (Fig. 3A). Gram-positive *M. flavus* was as efficient as laminarin (not shown). In 2-dimensional electrophoresis of bee haemolymph (Fig. S1), besides previously identified IRP30, CE, PGRP-S2 (Randolt et al., 2008) we observed a slight induction of a putative serine protease (XP_623150), which might be involved in phenoloxidase activation.

Time course experiments revealed that IRP30 becomes detectable 12 h post infection (p.i.) and its amounts increase until 24 h p.i. (Fig. 3B). Little to no increase has been observed between 24 and 48 h. The same time-dependent expression pattern was seen for CE-D (Fig. 3B).

3.2.2. *Expression of IRP30 in bee castes, during bee life and colony cycle*

Three morphological castes co-exist in a honeybee colony: male drones, female workers and queens. Haemolymph samples of adults representing each of the castes were tested by SDS–PAGE with the result that *IRP30* expression in all castes was consistently upregulated upon infection (Fig. S3).

Protein extracts from different developmental stages (egg to adult) were tested by immunoblotting with the antisera against IRP30 and CE-D. Neither IRP30 nor carboxylesterase could be detected (Fig. S4A). We further tested infected worker larvae for expression of IRP30. Clearly, bee larvae do not produce IRP30 even after an immune challenge (Fig. S4B).

Surprisingly, in contrast to summer bees, non-infected winter bees constitutively expressed small amounts of IRP30 in their haemolymph. The IRP30 amounts in these bees increased in response to bacterial challenge (Fig. 4A).

In order to quantify the IRP30 and CE-D mRNAs, we designed primers for quantitative real-time polymerase chain reaction (qRT-PCR, see Supplementary Table I). The primers were used with cDNAs prepared from different developmental stages of the bee. We observed only minuscule expression of both mRNAs in early pupa stages. Infection of adult bees resulted in nearly 500-fold and 1200-fold induction of IRP30 and CE-D mRNAs respectively (Fig. S5).

3.2.3. *Organs expressing IRP30 in the honeybee*

We next asked which organs/tissues synthesize the IRP30. Detection of IRP30 by immunoblotting yielded IRP30 signals in the extracts of three basic body segments (head, thorax, abdomen), which is not surprising because the haemolymph circulates

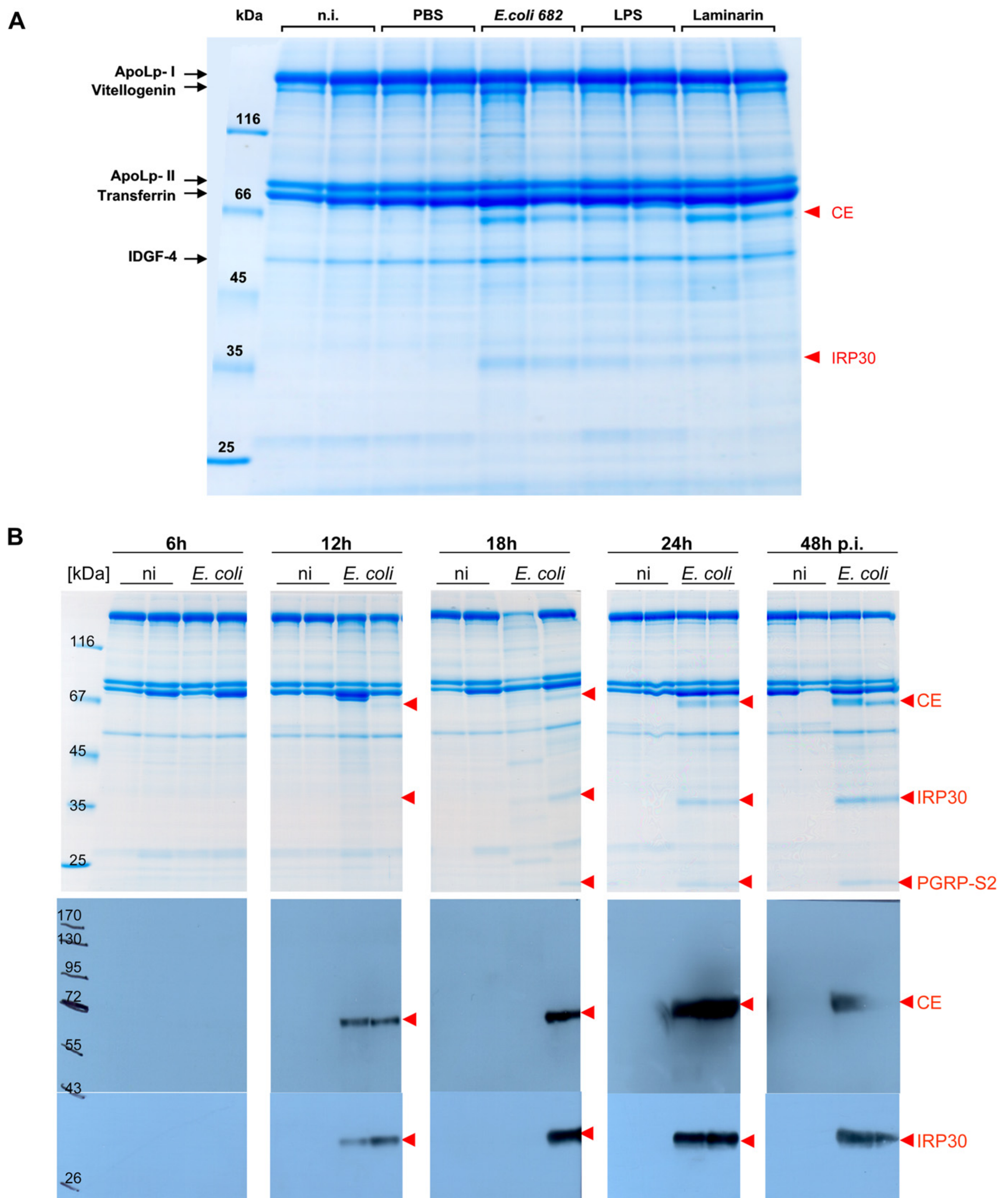

Fig. 3. Induction of IRP30 and carboxylesterase in bee haemolymph. A. IRP30 is induced by living bacteria, lipopolysaccharides (LPS) and laminarin. Freshly emerged honeybees (always a minimum of 10 animals) were: not treated (n.i.), injected with sterile PBS, *E. coli* (10^4 cells), LPS or laminarin resuspended in PBS. Haemolymph samples of at least 5 individuals were collected 24 h p.i., electrophoresed (10% SDS−PAGE) and stained with Coomassie blue G250. Each lane was loaded with 1.3 µl haemolymph of one individual. Constitutively expressed proteins like Apolipophorin-I and -II (ApoLp-I and -II), Vitellogenin, Transferrin and Imaginal Disc Growth Factor-4 (IDGF-4) are indicated by black arrows, infection-inducible proteins (IRP30 and CE-D) by red arrows. B. IRP30 levels increase continuously after 12 h post infection. Freshly emerged honeybees were either left not treated (n.i.) or injected with 10^5 cells *E. coli*. Haemolymph samples were collected at indicated time points (6, 12, 24 and 48 h p.i.). Each lane represents 1.3 µl haemolymph of one individual electrophoresed on two parallel gels. One gel was stained, the other one was analyzed by immunoblotting with antisera against IRP30 and CE-D. Both proteins appear simultaneously 12 h p.i. in *E. coli*-infected individuals and their amounts increase until 24 h p.i. A slight decrease of both proteins can be observed after 48 h. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

