Compatibility in the *Biomphalaria glabrata/Echinostoma caproni* model: new candidate genes evidenced by a suppressive subtractive hybridization approach

A. BOUCHUT[1], C. COUSTAU[2], B. GOURBAL[1] *and* G. MITTA[1]*

[1] *Parasitologie Fonctionnelle et Evolutive, UMR 5555, CNRS Université de Perpignan Perpignan, France*
[2] *U547, INSERM Institut Pasteur de Lille, Lille, France*

SUMMARY

In order to elucidate mechanisms underlying snail/echinostome compatibility, numerous molecular studies comparing transcripts and proteins of *Biomphalaria glabrata* susceptible or resistant to *Echinostoma caproni* were undertaken. These studies focused on plasma and haemocytes of the two strains and revealed that some transcripts and/or proteins were differentially expressed between strains. The aim of the present study was to develop a complementary transcriptomic approach by constructing subtractive libraries. This work revealed some candidate transcripts already identified in previous studies (calcium-binding proteins and glycolytic enzymes) as well as novel candidate transcripts that were differentially represented between strains of *B. glabrata*. Among these newly identified genes, we revealed several genes potentially involved in immune processes encoding proteases, protease inhibitors, a lectin, an aplysianin-like molecule, and cell adhesion molecules.

Key words: *Biomphalaria glabrata*, *Echinostoma caproni*, host-parasite interactions, suppression subtractive hybridization, compatibility.

INTRODUCTION

Snail-trematode immunobiological interactions have been studied in the fresh-water snail *Biomphalaria glabrata* following infection by trematodes like the human blood fluke *Schistosoma mansoni* or echinostome species (*Echinostoma paraensei* or *Echinostoma caproni*). As compatibility polymorphism was evidenced in these different models (Richards, 1975; Richards and Shade, 1987; Lewis *et al.* 1993; Langand and Morand, 1998), numerous molecular studies comparing susceptible and resistant strains of molluscs were undertaken to elucidate the mechanisms underlying differences in compatibility. These studies revealed many transcripts or proteins that were differentially expressed between strains of *B. glabrata* susceptible or resistant to *S. mansoni* or *E. caproni* (Miller *et al.* 2001; Raghavan *et al.* 2003; Lockyer *et al.* 2004; Nowak *et al.* 2004; Vergote *et al.* 2005; Jung *et al.* 2005; Bouchut *et al.* 2006 *a, b*).

In the *B. glabrata/E. caproni* system we are investigating, 2 snail strains have previously been selected for resistance or susceptibility to this parasite (Langand *et al.* 1998). *E. caproni* miracidia

are able to penetrate both susceptible and resistant snails. While the parasite undergoes normal development in susceptible snails, it is encapsulated and eliminated in resistant snails. Susceptibility/resistance of *B. glabrata* to *E. caproni* is inherited in a multigenic fashion (Langand and Morand, 1998) and has been shown to rely on both cellular and humoral factors (Ataev and Coustau, 1999). Regarding humoral factors, calcium-binding proteins and glycolytic enzymes differentially represented in the plasma of susceptible and resistant snails were identified using a comparative proteomic approach (Vergote *et al.* 2005). This study also showed that the genes encoding these plasmatic proteins are mainly expressed in the albumen gland, raising the question of the potential involvement of this organ in mechanisms underlying susceptibility/resistance to *E. caproni*. Regarding cellular factors, previous studies showed that excretory-secretory (ES) products from *in vitro* transformed *E. caproni* sporocysts inhibited key defence functions of susceptible host haemocytes such as adhesion and phagocytosis (Lie, 1982; Humbert and Coustau, 2001). Interestingly, haemocytes from resistant snails remained unaffected by these parasite ES products, suggesting that they exhibit constitutive differences with susceptible snails haemocytes (Humbert and Coustau, 2001). In order to investigate the molecular basis of these differences, we developed comparative molecular approaches on haemocytes of the two strains. A global

* Corresponding author: Parasitologie Fonctionnelle et Evolutive, UMR 5555, CNRS Université de Perpignan, 52 Ave Paul Alduy, 66 860 Perpignan cedex, France. Tel: +33 (0) 4 68 66 21 85. Fax: +33 (0) 4 68 66 22 81. E-mail: mitta@univ-perp.fr

proteomic approach on this cell type allowed the identification of several proteins that were differentially represented in haemocytes from the susceptible and resistant snails (Bouchut *et al.* 2006*b*). Another study, focusing on genes involved in adhesion processes confirmed the potential importance of several candidate genes like dermatopontins and matrilin-like molecules (Bouchut *et al.* 2006*a*) This transcriptomic approach also revealed several genes that were differentially expressed in whole bodies of *E. caproni*-susceptible and -resistant *B. glabrata*.

These different approaches taken together allowed identification of several candidate genes and suggested a potential involvement of tissues other than haemocytes in susceptibility/resistance processes. In order to investigate this latter point, we compared in the present work the transcripts of the two snail strains at the whole body level using a Suppression Subtractive Hybridization (SSH) approach. Herein, we describe 3 main categories of ESTs: those directly related to immune function, ESTs related to calcium homeostasis and ESTs encoding glycolytic enzymes. Quantitative analyses of transcript levels confirmed their differential expression and the validity of this SSH approach.

MATERIALS AND METHODS

Biomphalaria glabrata *strains*

Biomphalaria glabrata snails used in this study belonged to the EAF-S or CB-R strains selected, respectively, for their susceptibility or resistance to *E. caproni* (Langand *et al.* 1998; Humbert and Coustau, 2001). The selection was performed from the polymorphic Bg.Bra strain (Langand *et al.* 1998). At the time of the study, the percentages of adult snails susceptible to *E. caproni* were: Bg.Bra = 70%, EAF-S = 98% and CB-R = 1%.

RNA isolation and substractive cDNA library construction

Total RNA from whole body of unexposed (to *E. caproni*) EAF-S and CB-R snails was extracted according to previously described procedures (Guillou *et al.* 2004). The total RNAs from 10 individuals were pooled for each strain. For subtractive library construction, poly(A)$^+$ RNA was purified using NucleoTrap mRNA purification kit (Clontech). Two forward and reverse libraries were constructed by subtracting mRNA from EAF-S and CB-R strains (tester) with mRNA from CB-R and EAF-S strains (driver), respectively. SSH libraries were produced using the PCR-Select cDNA subtraction kit (Clontech). Tester and driver cDNA were prepared using 2 μg of poly(A)$^+$ RNA. Enzyme digestion, adapter ligation, hybridization, and PCR

amplification were performed according to protocols provided by the manufacturer (Clontech). PCR products were cloned into pCR4-TOPO cloning vector using TOPO TA cloning kit (Invitrogen) and transformed into One Shot TOP10 chemically competent *Escherichia coli* cells (Invitrogen). The 2 libraries obtained were named EAF-S and CB-R libraries and contained transcripts potentially over-represented in the 2 corresponding strains respectively.

A total of 256 EAF-S library clones and 257 CB-R library clones were randomly selected and single-pass sequenced using a dideoxy-dye-terminator method (CEQ^{TM} DTCS-Quick Start kit, Beckman Coulter) and a CEQ^{TM} 8000 apparatus (Beckman Coulter). Vector and adaptor sequences were trimmed from all sequences using Sequencer^{TM} software (Gene Codes Corporation). All sequences were then examined for possible sequencing errors. High quality ESTs, longer than 150 bp, in length, were assembled in clusters (Sequencher^{TM} software). Consensus sequences from clusters or unique sequences from singletons were submitted to database searches using BLASTX and BLASTN programs (http://blast.genome.jp/ or http://www.ncbi.nlm.nih.gov/BLAST/). Specific domain searches were performed using the RPS-BLAST program (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). EST sequences have been submitted to the dbEST databases (GenBank Accession nos. from EE049524 to EE049639 for the EAF-S library and from EE049640 to EE049693 for the CB-R library).

Real-Time PCR (Q-PCR)

Considering their similarities and their putative interest in our system, genes encoding proteins like CaBP, carbohydrate hydrolytic enzyme genes belonging to immunity class, cytoskeleton structure, cell adhesion, transport, cholesterol homeostasis and transcription regulation were selected for quantitative analysis of their transcripts. Real-Time RT-PCR analysis was performed on total RNAs extracted from EAF-S and CB-R snails. Reverse transcription was performed according to previously described procedures (Guillou *et al.* 2004). Specific primers for Q-PCR were edited using the Light Cycler Probe Design Software version 1.0 (Roche Diagnostics) (Table 1). PCR reactions were set up according to the LightCycler Manual (Roche Diagnostics). The following LightCycler run protocol was used: denaturation programme (95 °C, 10 min), amplification and quantification programmes repeated 40 times (95 °C for 15 s, 60 °C for 5 s, 72 °C for 16 s), melting curve programme (60–95 °C with a heating rate of 0·1 °C per second and continuous fluorescence measurement), and a cooling step to 40 °C. Amplification of single highly specific products was verified (melting curve analysis). For each reaction, the

Table 1. *Primer sequences used for Q-PCR*

Cluster name	Forward	Reverse
1 EAF-S	ATCGAGATGATGACAGGC	CTGTAGTCGTTATTGGTAGTG
2 EAF-S	CATCCTTACATCTGGGCA	TCAATAGGCGAACTTAAAGTC
3 EAF-S	AGAGAACGCCTATGCT	CCCGGTGGTTGAGAAA
5 EAF-S	GAAGAATTTGGCGGGAG	ATGGTCCAGTCATCGT
7 EAF-S	ACATAGAGCTGTTTCAAGACT	CGTGATGAAGCTCGGC
8 EAF-S	AGCATTTAGCACTAAGCAG	CGAAAGTTTTCCTTGGAAGAG
9 EAF-S	CCTGCCTCACCCTTAC	GGATAAATGCCGTCTGG
12 EAF-S	ATTCGCTGTTCCTGAGA	GGTTTCATTGTTGCGTCC
13 EAF-S	ACACAGACTTCAGGAGC	CAAGTGATCTGCCAAGGA
14 EAF-S	TAGACACCCATCACGC	GAATGCCAGTCAGATCGTA
15 EAF-S	TTAGTGGACAAGGATCTCAAC	AGAATATGGTCGCCGT
16 EAF-S	GACGGTGATGTGACCA	AGCTTCCCAGTTGTATCT
17 EAF-S	AAGGTGTGATGCCAGA	CCCAAGAGCCGTGTTC
18 EAF-S	CACCAGTGTGTGGTTAT	CCCCTATTCACAGAGGTT
19 EAF-S	TAGGGCCAACATGCAAG	GTAGACGTATAGGCAGACT
21 EAF-S	GCACTCAGCAGGAGACTAT	GACCATTCCAATTGTAATTTTTG
23 EAF-S	CTGACCAGGTATGTCGG	GTCATCCAGCTCTAGGTT
24 EAF-S	ACAGAGGCTCATGTGG	CATAGCCGTGAGAGTTG
28 EAF-S	CACTATCCTGGGCAAC	ATGCAAACCAGCCAGA
29 EAF-S	GTCCTACTACTACGGTGATAAG	GTGAGGCTGTTGAAGTC
30 EAF-S	ATGACCTCTCCAAATGTG	GGGTGAAGGGTAGGAC
31 EAF-S	ACATTTGAAGGACGCTG	TTGGGAGGTAAACAAGGAA
32 EAF-S	CAATTTTCACGAATACATCTTGC	CTCTGTGAACACGTTGC
42 EAF-S	TGGTGTCCGTCAGTCT	GGTGACCGTCTATGCT
43 EAF-S	CATGGAGAGAGGGCTG	CTTGAACTGATCCGCTATT
1 CB-R	AATCGGGAGAATACGTGA	CAATGATGGTGTCCAGTAAT
2 CB-R	ACAAGAGAAGTTGATAACGC	CGTGATCTAATGTTTCTGGC
3 CB-R	ATATCCTCCCTCAAGTCAC	TGAATGACGAGCAGAGTTA
4 CB-R	CATCGTTGACCATCCTTTC	ACATCATTTCATCATGTCAAGA
5 CB-R	TTACATCGCTGGCACC	GCACTGTCAGCTTATCAC
6 CB-R	TTGCTATGACACCGGC	TTAGGATGAAATGGAGGACT
10 CB-R	ATGGAAGACGCCATTG	ACCGAAGCAGTTGAAGA

crossing point (CP) was determined using the 'Fit Point Method' of the LightCycler Software 3.3 (Roche Diagnostics). Three independent PCR reactions were performed in duplicate and the mean values of the CPs were calculated. For each mRNA to be analysed, the absence of contaminating genomic DNA was verified by running a no-RT control using primers for the ribosomal protein S19 (S19, Accession number CK988928).

Amplification efficiencies (E) of each PCR product were determined according to previously described procedures (Guillou *et al.* 2004). For each sample, the level of transcription of the target gene (Tg) was compared to the transcription of the ribosomal protein S19. The transcription ratio (R) was calculated according to the formula:

$$R = (E_{Tg})^{CPTg} / (E_{S19})^{CPS19}.$$

Statistical significance was assessed using the Mann-Whitney U-test for non-paired series and the Wilcoxon test for paired series. *P* values <0·05 were considered as significant. All tests were performed using STATVIEW v. 4.5 statistic software package.

RESULTS

EST sequencing and general characteristics of EAF-S and CB-R libraries

In total 256 EAF-S library clones and 257 CB-R library clones were sequenced. Using the CEQ™ 8000 sequence analysis software, 227 and 235 cDNA high quality sequences were obtained as one-path reads from the 2 libraries respectively (Table 2). ESTs were aligned and assembled in clusters using the Sequencher™ software. For the EAF-S library, ESTs coalesced into 41 contigs and 84 singletons, suggesting that the overall redundancy of the library was 63·4%. For the CB-R library, ESTs coalesced into 20 contigs and 47 singletons for an overall redundancy of 80% (Table 2).

All clusters, consensus sequences from contigs or unique sequences from singletons, were submitted to database searches using BLASTX and BLASTN programs (http://blast.genome.jp/ or http://www. ncbi.nlm.nih.gov/BLAST/). Clusters corresponding to contaminant sequences – vector, repeated DNA, rRNA, xenocontaminants (sequences from foreign genome) – were removed from the analysis. In order to minimize redundancy in the EST database,

Table 2. General characteristics of *Biomphalaria glabrata* EAF-S and CB-R SSH libraries

	EAF-S Lib	CB-R Lib
Total number of cDNA sequenced	256	257
Total number of cDNA analysed	227	235
Average insert size[a]	650 bp	560 bp
Average EST length[b]	370 bp	450 bp
EST contigs	41	20
Singletons	84	47
Redundancy[c]	63·4%	80%

[a] The average insert size was calculated from 36 randomly selected cDNA clones after insert digestion.
[b] The average EST length was calculated from 100 randomly selected cDNA ESTs.
[c] Redundancy = number of ESTs assembled in clusters/total EST.

sequences displaying 100% identity were submitted as a single sequence. ESTs aligning in the same contig but displaying differences in their nucleotidic sequence were submitted individually to the database. A total of 116 and 54 ESTs were submitted to the dbEST section of the NCBI/GenBank database for the EAF-S and CB-R libraries, respectively (Accession numbers from EE049524 to EE049639 and from EE049640 to EE049693).