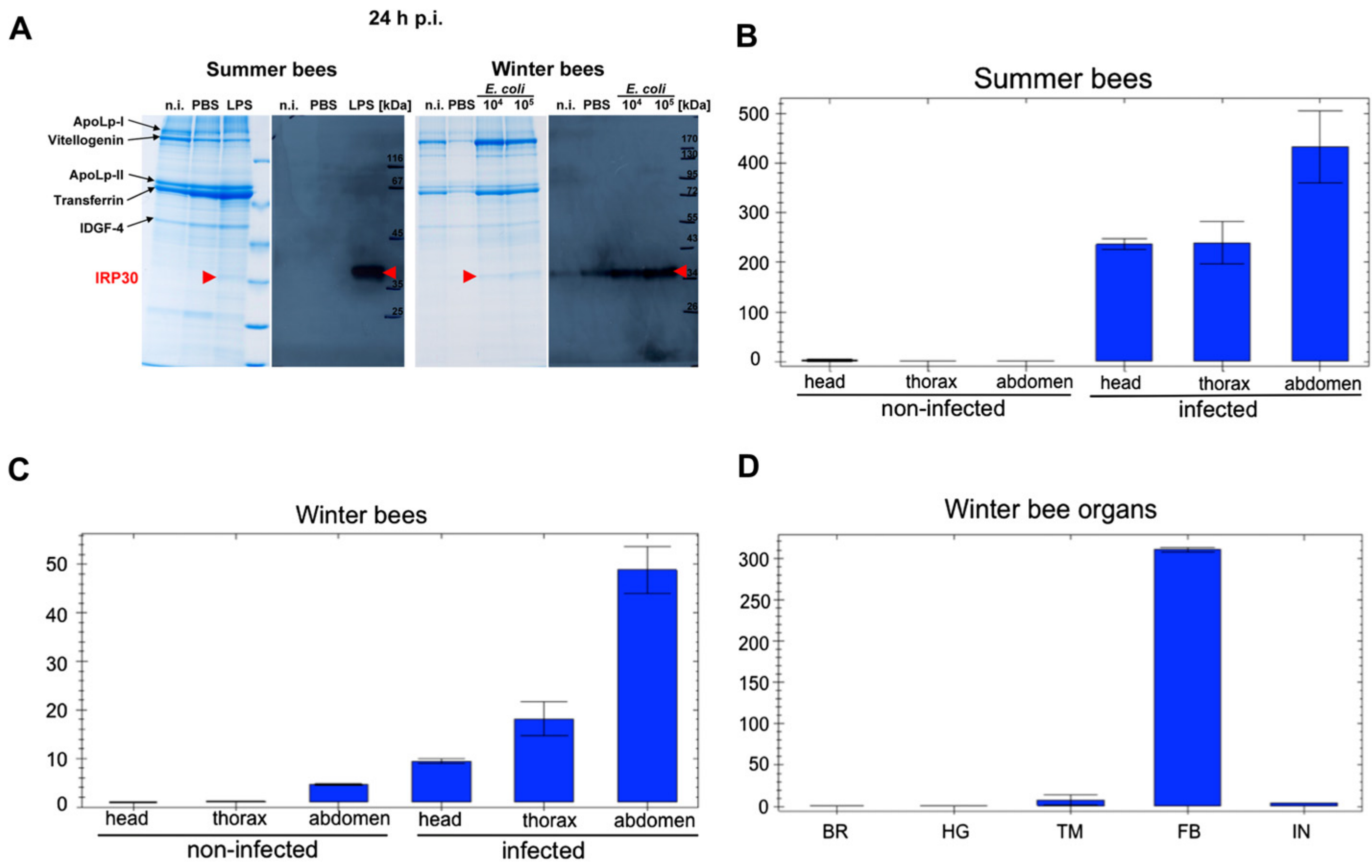

Fig. 4. Regulation of IRP30 at the protein and mRNA level. A. Non-infected winter bees express IRP30 constitutively in their haemolymph. Freshly emerged summer bees (June) were either not treated (n.i.), injected with sterile PBS or LPS (0.5 mg/ml). Winter bees (February) were collected directly from the hive and either left non-treated (n.i.), injected with sterile PBS or with two different concentrations of living *E. coli* (10^4 and 10^5 cells). Haemolymph samples were collected 24 h p.i., each lane represents one individual. 1.3 µl haemolymph of summer bees and 0.3 µl haemolymph of winter bees were electrophoresed and stained (left panels) or immunoblotted with IRP30 antiserum (right panels). B–C. IRP30 expression triggered by immune challenge. Summer (B) and winter workers (C) were injected with a suspension of *E. coli* (10^5 cells). The relative *IRP30* RNA expression level (*y*-axis) in dissected body parts was determined by qRT-PCR 20 h later. Non-infected summer bees do not express *IRP30*, whereas winter bees express IRP30 constitutively in their abdomen. Immune challenge induces *IRP30* expression in all body parts, predominantly in the abdomen. Induction is stronger in summer bees. HE = head, TH = thorax, AB = abdomen. All experiments were repeated at least four times, 5 animals were dissected and pooled for each experiment. D. IRP30 is constitutively expressed by the fat body of winter bees. As infected bees express IRP30 in all body sections (Fig. 4B), organs expressing IRP30 were sought in non-infected winter bees. Fat body expresses several hundreds-fold more IRP30 than the other organs. BR = brain, HG = hypopharyngeal glands, TM = thoracal muscle, FB = fat body, IN = intestine.

through the whole insect body. Quantitative real-time PCR (qRT-PCR) was used to determine the tissue of IRP30 expression because mRNAs do leave the location of their synthesis. We tested total cDNAs of three body segments of infected bees. *IRP30* mRNA was expressed to similar levels in all body parts (Fig. 4B). However, in non-infected winter bees, only the abdomen expressed *IRP30* (Fig. 4C). This result was obtained repeatedly with winter bees from different colonies, which showed no signs of disease. The abdominal cavity of insects encloses digestive and reproductive organs, several glands (i.e. wax, Nasanov or venom glands) and, importantly, the fat body, which was shown to play a central role in insect immunity (Wilson-Rich et al., 2008). Winter bees possess much larger fat bodies than summer bees. We dissected brain, hypopharyngeal glands, thoracal muscle, fat body and intestine from winter bees and prepared cDNAs. These were used for qRT-PCR analysis. Fat bodies of winter bees expressed high amounts of IRP30 mRNA (Fig. 4D) unlike all other tested organs. Bacterial challenge of winter bees induced the expression of IRP30 in all body segments, with highest expression in the abdomen (Fig. 4C).

In summary IRP30 is expressed in fat bodies of adult winter bees. Bacterial challenge might induce the expression of IRP30 in either haemocytes or other tissue(s) wide-spread in the whole body, such as epithelial cells.

3.3. *IRP30 presence and function across the Hymenoptera*

3.3.1. *IRP30 in other social bees*

Data above suggest that IRP30 could be an important component of bee immune response. We set out to characterize *IRP30* genes in other honeybee species. Using the primers derived from the *A. mellifera* IRP30, we amplified homologous sequences from genomic DNAs of *Apis dorsata*, *Apis florea* and *Apis cerana*. All bee IRP30 sequences code for highly related proteins (85–92% identity, 92–97% similarity). Amplified gene fragments contain a single intron (376–486 bp, see Table 1), inserted in the same position and in the same translational phase 0 (between two coding triplets, see Fig. 1A dark triangle).