Exclusion of false-positive and functional groups of ESTs from the two libraries

As a first approach to exclude false-positives, previously selected clusters of the 2 libraries were aligned and compared using the Sequencher™ software. All sequences common to the 2 libraries were eliminated from the analysis.

Selected ESTs were subjected to BLASTN and BLASTX search. Sequence similarities were considered to be significant when the expected value was less than 10^{-2}. Comparison of the ESTs against non-redundant protein and nucleic databases revealed that 38·8% of EAF-S and 26·3% of CB-R clusters did not align significantly with known genes, 16·5% of EAF-S and 29% of CB-R significantly aligned with genes of unknown function, and the remaining 44·7% of clusters (for both strains) were significantly similar to genes of known function (Fig. 1). All clusters corresponding to this last category were assigned functions as predicted from sequence similarity and subsequently classified into distinct functional classes. ESTs from the EAF-S library were assigned to 6 major classes (Fig. 2A): (1) immunity, (2) calcium homeostasis, (3) cell structure, shape and mobility, (4) energy metabolism and hydrolytic enzymes, (5) protein biosynthesis, (6) nucleic acid biosynthesis and metabolism. ESTs
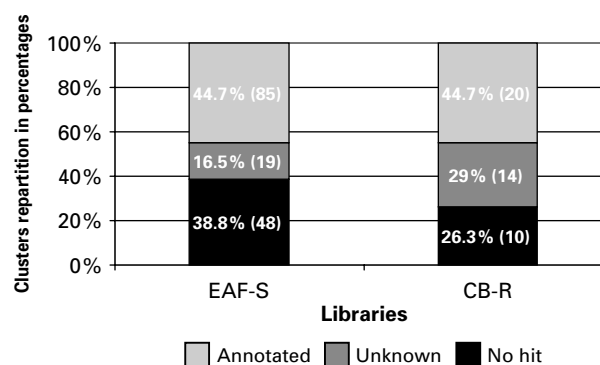

Fig. 1. Cluster distribution of the EAF-S and CB-R libraries after BLASTN and BLASTX analysis. 'Annotated' clusters are clusters displaying similarities for genes of known function. 'Unknown' clusters display similarities for sequences present in the dbEST database of Genbank showing no similarities for genes of known function. The last group 'No hit' corresponds to clusters displaying no homology for any sequences present in any non-redundant databases. The percentage represents the number of clusters for each category. The number given in parentheses is the total number of sequences for each category.

from the CB-R library were grouped into 4 classes (Fig. 2B): (1) immunity, (2) energy metabolism and hydrolytic enzymes, (3) protein biosynthesis and (4) nucleic acid biosynthesis and metabolism.

ESTs of interest and selection of ESTs for which transcripts are differentially represented between strains

Among the genes identified, several functional groups are of particular interest. Some of these genes encoding proteins like CaBP or carbohydrate hydrolytic enzymes (Table 3) were identified in a previous proteomic study (Vergote *et al.* 2005), therefore supporting the potential importance of their corresponding functional groups. Indeed, the same genes (cluster numbers 21 and 22, Table 3) and novel genes of the same families (clusters 14, 15, 16 and 23, 24, 25, 26, Table 3) were identified in the present transcriptomic study. In addition, we identified here novel genes belonging to other functional groups of potential interest in our system (genes belonging to immunity class, cytoskeleton structure, cell adhesion, transport, cholesterol homeostasis and transcription regulation). All the genes corresponding to these different functional groups were selected for further analysis. Genes involved in protein and nucleic acid biosynthesis were arbitrarily excluded from the analyses.

In order to assess the relative expression of selected genes and to confirm their differential expression between strains, quantitative RT-PCR experiments were carried out on total RNA extracted from EAF-S and CB-R snails (Figs 3 and 4). The transcripts

A

Nucleic acid
biosynthesis and
metabolism (9.3%)

Protein biosynthesis
(14%)

Immunity
(30.2%)

Calcium
homeostasis
(7%)

Cell structure,
shape and mobility
(9.3%)

Energy, metabolism
and hydrolytic
enzymes (30.2%)

EAF-S

B

Nucleic acid
biosynthesis and
metabolism (11.8%)

Immunity
(35.3%)

Protein
biosynthesis
(29.4%)

Energy, metabolism
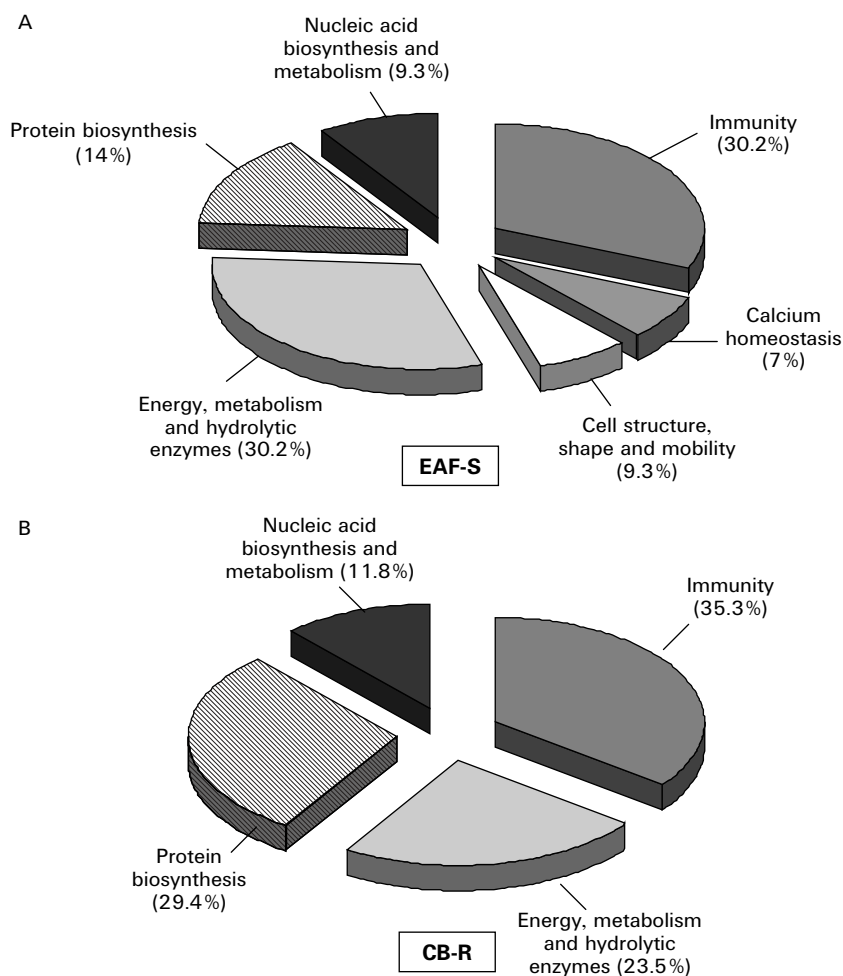and hydrolytic
enzymes (23.5%)

CB-R

Fig. 2. Distribution of the annotated clusters from the EAF-S (A) and CB-R (B) libraries into functional classes. The percentages of clusters for each class are given in parentheses.

corresponding to cluster numbers 4, 6, 10, 11, 20, 22, 25–27, 33, 40 and 41 of the EAF-S library were not detected, nor the cluster numbers 7–9 of the CB-R library. These results show that the SSH approach was successful. Indeed, 80% of the transcripts corresponding to EAF-S clusters display a higher concentration in EAF-S strain as compared to the CB-R strain and the relative expression ratio is greater than 2·5-fold for 60% of them (Fig. 3). For the CB-R library, 100% of the clusters are over-expressed in the CB-R strain and 70% of them displayed a relative expression ratio greater than 2·5 fold (Fig. 4). All these differences are statistically significant ($P < 0.05$).

The EAF-S clusters displaying the more significant differences (transcript concentration 2·5-fold higher in EAF-S snail, Fig. 3) correspond to (i) proteases (cluster 3 and 5) and cell adhesion molecules (cluster 9 and 12), both included in the immunity class, (ii) CaBPs (cluster 14–16), (iii) carbohydrate hydrolytic enzymes like mannanases (cluster 21 and 23) and a glucanase (cluster 24), (iv) transport proteins (cluster 28 and 29) and (v) proteins involved in cholesterol homeostasis (cluster 30–32).

Regarding the CB-R library, the clusters displaying the most significant differences (transcript concentration 2·5 fold higher in CB-R snail, Fig. 4) correspond to putative proteins all belonging to the immunity class: an antimicrobial protein (Bg Aplysianin-A like, cluster 1), a protease (Bg coagulation factor V like, cluster 2), protease inhibitors (Bg serpin like 3 and Bg serpin like 3, cluster 3 and 4, respectively) and a pattern recognition receptor (Bg C-type lectin like, cluster 5).

In order to further investigate the putative function of these different clusters of interest, similarities were analysed in conjunction with conserved deduced domains in the corresponding proteins (Tables 3 and 4 for EAF-S and CB-R, respectively). For example, cluster 3 (Table 3) displays similarities for the cathepsin L of the sea urchin *Strongylocentrotus purpuratus* and the domain corresponding to this class of protease (peptidase C1A subfamily) was detected in the deduced amino acid sequence of this cluster. Consequently, the cluster 3 was named Bg cathepsin L-like. Another example of interest, cluster 5 (Table 4) displays similarities for the lectin 1 of the planaria *Girardia tigrina* (Shagin *et al*. 2002).
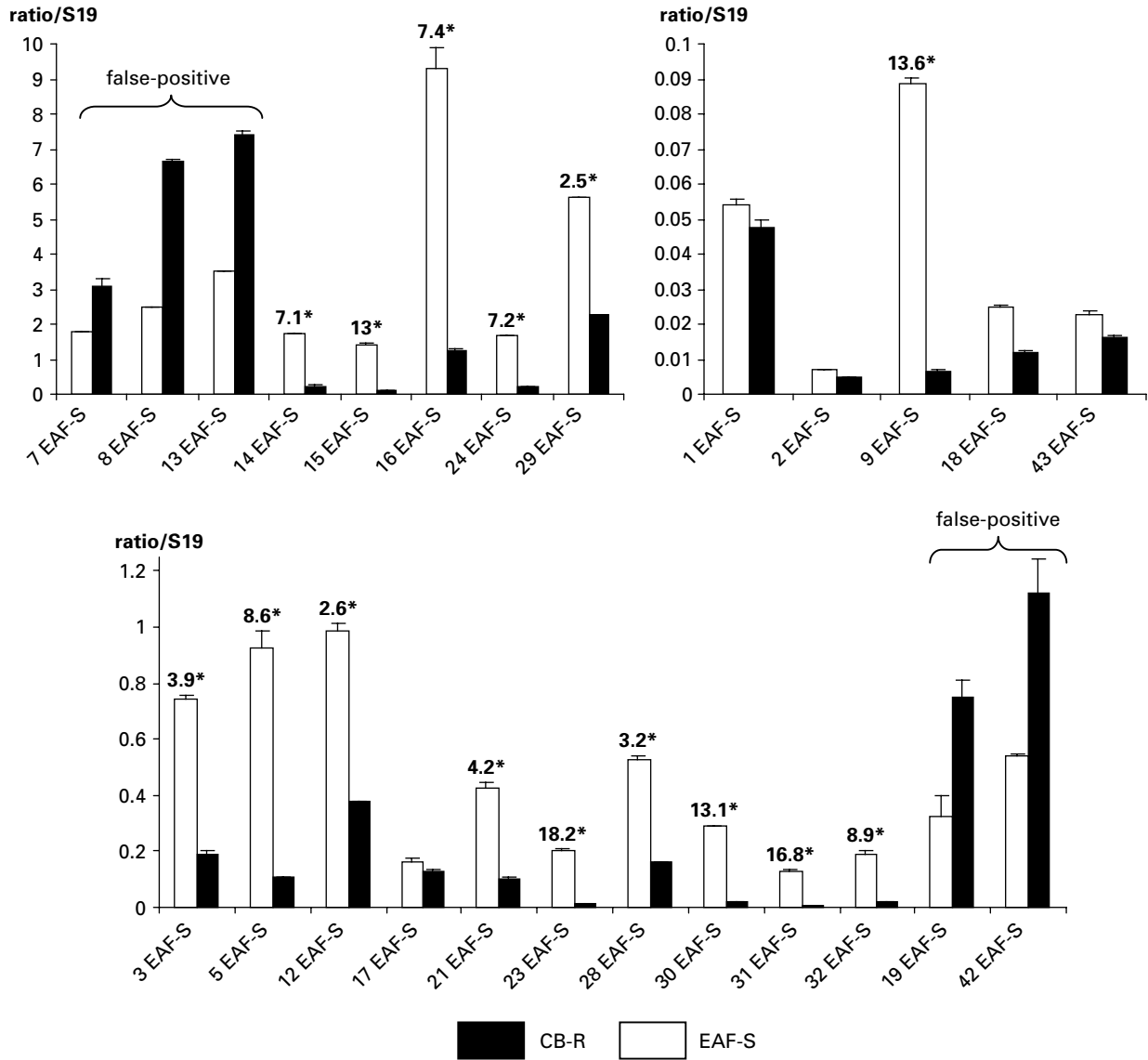
Table 3. Genes identified from the EAF-S library