To test the possibility that IRP30 is an *Apis* specific immune response gene, we searched for immunoreactive IRP30 homologs in the haemolymph samples of another Apidea bee, the bumblebee (*B. terrestris*). Experimental infection of adult bumblebees with *E. coli* led to strong induction of antimicrobial peptides, CE-D and a protein with a molecular mass similar to IRP30 (Fig. 5A and B). The protein was recognized by α-IRP30 antiserum (Fig. 5C). The same antiserum did not recognize any protein in the wasp haemolymph. To identify the *IRP30* gene sequence, we found a EST fragment (accession number EE605521) coding for an IRP30 homolog in Asiatic bumblebee *B. ignitus* (Kim et al., 2006). Using RACE protocol, we

Table 1
Overview of *IRP30* genes from different Hymenopterans.

Species	GenBank ID	Protein length (aa)	Intron length (nt)	Reference	Note
Apis mellifera	JN104327	280	401	(1,2)	
Apis florea	JN181867	281	376/378[c]	This work	LS-PCR
Apis dorsata	JN181868	262[a]	486	This work	LS-PCR
Apis cerana	JN181869	173[a]	432	This work	LS-PCR
Bombus terrestris	JN181870	275	620/625/627[c]	This work	LS-PCR
Bombus ignitus	JN181871	275	n.d.	This work, (3)	PCR
Bombus impatiens	BK008062	275	699	This work	*in silico*
Megalopta genalis		276[a]	n.d.	(4)	
Vespula squamosa	JN181872	274	610	This work, (5)	PCR
Vespula germanica		155	n.d.	This work	LS-PCR
Vespa mandarinia	AB645728	274	720	This work	LS-PCR
Polistes canadensis	JN181873	276	509	This work	
Polistes dominula	JN181874	268[a]	449	This work	LS-PCR
Harpegnathos saltator	EFN75616	280	1674	(6)	*in silico*
Atta cephalotes	JN181880–JN181885	–[b]	295/296	This work	PCR
Atta sexdens	JN181875	268	291	This work	LS-PCR
Atta laevigata	JN181876	268	294/297/293[c]	This work	LS-PCR
Atta vollenweiderii	JN181877	268	292	This work	LS-PCR
Acromyrmex lundii	JN181878	274	290	This work	LS-PCR
Acromyrmex ambiguus	JN181879	274	277	This work	LS-PCR
Acromyrmex echinatior	EGI61553	279	286		*in silico*
Pogonomyrmex barbatus	ADIH01023411	274	299	(7)	*in silico*
Nasonia vitripennis	XP_001603643	280	35; 86	(2)	*in silico*

n.d.: not determined.
LS-PCR: low stringency polymerase chain reaction.
(1) (Randolt et al., 2008).
(2) (Fujiyuki et al., 2009).
(3) (Kim et al., 2006).
(4) (Woodard et al., 2011).
(5) (Hoffman and Goodisman, 2007).
(6) (Bonasio et al., 2010).
(7) (Smith et al., 2011).
[a] The length of encoded protein could not be determined from the available sequence information.
[b] *Atta cephalotes IRP30* gene does not code for a functional protein.
[c] Numbers separated by slash (/) indicate the lengths of introns found in different alleles.

completed the full cDNA sequence of *B. ignitus IRP30*. Employing PCR primers derived from *B. ignitus IRP30* we amplified an *IRP30* homolog from the genomic DNA of *B. terrestris*. These two bumblebee genes encode nearly identical proteins. Recent transcriptome characterization of different bee species (Woodard et al., 2011) revealed additional IRP30-like sequence fragments in many bee species. One IRP30 cDNA (*Megalopta genalis*) could be nearly completely reconstituted. We conclude that the *IRP30* gene is expressed in many bee species, irrespective of their stage of sociality.

3.3.2. IRP30 in social wasps

To determine whether IRP30 was specific to bees, we tested our bee IRP30 antiserum on vespid wasp *V. germanica*. The antiserum did not recognize any protein of *V. germanica* haemolymph (Fig. 5C). However, we found a IRP30-like fragment among ESTs of the wasp *V. squamosa* (Hoffman and Goodisman, 2007). The complete sequence of *V. squamosa* IRP30 cDNA has been determined using RACE protocol. Genomic locus of *V. squamosa* IRP30 has been amplified using primers derived from the cDNA and subsequently sequenced. Employing primers derived from the regions conserved between bee and wasp IRP30 we successfully amplified *IRP30* from V*espa mandarinia* and *V. germanica* (Table 1). An additional *IRP30* homolog was cloned from the two paper wasps *P. canadensis* and *P. dominulus* indicating that IRP30 is broadly distributed in different genera of wasps. However its regulation/function in wasps might be different given that infection did not elicit any peptide with molecular mass similar to IRP30 in *V. germanica* (Fig. 5C).

3.3.3. IRP30 in ants

We further explored the presence and function of IRP30 in ants (Formicidae). IRP30 appears to be absent from the Formicidae currently sequenced (the carpenter ant *Camponotus floridanus* (Bonasio et al., 2010) and the Argentine ant *Linepithema humile* (Smith et al., 2011)). But we found homologous sequences in the primitive Ponerine ant *Harpegnathos saltator*, and close homologs in two Myrmicines – *Pogonomyrmex barbatus* (Smith et al., 2011) and the leafcutting ant *A. cephalotes* (Suen et al., 2011). Interestingly, the *A. cephalotes* sequence includes three frame-shifting mutations: one deletion of four nucleotides, another deletion of a single nucleotide and one insertion of a single nucleotide (Fig. S6). To verify that the mutations are not due to sequencing errors in the genomic sequence of *A. cephalotes*, the PCR-amplified *IRP30* gene was sequenced again. Direct sequencing of the genomic PCR product (Fig. 6A) confirmed two of three frame-shifts (Fig. S7). The insertion of four nucleotides found in the sequenced *Atta* genome was not confirmed. Intriguingly, *IRP30* gene is expressed by *A. cephalotes* workers in an infection-dependent manner (Fig. 6B). In addition, the intron was apparently removed by RNA splicing, as the observed size of the PCR products differ between cDNA and genomic DNA (Fig. 6A). The cDNA amplificate of *A. cephalotes IRP30* was sequenced. The RNA splicing and the two frame-shifting mutations were confirmed. No signs of multiple alleles were observed in sequenced PCR products.

The *A. cephalotes* samples used for genome sequencing (Suen et al., 2011) and the colony used in this work, both originate from Panama. The collection sites (Gamboa and Barro Colorado island) are about 11 km distant from each other (B. Hölldobler, personal communication). It is possible that defective *IRP30* alleles accumulated in the local population of *A. cephalotes* and for some reason were not eliminated by natural selection. In order to gain an insight into *A. cephalotes IRP30* we sequenced the alleles amplified from a colony originating from a geographically isolated population, from Trinidad and Tobago (>2000 km from Panama).