Class	Functional group	Cluster #	Accession no. in dbEST	Name	Annotated BLASTX similarity (accession #, E-value)[a]	RPS-BLAST detected domain (E-value)[b]
Immunity	Cellular defence effector	1	EE049557	Bg g-type lysozyme like	g-type lysozyme (AAX09979, A. irradians, 4e-20)	GEWL, Goose Egg White Lysozyme domain, gnl|CDD|29563 (6e-22)
	Protease	2	EE049531	Bg serine proteinase like	serine proteinase (CAA89967, *A. gambiae*, 6e-13)	Trypsin-like serine protease, gnl|CDD|29152 (1e-18)
		3	EE049537	Bg cathepsin L like	similar to Cathepsin L precursor (XP_780713, *S. purpuratus*, 9e-63)	Peptidase C1A subfamily, gnl|CDD|30292 (7e-29)
		4	EE049626	Bg serine protease alpha like	serine protease alpha (AAG40233, *B. glabrata*, 5e-16)	ND
		5	EE049544	Bg serine protease beta like 1	serine protease beta (AAG40234, *B. glabrata*, 6e-48)	Trypsin-like serine protease, gnl|CDD|29152 (6e-13)
		6	EE049545	Bg serine protease beta like 2	serine protease beta (AAG40234, *B. glabrata*, 2e-30)	Trypsin-like serine protease, gnl|CDD|29152 (5e-20)
	Protease inhibitor	7	EE049561	Bg serpin like 1	serine proteinase inhibitor (AAM05994, *M. acetivorans*, 5e-08)	SERPIN, SERine Proteinase INhibitors (serpins), gnl|CDD|29117 (3e-13)
		8	EE049549	Bg serpin like 2	probable serine protease inhibitor (NP_924916, *G. violaceus* PCC7421, 4e-08)	SERPIN, SERine Proteinase INhibitors (serpins), gnl|CDD|29117 (1e-13)
	Cell adhesion	9	EE049526	Bg DEC-1 like	DEC-1 (BAD16597, *L. stagnalis*, 3e-13)	Chitin-binding domain type 2, gnl|CDD|24293 (3e-05)
		10	EE049566	Bg dermatopontin like 1	Dermatopontin 2 (BAD97854, *B. glabrata*, 2e-12)	ND
		11	EE049554	Bg salivary proline rich protein like	salivary proline-rich protein precursor (AAA60189, *H. sapiens*, 2e-07)	ND
		12	EE049565	Bg ependymin like protein	ependymin related protein 2 (NP_001007626, *R. norvegicus*, 2e-05)	ND
	Superoxide metabolism	13	EE049528	Bg Cu/Zn SOD like	Superoxide dismutase (XP_996209, *M. Musculus*, 3e-24)	Copper/zinc superoxide dismutase (SODC), gnl|CDD|40181 (4e-32)
Calcium homeostasis	Ca^{2+} binding protein	14	EE049534	Bg calcium binding protein 3	calcium binding protein 1 (AV91525, *B. glabrata*, 1e-16)	ND
		15	EE049535	Bg calcium binding protein 4	calcium binding protein 2 (AV91522, *B. glabrata*, 1e-03)	ND
		16	EE049539	Bg DEC-3 like	DEC-3 protein (BAD16599, *L. stagnalis*, 7e-36)	EF-hand, calcium binding motif, gnl|CDD|28933 (4e-04)
Cell structure, shape and mobility	Cytoskeleton	17	EE049533	Bg actin like	cytoskeletal actin 2 (AAC28359, *M. occulta*, 4e-75)	Actin, gnl|CDD|28896 (1e-61)
		18	EE049555	Bg actin-interacting protein 1 like	actin-interacting protein 1 (Q11176, *C. elegans*, 1e-11)	WD40 domain, gnl|CDD|29257 (4e-04)
		19	EE049548	Bg alpha 2-tubulin like	alpha 2-tubulin (AAS55708, *L. striatellus*, 3e-50)	alpha_tubulin, gnl|CDD|30217 (2e-55)
	Cell adhesion	20	EE049553	Bg collagen like	COL11A2 (BAD93256, *O. latipes*, 6e-05)	ND
Energy, metabolism and hydrolytic enzymes	Carbohydrate metabolism	21	EE049529	Bg endo-1,4-beta-mannanase 1	endo-1,4-beta-mannanase 1 (AAV91523, *B. glabrata*, 7e-70)	ND
		22	EE049530	Bg endo-1,4-beta-mannanase 1	endo-1,4-beta-mannanase 1 (AAV91523, *B. glabrata*, 1e-61)	ND

		23	EE049558	Bg endo-1,4-beta-mannanase 3	endo-1,4-beta-mannanase 2 (AAV91524, *B. glabrata*, 5e-28)	ND
		24	EE049550	Bg endo-1,4-beta-D-glucanase like 1	endo-1,4-beta-D-glucanase (CAC59695, *M. edulis*, 1e-43)	ND
		25	EE049552	Bg endo-beta-1,4-glucanase like	putative endo-beta-1,4-glucanase (BAD12006, *N. koshunensis*, 1e-09)	Glycosyl hydrolase family 9, gnl|CDD|40839 (1e-09)
		26	EE049551	Bg cellulase like	cellulase (BAC67186, *H. discus*, 1e-08)	ND
	Transport	27	EE049547	Bg hemoglobin like	hemoglobin (CAI26300, *B. glabrata*, 5e-06)	ND
		28	EE049546	Bg hemoglobin	hemoglobin (CAH23232, *B. glabrata*, 8e-45)	Globins gnl|CDD|29979 (3e-10)
		29	EE049542	Bg Fatty acid binding protein like	Fatty acid-binding protein, liver (P81653, *H. bivius*, 9e-06)	ND
	Cholesterol homeostasis	30	EE049540	Bg phosphoribosyl-formylglycinamidine synthase like	similar to phosphoribosyl-formylglycinamidine synthase (XP_966891, *T. castaneum*, 8e-19)	Low Density Lipoprotein Receptor Class A domain, gnl|CDD|29012 (2e-05)
		31	EE049541	Bg Low-density lipo-protein receptor-related protein 1B precursor like	similar to Low-density lipoprotein receptor-related protein 1B precursor (XP_422146, *G. gallus*, 9e-23)	Low Density Lipoprotein Receptor Class A domain, gnl|CDD|29012 (9e-06)
		32	EE049559	Bg proprotein convertase subtilisin/kexin type 9 like	proprotein convertase subtilisin/kexin type 9, (Q80W65, AAH38085, *M. musculus*, 1e-03)	ND
	Mitochondrial protein	33	EE049527	Bg mitochondrial protein 18 kDa like	mitochondrial protein 18 kDa isoform a (XP_781554, *S. purpuratus*, 1e-03)	ND
Protein biosynthesis	Ribosomal protein	34	EE049556	Bg ribosomal protein L21 like	ribosomal protein L21 (AAN05604, *A. irradians*, 3e-53)	Ribosomal protein L21e, gnl|CDD|1707 (3e-26)
		35	EE049595	Bg 40S ribosomal pro-tein S2 like	similar to ribosomal protein S2, partial (XP_795432, *S. Purpuratus*, 8e-51)	Ribosomal protein S5, gnl|CDD|9973 (2e-23)
		36	EE049584	Bg ribosomal protein L7 like	Ribosomal protein L7a (AAH59533, *D. rerio*, 1e-57)	ND
		37	EE049617	Bg ribosomal protein S2 like	ribosomal protein S20 (AAM28852, *B. belcheri tsingtaunese*, 8e-32)	Ribosomal protein S10p/S20e, gnl|CDD|40433 (3e-17)
		38	EE049623	Bg ribosomal protein L11 like 1	ribosomal protein L11 (CAD91442, *C. gigas*, 6e-08)	ND
		39	EE049625	Bg ribosomal protein L37A like 1	60S ribosomal protein L37A (AAK52799, *B. belcheri*, 2e-05)	Ribosomal L37ae protein family, gnl|CDD|41808 (1e-03)
Nucleic acid biosynthesis and metabolism	Endonuclease	40	EE049562	Bg deoxyribonuclease like	deoxyribonuclease I (BAD83646, *P. major*, 5e-06)	DNaseIc, deoxyribonuclease I, gnl|CDD|47782 (3e-09)
	Reverse transcriptase	41	EE049564	Bg reverse transcriptase like	similar to reverse transcriptase family member (XS300) (XP_799214, *S. purpuratus*, 44e-02)	ND
	Transcription factor	42	EE049524	Bg Myc homolog like 1	Myc homolog (AAB34577, *C. virginica*, 4e-03)	ND
		43	EE049525	Bg Myc homolog like 2	Myc homolog (AAB34577, *C. virginica*, 5e-06)	ND

^a = The E-value is the probability due to chance, that there is another alignment with a similarity greater than the given one, obtained with BLASTX.