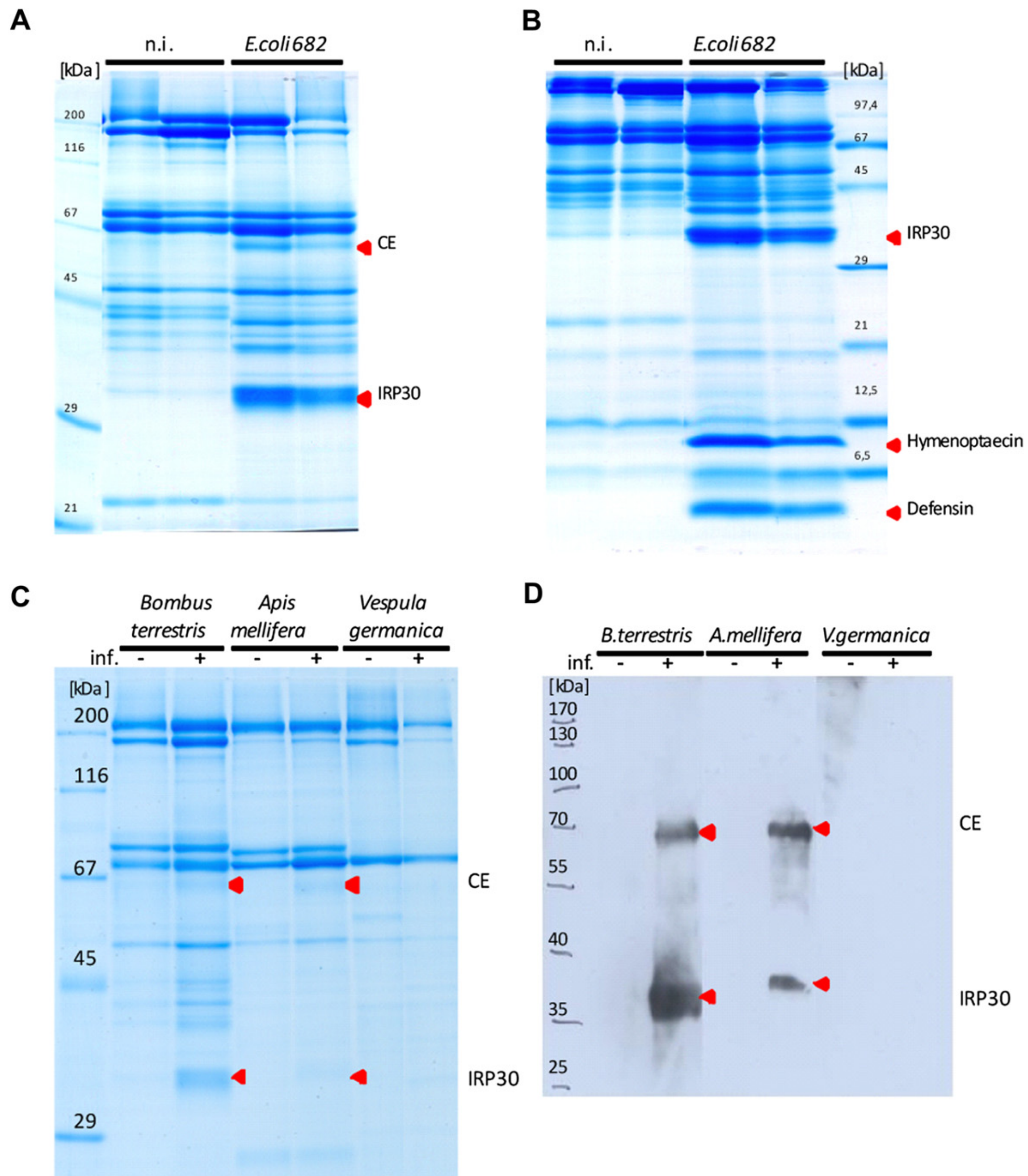

Fig. 5. IRP30 antiserum recognizes a protein in bumblebee but not in wasp. A–B. Infection induced proteins and peptides of bumblebee. Bumblebees *B. terrestris* were either not infected (n.i.) or infected with *E. coli* 682 (10^5 cells). 0.5 μl of haemolymph samples collected 48 h p.i. were electrophoresed (A) in 10% SDS–PAGE (Laemmli, 1970) and (B) 15% Tricine SDS–PAGE (Schagger and von Jagow, 1987). Gels were stained with Coomassie G250. Marked proteins and peptides induced by infection were identified by MS–MS (Randolt et al., 2008). C–D. Detection of IRP30 and CE-D in bumblebees and wasps by immunoblotting. 1 μl of bee, 0.5 μl of bumblebee and 1 μl of wasp haemolymphs were size-separated by 10% SDS–PAGE and stained (C), or electroblotted and incubated with an antiserum against IRP30 and CE-D (D). Infected bumblebees but not wasps express large amounts of an IRP30- and CE-D-antiserum reactive proteins.

Interestingly, we found that *IRP30* from this colony is a mixture of alleles, which differ in the length of guanine stretch (3 vs. 4) at the nucleotide 878 i.e. one allele is free of frame-shift at this position. However, deletion of guanine at 648th nucleotide was confirmed in all sequences. In addition, three novel mutations have been found in Trinidad and Tobago colony (Table 2). Taken together, *A. cephalotes* from Trinidad and Tobago also contain only non-functional *IRP30* alleles, albeit at different stage of decomposition.

This suggests that *A. cephalotes IRP30* passed the initial steps on the way toward its inactivation and is likely to be non-functional in this species. Of note, in the intron region of IRP30 a single deletion of adenine 273 in clone # 6 was detected, whereas coding region accumulates four indels, one precautious STOP codon plus 11 substitutions. This suggests there has been selection for inactivating mutations of IRP30 in *A. cephalotes*.

We tested the fate of *IRP30* in other leafcutter ants belonging to the genus *Atta* and closely related genus *Acromyrmex*. Genomic DNAs of three *Atta* species – *Atta laevigata*, *Atta vollenweideri*, *Atta sexdens* – plus two *Acromyrmex* species, *Acromyrmex lundii* and *Acromyrmex ambiguus* were subjected to PCR amplification using

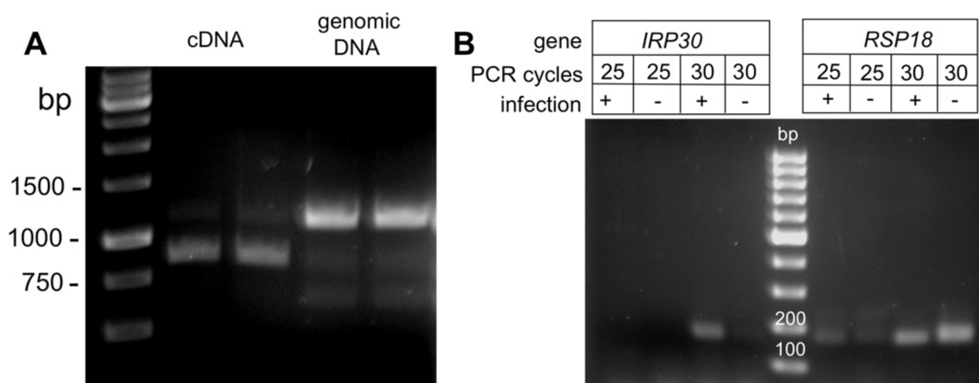

Fig. 6. Defective *IRP30* mRNA of ant *Atta cephalotes* is expressed in response to infection and correctly processed. A. Primers for amplification of defective IRP30 of *A. cephalotes* amplify differently sized products from cDNA and genomic DNA. Both products were sequenced and two frame-shifting mutations confirmed. Size difference is due to intron splicing in the process of RNA maturation. B. *IRP30* mRNA is expressed in infection-dependent manner. Primers were designed, which flank ~180 bp fragments of *rps18* and *IRP30* of *A. cephalotes*. Five major workers were infected with *E. coli* and used for RNA isolation 24 h later together with non-infected controls. Semi-quantitative PCR was run over 25 and 30 cycles with both cDNAs to quantify the expression of *IRP30* and *rps18*. Whereas both samples contain similar amounts of *rps18*, *IRP30* was amplified only from the cDNA of infected ants.

primers derived from *A. cephalotes IRP30* and sequenced. The sequences were similar to *IRP30* of *A. cephalotes*, but none of them contained any of the frame-shift mutations found in *A. cephalotes* (Fig. S7). Thus, in contrast to *A. cephalotes*, *IRP30* genes of these five ant species code for fully functional proteins.

These findings suggest that IRP30 may be of general functional importance across the social Hymenoptera, but that there are both sub-family specific (e.g. the Myrmicinae) and species-specific (*A. cephalotes*) losses of the gene/functionality. It remains to be seen why this has happened in *A. cephalotes* but not in close relatives with similar ecology.

3.4. Evolution of IRP30

3.4.1. Phylogeny and origin of IRP30 proteins

Using database searches and low stringency PCR, we characterized IRP30 from 23 organisms, all belonging to the *Aculeata* suborder of Hymenoptera. One of us (Fujiyuki et al., 2009) identified a putative homolog of IRP30 in a distant non-social hymenopteran, *Nasonia vitripennis* belonging to *Parasitica* grade of *Apocrita* (for recent phylogeny of Hymenoptera, see Davis et al. (2010)).

Importantly, we found no IRP30-like sequences in the genomes of 27 non-hymenopteran arthropod genomes sequenced so far.

Detailed view at aligned IRP30 sequences confirms that critical residues of LRR and LZ domains are conserved in IRP30s of all tested species (Fig. S8).

Aligned IRP30 sequences were subjected to phylogenetic analysis (Fig. 7A). Interestingly, the *N. vitripennis* homolog of IRP30, NvIRP30, seemed phylogenetically very distant. In order to learn more about the evolution of *IRP30* of *Nasonia*, we investigated the genomic organization of the *IRP30* loci in the genomes of different species. The outstanding position of *Nasonia* IRP30 could be confirmed by this analysis. Whereas all other *IRP30* genes possess a single intron of different length interrupting their coding sequences at the position between the codons 39 and 40 of *A. mellifera* IRP30, the *NvIRP30* contains two short introns, which are both located in different positions (Fig. 7B).

4. Discussion

At the outset of our studies IRP30 appeared to be a unique infection responsive protein for *A. mellifera* with no clear homologs in other metazoans. This was particularly interesting since the honeybee genome sequencing identified significantly lower numbers of immune-related genes than in other insects (Evans et al., 2006). For this reason it has been speculated that IRP30 might be a bee-specific immune factor opening the door toward special immunity of this social insect (social immunity).

Here we provide the first evidence that IRP30 is present in a number of species across the three families with social species (Apidae, Formicidae, Vespidae), but found it lacking in the only non-social Hymenoptera genome currently available. We provide a mixture of experimental and bioinformatic analyses that strongly imply that IRP30 is participating in immune response in social Hymenoptera.

4.1. Genetic and biochemical properties of IRP30

IRP30 was initially found together with other immune responsive proteins (PO, PGRP-S2 and CE) by proteomic profiling of adult bee haemolymph upon bacterial challenge (Randolt et al., 2008). We now provide evidence for simultaneous time-dependent expression of *IRP30* and *CE-D* (Fig. 2B) in response to inoculation

Table 2
Observed non-sense mutations in different alleles of *Atta cephalotes IRP30*.

Nucl.position →	508	648	843	846	878
*A.ceph*_Panama*	Δ ATGC	Δ G	–	–	GGG->GGGG
*A.ceph*_Panama	–	Δ G (–/–)	–	–	GGG->GGGG (–/–)
*A.ceph*_Trin&Tob	–	Δ G (–/–)	Δ TACTGATGTCG (–/+)	CGA->TGA (–/+)	GGG->GGGG (–/+)

*A.ceph*_Panama* – sequence from *A. cephalotes* genome sequencing (Suen et al., 2011) (from Panama).
*A.ceph*_Panama – locally cultivated colony (from Panama).
*A.ceph*_T Trin&Tob – colony held at Zoo London (from Trinidad & Tobago).
(–/–) – present in all sequenced alleles from a given colony.
(–/+) – present in some alleles from a given colony.