^b = The E-value is the probability due to chance, that there is another alignment with a similarity greater than the given one, obtained with RPS-BLAST.

Fig. 3. Transcript ratios corresponding to selected clusters of the EAF-S library in whole body susceptible EAF-S snails (light boxes) and resistant CB-R snails (dark boxes). Ratios were determined using real-time quantitative PCR and are expressed relative to S19 expression levels (ratio/S19). Each histogram is the average value ($\pm$ S.D.) of 3 independent experiments each performed in duplicate. Differential transcription levels between strains are indicated above the histograms. Since transcription levels can differ by up to 2 orders of magnitude, 3 different graphs were used for sake of clarity. *Genes showing differences in transcript levels that were significant ($P < 0.05$) and higher than 2·5-fold.

Candidate Genes:

1 EAF-S: Bg g-type lysozyme like
2 EAF-S: Bg serine proteinase like
3 EAF-S: Bg cathepsin L like
5 EAF-S: Bg serine protease beta like 1
7 EAF-S: Bg serpin like 1
8 EAF-S: Bg serpin like 2
9 EAF-S: Bg DEC-1 like
12 EAF-S: Bg ependymin like protein
13 EAF-S: Bg Cu/Zn SOD like
14 EAF-S: Bg calcium binding protein 3
15 EAF-S: Bg calcium binding protein 4
16 EAF-S: Bg DEC-3 like
17 EAF-S: Bg actin like

18 EAF-S: Bg actin-interacting protein 1 like
19 EAF-S: Bg alpha 2-tubulin like
21 EAF-S: Bg endo-1,4-beta-mannanase 1
23 EAF-S: Bg endo-1,4-beta-mannanase 3
24 EAF-S: Bg endo-1,4-beta-D-glucanase like 1
28 EAF-S: Bg hemoglobin
29 EAF-S: Bg Fatty acid binding protein like
30 EAF-S: Bg phosphoribosylformylglycinamidine synthase like
31 EAF-S: Bg Low-density lipoprotein receptor-related protein 1B precursor like
32 EAF-S: Bg proprotein convertase/kexin type 9 like
42 EAF-S: Bg Myc homolog like 1
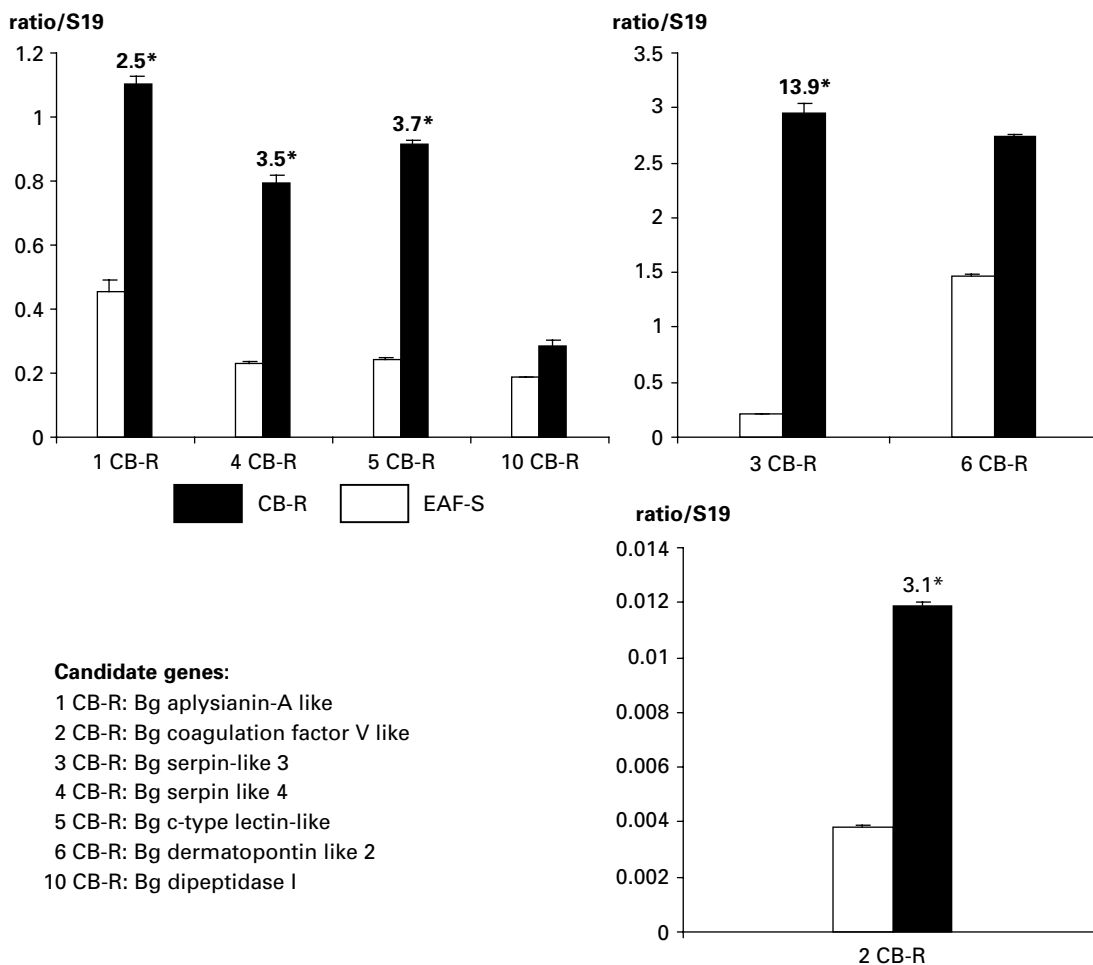43 EAF-S: Bg Myc homolog like 2

Fig. 4. Transcript ratios corresponding to selected clusters of the CB-R library in whole body susceptible EAF-S snails (light boxes) and resistant CB-R snails (dark boxes). Ratios were determined using real-time quantitative PCR and are expressed relative to S19 expression levels (ratio/S19). Each histogram is the average value ($\pm$ S.D.) of 3 independent experiments each performed in duplicate. Differential levels of the transcripts between strains are indicated above the histograms. Since transcription levels can differ by up to 2 orders of magnitude, 3 different graphs were used for sake of clarity. *Genes showing differences in transcript levels that were significant ($P<0.05$) and higher than 2·5-fold.

The deduced amino acid sequence from this cluster presents a C-type lectin domain. This protein was called Bg C-type lectin like. An extensive analysis of the deduced amino acid sequence also revealed 2 groups of 3 residues which are known to be involved in calcium binding and carbohydrate binding. One of these groups of residues (EPN) suggests that the molecule identified in this study could be a mannose-binding lectin. The same procedure was applied to all the clusters.