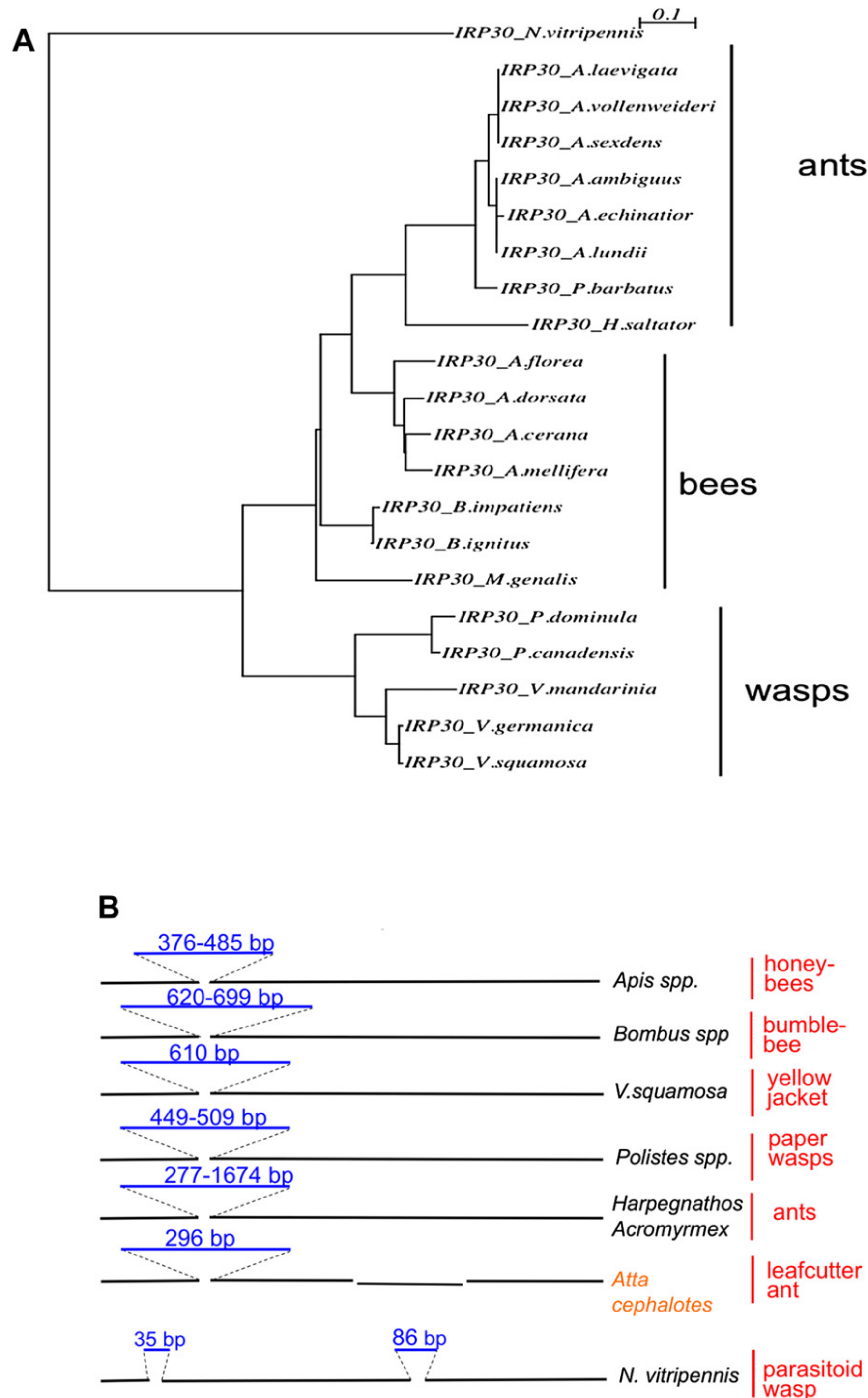

Fig. 7. Phylogeny of Hymenopteran IRP30 proteins. A. Protein sequences of IRP30 from different species were aligned using CLUSTAL-W algorithm. The calculated neighbor-joining tree confirms major clades (wasps, ants, bees) and reveals the large phylogenic distance of *Nasonia IRP30*. All major clades are supported by high bootstrap values. B. Comparison of *IRP30* gene structures. Blue bars above the lines represent the introns. Intron sizes (in nucleotides) are indicated above. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with living bacteria or purified components of their cell walls. We show that after infection the summer honeybees express higher levels of IRP30 in their abdomen compared to other body segments (Fig. 4B). The same is true for winter honeybees (Fig. 4C), but contrastingly the non-infected winter honeybees only expressed *IRP30* in the fat body (Fig. 4D).

IRP30 mRNA was independently found by one of us to be highly induced in bee brains infected with Kakugo virus (Fujiyuki et al., 2009). On the other hand, (Scharlaken et al., 2007) did not identify IRP30 among bee head proteins that are differentially expressed after infection with *E. coli*. Unfortunately, Scharlaken et al. did not use molecular weight markers in their two-dimensional gels making the identification of the IRP30 spot *ex post* impossible.

Immunoreactive IRP30 was identified in the haemolymph of bacterially infected adult bees of all castes and bumblebees. Centrifugation did not pellet it from the bee haemolymph (see

Fig. S2), which indicates that this protein does not stably associate with haemocytes. Thus IRP30 is a soluble extracellular protein.

A leucine-rich repeat domain located between amino acids 71–130 is a common interacting domain found in many proteins. It forms a crescent-shaped structure with two different surfaces. The outer convex surface is usually hydrophilic in its nature, whereas the inner space of a crescent is often the place where hydrophobic interactions with ligands take place (Bella et al., 2008). The second remarkable structural element found in IRP30 is the C-terminal leucine zipper. It folds into α-helix and serves as a structural element conveying interaction with another protein bearing the same structural motif. Two interacting leucine zippers form intertwirled α-helices held together by a hydrophobic stacking of leucines in the interaction interface. Knowing that, we sought for interacting proteins by different approaches ranging from immune precipitation of bee haemolymph (Fig. S2), over Far-Western blotting to adsorption onto living bacteria (Fig. S4). The absence of interacting proteins can be explained in several ways: i) the size of interacting partners. Small peptides would yield only a weak, if any, signal in the Coomassie-stained gel due to their size and amount of bound dye ii) peptides of ~50 and ~25 kDa could be overlaid by the subunits of immunoglobins in immunoprecipitation iii) last possibility would be that the interaction is of temporary nature, not strong enough to resist the experimental procedures used. Antibodies against carboxylesterase did not detect this protein in an immunoprecipitation assay. Thus, carboxylesterase and IRP30, despite identical regulation and presence of leucine zipper in their primary structures, do not associate in bee haemolymph.

Leucine-rich repeat motif-containing proteins were described as an essential component of insect immunity. For example, LRR-proteins (LRIM1, APL1, LRRD7) form a complex, which confers immunity of mosquitoes against Plasmodium ookinetes by mediating their melanization. The same proteins regulate also the endocytosis of bacteria by *Anopheles* haemocytes (Povelones et al., 2009).

Another leucine-rich repeat containing protein, Leureptin, was identified recently in the haemolymph of *Manduca sexta* (Zhu et al., 2010). It shares some similarities with IRP30, leureptin is induced by bacterial infection and it seems to be involved in pathogen recognition as it binds bacterial LPS. However there are hardly further similarities between leureptin and IRP30. With its 13 repeats, the LRR of leureptin is significantly larger, than that of IRP30 (5 repeats). The C-terminus of leureptin does not contain any leucine zipper. Instead, polar and charged amino acids accumulate there (Zhu et al., 2010).

4.2. Immune responsiveness of IRP30

Leureptin is constitutively expressed in *M. sexta* larvae; its amounts in haemolymph decrease continuously after infection as it binds to haemocytes and accumulates in the fat body (Zhu et al., 2010). In contrast, IRP30 was neither constitutively expressed nor could its expression be induced by infection during larval development (Figs. S4, S5). Only in adult bees its expression can be induced by bacterial infection. In agreement with our observations IRP30 was not identified in the haemolymph proteome of bee larvae infected with *Paenibacillus larvae* (Chan et al., 2009). Immunoreactive IRP30 is first detected 12 h p.i., and reaches its maximum 24–48 h p.i. Winter bees constitutively express small amounts of IRP30, which can be boosted by bacterial infection (Fig. 4). Supposing a role of IRP30 in bee immunity, its constitutive expression in winter bees may indicate that immunity of winter bees is constitutively upregulated to ensure survival of the colony under hostile living conditions. It would be interesting to see how colony fitness affects the *IRP30* expression during winter and if IRP30 level correlates with disease resistance and winter survival.

Carboxylesterases have been linked to immune response in different organisms. Human carboxylesterase HSME1 was found in the liver, the human analog to the fat body (Munger et al., 1991) and in monocytes and macrophages (components of the cellular immune system (Zschunke et al., 1992)). In insects carboxylesterases are thought to be part of the haemocytes together with other hydrolytic enzymes like acidic phosphatases and proteases (Lanzrein et al., 1998). Our new findings together with the fact that this carboxylesterase (gi|66512983) is the only one out of 24 in *Apis* with a leucine zipper motif give a support to the idea of carboxylesterase as an immune protein.

The pattern and kinetics of IRP30 and CE-D synthesis do not reconcile with their role in primary immune response, which acts very rapidly and effectively, eliminating >99% of invading bacteria within minutes (Haine et al., 2008). However it overlaps with the immune peptides synthesis (see Fig. 5B), which are major players in the second delayed line of immune defense. But unlike immune peptides, IRP30 does not participate in larval immune response. Does it reflect a weak point of the larval immunity? For example, only larvae are susceptible to *P. larvae*.

4.3. The role of IRP30 in social evolution

Our sequencing of *IRP30* in two populations of *A. cephalotes* provides an excellent example of initial stages of gene inactivation, drifting it on a heterochromatization trajectory. The presence of three frame-shifting mutations in its genomic sequence was unusual since three frame-shifting insertions or deletions (indels) in ~1100 nucleotides is far beyond any acceptable sequencing error rate (less than 1 error in 10,000 bases).

Our sequencing confirmed existence of two indels also in the processed *IRP30* transcript. In addition we identified three additional non-sense mutations in a geographically distant colony from Trinidad & Tobago. Despite our efforts we did not identify any *IRP30* allele in *A. cephalotes* coding for an intact protein (Table 2). Depending on the presence of the first non-sense mutation in a given allele, the truncated *A. cephalotes* IRP30 would contain either no or only truncated LRR. It is highly improbable that such a protein would be functional. Nevertheless, the *A. cephalotes IRP30* mRNA is still induced by bacterial infection and its primary transcript undergoes RNA splicing removing the conserved intron as confirmed by sequencing of both genomic DNA and cDNA (see Fig. 6). Extensive gene loss in *A. cephalotes* has been assigned to its mutualism with a fungus (Suen et al., 2011). It is conceivable that permanent exposition to fungus could lead to inactivation of an immune gene such as *IRP30* in this species, However intact *IRP30* genes in other fungus-growing *Atta* and *Acromyrmex* species controvert this idea. Thus, why only *A. cephalotes* and not its close or distant relatives accumulated these mutations remains a mystery. Accumulation of indel mutations in the coding region of *A. cephalotes IRP30* as compared to non-coding intron indicates that selection for *IRP30* inactivation might have been acting in this species.

The neighbor-joining tree calculated from aligned sequences of different *IRP30*s shows clusters, which reflect subdivision of tested species into major *Aculeata* subfamilies (the sequences within of *Aculeata* subfamilies such as wasps are more similar to each other than to other *IRP30*s). In line with stated above, *N. vitripennis IRP30* is the most distant *IRP30*. SLIT-homolog, a protein consisting of three LRR domains was found by BLAST as a closest non-IRP30 relative (lacking the typical IRP30 size and buildup, LRR plus LZ at the C-terminus). Instead SLIT-homolog contains three separate LRR motifs. When SLIT-homolog was included in the analysis, it appeared to be less distant than the putative IRP30 of *N. vitripennis* (Fig. 7). Another hint comes from the intron structure of *IRP30* genes, where all *IRP30*s have a single intron residing in the same

position. The only exception is *Nasonia IRP30*, which contains two introns, each in a position different from that of other *IRP30*s (Fig. 7, Table 1).

Sociality has evolved at least 11 times in the Hymenoptera (Hughes et al., 2008). *N. vitripennis* is a highly derived parasitic wasp, but is the only non-social hymenoptera for which genome sequence data is available. It is known that positions and translational phases of introns are difficult to change without damaging the encoded protein. Therefore the most parsimonious explanation is that the *IRP30* of *Nasonia* does not share a common origin with other *IRP30* genes, i.e. *N. vitripennis IRP30* is a result of convergent evolution. However, until more data become available for non-parasitic, non-social Hymenoptera, we cannot determine whether the IRP30 and its function as we find in the social Hymenoptera is a socially-derived trait, that has evolved multiple times through convergent evolution in social lineages, or whether it was present in a common ancestor of all Hymenoptera, and subsequently lost/modified in some lineages (e.g. *Nasonia*, *Myrmicine* ants). It remains to be seen why IRP30 and its immune function has been conserved in some lineages and not other. Moreover, it will be interesting to determine whether this gene, if present, also has a role in immunity in solitary Hymenoptera.

5. Conclusions

In conclusion, we present a comprehensive physiological, genomic and phylogenetic assessment of a wide-spread and important immune protein. We characterize the protein and describe its biochemistry in the honeybee, demonstrate its responsiveness to infection, detect homologs in several species of social Hymenoptera, and compare gene structure in these species, and examine the evolution of this gene across insects, and document cases where the gene has been lost or has become non-functional. Finally, we show that IRP30 of *N. vitripennis* possibly does not share the origin with other IRP30s raising the intriguing question of whether IRP30 has evolved a key role and function in social living.

Acknowledgments

This work was supported by a research grant from the Deutsche Forschungsgemeinschaft, SFB 567, project A9.

We wish to thank Dirk Ahrens for help with beekeeping, members of the BEEgroup and Department of Pharmaceutical Biology for stimulating discussions.

We are indebted to Michael Goodishman and Bung-Rae Jin for donation of cDNA libraries and genomic DNA, Rita Cervo for providing the *Polistes dominulus* samples, Drs Masato Oto and Tetsuhiko Sasaki for *Vespa mandarinia*, Jürgen Gadau for sharing unpublished results, Annette Laudahn for providing ants specimens, Bert Hölldobler for valuable information on ants.

Appendix. Supplementary data

Supplementary data related to this article can be found online at doi:10.1016/j.ibmb.2011.09.006.

References

Albertova, V., Su, S., Brockmann, A., Gadau, J., Albert, S., 2005. Organization and potential function of the mrjp3 locus in four honeybee species. J. Agric. Food Chem. 53, 8075–8081.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Bella, J., Hindle, K.L., McEwan, P.A., Lovell, S.C., 2008. The leucine-rich repeat structure. Cell. Mol. Life Sci. 65, 2307–2333.

Bonasio, R., Zhang, G., Ye, C., Mutti, N.S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., Zhang, P., Huang, Z., Berger, S.L., Reinberg, D., Wang, J., Liebig, J., 2010. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Science 329, 1068–1071.

Cerenius, L., Söderhall, K., 2004. The prophenoloxidase-activating system in invertebrates. Immunol. Rev. 198, 116–126.

Chan, Q.W.T., Melathopoulos, A.P., Pernal, S.F., Foster, L.J., 2009. The innate immune and systemic response in honey bees to a bacterial pathogen, *Paenibacillus larvae*. BMC Genomics 10, 387.

Consortium, H.G.S., 2006. Insights into social insects from the genome of the honeybee *Apis mellifera*. Nature 443, 931–949.

Davis, R.B., Baldauf, S.L., Mayhew, P.J., 2010. The origins of species richness in the Hymenoptera: insights from a family-level supertree. BMC Evol. Biol. 10, 109.

Evans, J.D., Aronstein, K., Chen, Y.P., Hetru, C., Imler, J.L., Jiang, H., Kanost, M., Thompson, G.J., Zou, Z., Hultmark, D., 2006. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. Insect Mol. Biol. 15, 645–656.

Fujiyuki, T., Matsuzaka, E., Nakaoka, T., Takeuchi, H., Wakamoto, A., Ohka, S., Sekimizu, K., Nomoto, A., Kubo, T., 2009. Distribution of Kakugo virus and its effects on the gene expression profile in the brain of the worker honeybee *Apis mellifera* L. J. Virol. 83, 11560–11568.

Haine, E.R., Moret, Y., Siva-Jothy, M.T., Rolff, J., 2008. Antimicrobial defense and persistent infection in insects. Science 322, 1257–1259.

Hoffman, E.A., Goodisman, M.A.D., 2007. Gene expression and the evolution of phenotypic diversity in social wasps. BMC Biol. 5, 23.