DISCUSSION

In order to investigate the molecular basis of susceptibility/resistance in the *B. glabrata/E. caproni* model, several molecular approaches were undertaken to compare susceptible and resistant snail strains (Vergote *et al*. 2005; Bouchut *et al*. 2006 *a, b*). These studies focused principally on haemocytes and humoral factors which have been shown to

play key roles in the susceptibility/resistance process (Ataev and Coustau, 1999). Regarding humoral factors, the proteomic approach revealed several proteins differentially represented in the plasma of susceptible and resistant snails (Vergote *et al*. 2005). They corresponded to 2 isoforms of a glycolytic enzyme (endo-1,4-β-mannanase 1 and -2), 2 isoforms of a calcium-binding protein (Bg CaBP-1 and -2) and an inhibitor of cysteine protease (Bg type-2 cystatin). All these proteins are expressed in the albumen gland (Vergote *et al*. 2005). Regarding haemocytes, we developed 2 molecular approaches based on the comparison of transcript and protein expression in haemocytes of both strains. The first transcriptomic approach focused on genes involved in adhesion processes and revealed 4 transcripts which were differentially represented between haemocytes from resistant and susceptible snails (Bouchut *et al*. 2006 *a*). These genes encode 2 dermatopontin-like proteins, a matrilin-like and a cadherin-like protein.

Table 4. Genes identified from the CB-R library

Class	Functional group	Cluster no.	Accession no. in dbEST	Name	Annotated BLASTX similarity (accession #, E-value)[a]	RPS-BLAST detected domain (E-value)[b]
Immunity	Cellular defence effector	1	EE049642	Bg aplysianin-A like	Aplysianin-A precursor (Q17043, *A. kurodai*, 4e-07)	ND
	Protease	2	EE049646	Bg coagulation factor V like	similar to coagulation factor V (XP_224967, *R. norvegicus*, 4e-18)	ND
	Protease inhibitor	3	EE049644	Bg serpin-like 3	Serpin (serine protease inhibitor) (AAX95869, *O. sativa* (japonica cultivar-group), 5e-04)	SERPIN, SERine Proteinase INhibitors (serpins), gnl|CDD|29117 (3e-06)
		4	EE049641	Bg serpin like 4	serpin-2 (BAD11156, *H. longicornis*, 0·12)	SERPIN, SERine Proteinase INhibitors (serpins), gnl|CDD|29117 (1e-05)
	Pattern Recognition Receptor	5	EE049640	Bg c-type lectin-like	lectin 1 (AAL29940, *G. tigrina*, 6e-05)	C-type lectin (CTL) and CTL-like domains, gnl|CDD|28919 (6e-04)
	Cell Adhesion	6	EE049651	Bg dermatopontin like 2	dermatopontin 2 (AAZ80786, *B. glabrata*, 5e-20)	ND
Energy, metabolism and hydrolytic enzymes	ATP biosynthesis	7	EE049643	Bg F1-ATP synthase subunit B like	F1-ATP synthase, beta subunit (CAI07821, *Azoarcus sp.*, 8e-14)	AtpD, F0F1-type ATP synthase, beta subunit [Energy production and conversion] gnl|CDD|30404 (2e-17)
		8	EE049692	Bg cytochrome b	cytochrome b (AAQ74246, *B. glabrata*, 5e-48)	Cytochrome b, gnl|CDD|29371 (7e-11)
		9	EE049687	Bg cytochrome c oxidase subunit III	cytochrome c oxidase subunit III (AAQ74252, *B. glabrata*, 2e-20)	Cyt_c_Oxidase_III, Cytochrome c oxidase subunit III, gnl|CDD|29485 (3e-14)
	Protease	10	EE049650	Bg dipeptidase I	Renal dipeptidase (RDP) isoform 1 (XP_780043, *S. Purpuratus*, 2e-41)	Renal dipeptidase (rDP), gnl|CDD|30044 (3e-47)
Protein biosynthesis	Ribosomal protein	11	EE049684	Bg ribosomal proteine S30 like	ribosomal protein S30 (BAE16016, *C. gigas*, 5e-16)	Ribosomal protein S30, gnl|CDD|44665 (4e-05)
		12	EE049686	Bg 40S ribosomal proteine S6 like 1	ribosomal protein S6 (AAG60623, *A. californica*, 6e-26)	ND
		13	EE049688	Bg 40S ribosomal proteine S6 like 2	40S ribosomal protein S6 (AAK95188, *I. punctatus*, 4e-35)	Ribosomal protein S6e, gnl|CDD|41158 (5e-30)
		14	EE049690	Bg ribosomal proteine L11 like	similar to ribosomal protein L11 (XP_797802, *S. purpuratus*, 2e-21)	Ribosomal L5P family, gnl|CDD|40755 (3e-09)
		15	EE049693	Bg ribosomal proteine L37A	Ribosomal protein L37A CG5827-PB, isoform B (NP_723060, *D. melanogaster*, 3e-33)	Ribosomal L37ae protein family, gnl|CDD|41808 (2e-23)
Nucleic acid biosynthesis and metabolism	Endonuclease	16	EE049648	Bg endonuclease G like	endonuclease G (NP_787017, *B. taurus*, 8e-19)	NUC, DNA/RNA non-specific endonuclease, gnl|CDD|28975 (2e-17)
		17	EE049647	Bg endonuclease G2-like	Endonuclease G (NP_787017, *B. Taurus*, 1e-11)	NUC, DNA/ RNA non-specific endonuclease, gnl| CDD|28975 (1e-03)

[a] = The E-value is the probability due to chance, that there is another alignment with a similarity greater than the given one, obtained with BLASTX.
[b] = The E-value is the probability due to chance, that there is another alignment with a similarity greater than the given one, obtained with RPS-BLAST.

The same study (Bouchut *et al.* 2006*a*) also revealed different expression levels in whole snails of different genes potentially involved in extracellular matrix (ECM) structuring or coagulation. A second approach we developed aimed at identifying proteins differentially represented in haemocytes of the 2 strains. This study (Bouchut *et al.* 2006*b*) revealed an aldolase, an intermediate filament protein, a cytidine deaminase.

In the present work, we used another molecular comparative method termed SSH to identify transcripts differentially expressed between the 2 mollusc strains of interest. This approach was performed on entire snails in order to investigate the potential involvement of genes expressed in other snail tissues. This approach revealed several genes identified in the above-mentioned studies but also new candidates belonging to novel functional groups of interest.

The functional groups of – glycolytic enzymes and calcium-binding proteins (CaBP) – were identified in the present work but also in plasma of *B. glabrata* as previously shown (Vergote *et al.* 2005). Bg endo-1,4-β-mannanase 1 was found in both cases. In addition, a novel isoform of the same family of proteins, the Bg endo-1,4-β-mannanase 3 was identified in the present study. In both works, the transcripts corresponding to these enzymes were over-represented in susceptible snails. This type of glycosidase randomly cleaves within the β-1,4-mannan main chain of galactomannan, glucomannan, galactoglucomannan and mannan. In the present work another glycolytic enzyme, named Bg endo-1,4-β-D-glucanase like 1, was revealed. Bg endo-1,4-β-D-glucanase like 1 transcripts were also over-represented in susceptible snails. These data taken together reinforce the previously presented hypothesis that glycolytic enzymes could be involved in the *B. glabrata/E. caproni* interactions (Vergote *et al.* 2005).

For the functional group corresponding to CaBP, a similar situation is observed. Indeed, the comparative proteomic approach developed on plasma revealed an over-representation of transcripts and proteins of plasmatic CaBP (Bg CaBP-1 and Bg CaBP-2). In the present study, 3 novel members of this protein family were identified (Bg CaBP-3, -4 and Bg DEC-3) and their transcripts are also over-expressed in susceptible snails. Regarding links between differential representation of these molecules and susceptibility/resistance, a possibility is that these proteins could be involved in regulation of adhesion processes as described in a previous study (Vergote *et al.* 2005).