Hughes, W.O., Oldroyd, B.P., Beekman, M., Ratnieks, F.L., 2008. Ancestral monogamy shows kin selection is key to the evolution of eusociality. Science 320, 1213–1216.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., Gibson, T.J., 1998. Multiple sequence alignment with Clustal X. Trends Biochem. Sci. 23, 403–405.

Kim, Y.-J., Hwang, J.-S., Yoon, H.-J., Yun, E.-Y., Lee, S.B., Ahn, M.-Y., Kim, N.-S., Kim, I., Jeon, J.-P., Hwang, S.-J., 2006. Expressed sequence tag analysis of the diapausing queen of the bumblebee *Bombus ignitus*. Entomol. Res. 36, 191–195.

Klaudiny, J., Albert, S., Bachanova, K., Kopernicky, J., Simuth, J., 2005. Two structurally different defensin genes, one of them encoding a novel defensin isoform, are expressed in honeybee *Apis mellifera*. Insect Biochem. Mol. Biol. 35, 11–22.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.

Lanzrein, B., Pfister-Wilhelm, R., Wyler, T., Trenczek, T., Stettler, P., 1998. Overview of parasitism associated effects on host haemocytes in larval parasitoids and comparison with effects of the egg-larval parasitoid *Chelonus inanitus* on its host *Spodoptera littoralis*. J. Insect Physiol. 44, 817–831.

Mellroth, P., 2002. A scavenger function for a *Drosophila* peptidoglycan recognition protein. J. Biol. Chem. 278, 7059–7064.

Munger, J.S., Shi, G.P., Mark, E.A., Chin, D.T., Gerard, C., Chapman, H.A., 1991. A serine esterase released by human alveolar macrophages is closely related to liver microsomal carboxylesterases. J. Biol. Chem. 266, 18832–18838.

Nagai, K., Thogersen, H.C., 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*. Methods Enzymol. 153, 461–481.

Osta, M.A., Christophides, G.K., Kafatos, F.C., 2004. Effects of mosquito genes on *Plasmodium* development. Science 303, 2030–2032.

Pan, Y., Guo, H., Gao, X., 2009. Carboxylesterase activity, cDNA sequence, and gene expression in malathion susceptible and resistant strains of the cotton aphid, *Aphis gossypii*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 152, 266–270.

Povelones, M., Waterhouse, R.M., Kafatos, F.C., Christophides, G.K., 2009. Leucine-rich repeat protein complex activates mosquito complement in defense against *Plasmodium* parasites. Science 324, 258–261.

Randolt, K., Gimple, O., Geissendorfer, J., Reinders, J., Prusko, C., Mueller, M.J., Albert, S., Tautz, J., Beier, H., 2008. Immune-related proteins induced in the hemolymph after aseptic and septic injury differ in honey bee worker larvae and adults. Arch. Insect Biochem. Physiol. 69, 155–167.

Schagger, H., von Jagow, G., 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 368–379.

Scharlaken, B., de Graaf, D.C., Goossens, K., Brunain, M., Peelman, L.J., Jacobs, F.J., 2008. Reference gene selection for insect expression studies using quantitative real-time PCR: the honeybee, *Apis mellifera*, head after a bacterial challenge. J. Insect Sci. 8, 1–10.

Scharlaken, B., De Graaf, D.C., Memmi, S., Devreese, B., Van Beeumen, J., Jacobs, F.J., 2007. Differential protein expression in the honey bee head after a bacterial challenge. Arch. Insect Biochem. Physiol. 65, 223–237.

Schluns, H., Crozier, R.H., 2007. Relish regulates expression of antimicrobial peptide genes in the honeybee, *Apis mellifera*, shown by RNA interference. Insect Mol. Biol. 16, 753–759.

Shiotsuki, T., Kato, Y., 1999. Induction of carboxylesterase isozymes in *Bombyx mori* by *E. coli* infection. Insect Biochem. Mol. Biol. 29, 731–736.

Smith, C.R., Smith, C.D., Robertson, H.M., Helmkampf, M., Zimin, A., Yandell, M., Holt, C., Hu, H., Abouheif, E., Benton, R., Cash, E., Croset, V., Currie, C.R., Elhaik, E., Elsik, C.G., Fave, M.J., Fernandes, V., Gibson, J.D., Graur, D., Gronenberg, W., Grubbs, K.J., Hagen, D.E., Viniegra, A.S.I., Johnson, B.R., Johnson, R.M., Khila, A., Kim, J.W., Mathis, K.A., Munoz-Torres, M.C., Murphy, M.C., Mustard, J.A., Nakamura, R., Niehuis, O., Nigam, S., Overson, R.P., Placek, J.E., Rajakumar, R., Reese, J.T., Suen, G., Tao, S., Torres, C.W., Tsutsui, N.D., Viljakainen, L.,

Wolschin, F., Gadau, J., 2011. Draft genome of the red harvester ant *Pogonomyrmex barbatus*. Proc. Natl. Acad. Sci. U. S. A. 108, 5667–5672.

Steiner, H., 2004. Peptidoglycan recognition proteins: on and off switches for innate immunity. Immunol. Rev. 198, 83–96.

Suen, G., Teiling, C., Li, L., Holt, C., Abouheif, E., Bornberg-Bauer, E., Bouffard, P., Caldera, E.J., Cash, E., Cavanaugh, A., Denas, O., Elhaik, E., Favé, M.-J., Gadau, J., Gibson, J.D., Graur, D., Grubbs, K.J., Hagen, D.E., Harkins, T.T., Helmkampf, M., Hu, H., Johnson, B.R., Kim, J., Marsh, S.E., Moeller, J.A., Muñoz-Torres, M.C., Murphy, M.C., Naughton, M.C., Nigam, S., Overson, R., Rajakumar, R., Reese, J.T., Scott, J.J., Smith, C.R., Tao, S., Tsutsui, N.D., Viljakainen, L., Wissler, L., Yandell, M.D., Zimmer, F., Taylor, J., Slater, S.C., Clifton, S.W., Warren, W.C., Elsik, C.G., Smith, C.D., Weinstock, G.M., Gerardo, N.M., Currie, C.R., 2011. The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. PLoS Genet. 7, e1002007.

Tautz, J., 2008. The Buzz about Bees. Biology of a Superorganism. Springer, Heidelberg.

Wilson-Rich, N., Dres, S.T., Starks, P.T., 2008. The ontogeny of immunity: development of innate immune strength in the honey bee (*Apis mellifera*). J. Insect Physiol. 54, 1392–1399.

Woodard, S.H., Fischman, B.J., Venkat, A., Hudson, M.E., Varala, K., Cameron, S.A., Clark, A.G., Robinson, G.E., 2011. Genes involved in convergent evolution of eusociality in bees. Proc. Natl. Acad. Sci. U. S. A. 108, 7472–7477.

Zhu, Y., Ragan, E.J., Kanost, M.R., 2010. Leureptin: a soluble, extracellular leucine-rich repeat protein from *Manduca sexta* that binds lipopolysaccharide. Insect Biochem. Mol. Biol. 40, 713–722.

Zschunke, F., Salmassi, A., Kreipe, H., Parwaresch, M.R., Radzun, H.J., 1992. Heterogenous expression and putative structure of human monocyte/macrophage serine esterase 1. Res. Immunol. 143, 125–128.