Regarding novel functional groups of interest, the most promising candidates identified in the present study belong to immunity class. Some of them were characterized from the EAF-S library. The first functional group of interest identified in this library contains several proteases showing a higher expression in EAF-S snails: Bg cathepsin L like and Bg serine protease beta like 1. One attractive hypothesis is that a higher content in proteases could facilitate the parasite settlement in EAF-S snails. Indeed, as an alternative to produce their own proteases, some parasites (like pathogenic bacteria) can use proteases produced by host and turn them to their own advantage to facilitate invasion and dissemination (Armstrong, 2006). In addition, clusters corresponding to cell adhesion proteins were characterized from the EAF-S library. Two of them presented higher transcript levels in EAF-S snails. The first molecule, named Bg DEC-1 like, displays high similarities for a gene DEC-1 (dextral enriched clone-1) identified from the snail *Lymnaea stagnalis* (Harada *et al.* 2004). This molecule possesses chitin binding and von Willebrand type A (vWA) domains, suggesting a putative involvement in adhesion processes. The higher representation of this kind of molecule in susceptible snails can be linked to results obtained in a previous study (Bouchut *et al.* 2006*a*). In this last work, a series of genes encoding proteins containing vW domains potentially involved in coagulation processes have been shown to be more expressed (higher transcript level) in EAF-S snails. This led to the hypothesis that susceptible snails could possess a more potent collagen-independent filamentous network which could prevent haemocyte migration towards the parasite larvae and therefore facilitate parasite settlement in susceptible snails (Bouchut *et al.* 2006*a*). The second adhesion molecule, named Bg ependymin like, displays similarities for mammalian ependymin related proteins (MERPs). Ependymin was originally identified in the ependymal zone of goldfish brain and is a constituent of the extracellular matrix. Ependymins may be involved in cell contact mechanisms through anti-adhesive properties (Hoffmann and Schwarz, 1996). These anti-adhesive molecules could play a key role in establishing specific cell contacts during neural regeneration, differentiation and cell migration (Nimmrich *et al.* 2001). MERPs genes have been shown to be expressed at high levels in several haematopoietic cell lines as well as in non-haematopoietic tissues such as brain, heart, and skeletal tissues and by malignant tissues and malignant cell lines. This broad tissue distribution suggests a role for this MERP protein in processes that are important in many other types of cells and tissues (Apostolopoulos *et al.* 2001). The Bg ependymin-like transcript we identified in this study is over-represented in EAF-S snails. This candidate gene could be of potential interest in our system considering its putative involvement in anti-adhesive processes. The higher expression of these molecules and their associated anti-adhesive properties could prevent parasite encapsulation in EAF-S snails. Further studies investigating the tissue expression

of the Bg ependymin-like should be developed in order to test this hypothesis.

Several other novel clusters identified from the EAF-S library and belonging to another class (energy, metabolism and hydrolytic enzymes) also display a transcript content that are superior in EAF-S snails. These clusters correspond to functional groups like transport (Bg haemoglobin and Bg fatty acid binding protein) and cholesterol homeostasis (cluster 30–32). Considering their putative function, it is difficult to speculate on the potential role of these genes in our system.

Regarding genes identified from the CB-R library, several genes potentially involved in immune processes were also identified. Among them, we identified a cluster displaying similarities for a defence factor Aplysianin A purified from the albumen gland of *Aplysia kurodai* (Takamatsu *et al.* 1995). This glycoprotein inhibited the growth of both Gram-positive and Gram-negative bacteria. Another study developed in *Aplysia punctata* has evidenced a molecule called APIT (*A. punctata* ink toxin) displaying about 60% identity with Aplysianin A (Butzke *et al.* 2005). APIT is an L-amino acid oxidase, which produces hydrogen peroxide and was shown to lead to necrosis-like oxidative damage of Jurkat-T eukaryotic cells. The Bg Aplysianin A-like we identify in this study could be a defence factor involved in the anti-parasitic response of *B. glabrata*. Its higher abundance in CB-R snails could explain the efficiency of the resistant snail response against *E. caproni*. Further studies investigating its biological activity and tissue localization could help to strengthen this hypothesis.

The second gene of interest identified from the CB-R library is a transcript encoding a C-type lectin, we called Bg C-type lectin-like. This molecule displays an amino acid motif which suggests mannose-binding capabilities. The majority of C-type lectins characterized in invertebrates are humoral sugar-binding proteins with a demonstrated or suggested involvement in pathogen-associated molecular pattern recognition (e.g. from mollusc (Yuasa *et al.* 1998), crustaceans (Luo *et al.* 2003), flat worm (Duclermortier *et al.* 1999) and nematode (Loukas *et al.* 1999)). The putative function of this protein of *B. glabrata* and the over-abundance of the corresponding transcripts in CB-R snails suggest a potential involvement of this gene in resistance. Future functional studies should be developed to evaluate this putative role.

Finally, several clusters corresponding to protease inhibitors were characterized from the CB-R library. This kind of molecule often plays a crucial role in anti-parasitic defence. Two serpin-like molecules were identified from the CB-R library. Serpins (serine protease inhibitors) are protease inhibitors that are capable of inactivating and clearing the proteases involved in parasitic invasion and thus, that

are an important component of the armamentarium of the immune system to counter parasitic invasion (see Armstrong (2006) for review). Because proteases constitute one of the important classes of virulence and survival factors, the host has developed inhibitors of proteases which have evolved as important elements in the system of host defences against pathogens and as regulators of endogenous proteases. Inactivation of the secreted proteases of pathogens has the potential to restrict the steps of invasion and dispersal in the internal milieu, and parasite-mediated inactivation of other elements of host immunity. In invertebrates such as arthropod (see Kanost (1999) for review), serpins are likely to function in protecting their hosts from pathogens or parasites and thus play several roles in immunity: protection against microbial proteases, regulation of endogenous proteases involved in cascade responses against pathogens, haemolymph coagulation, phenoloxidase activation and proteolytic cytokines activation. Recently, a mosquito serpin, *Anopheles stephensi* SRPN6, has been found to be strongly implicated in the innate immune response against *Plasmodium* by RNA interference (Abraham *et al.* 2005). *As*SRPN6 knockdown significantly increased the number of developing parasites, which suggests that *As*SRPN6 could be involved in the parasite killing process. Higher abundance of 2 serpin-like molecules in CB-R snails and the putative function of these molecules make them very attractive candidates in our system. Future functional studies have to be applied to verify their potential interest.

In conclusion, this SSH approach has been successful in identifying new transcripts that were differentially represented in *E. caproni*-susceptible and resistant *B. glabrata*. Similarities of the corresponding genes provided informative data on the potential functions of these proteins. According to these predicted functions, various assumptions can be made on their potential involvement in susceptibility/resistance processes. Future functional studies are needed to clarify the role of these genes in the host-parasite interaction.

REFERENCES

Abraham, E. G., Pinto, S. B., Ghosh, A., Vanlandingham, D. L., Budd, A., Higgs, S., Kafatos, F. C., Jacobs-Lorena, M. and Michel, K. (2005). An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites. *Proceedings of the National Academy of Sciences, USA* **102**, 16327–16332. doi: 10.1073/pnas.0508335102.

Apostolopoulos, J., Sparrow, R. L., McLeod, J. L., Collier, F. M., Darcy, P. K., Slater, H. R., Ngu, C., Gregorio-King, C. C. and Kirkland, M. A. (2001).

Identification and characterization of a novel family of mammalian ependymin-related proteins (MERPs) in hematopoietic, nonhematopoietic, and malignant tissues. *DNA and Cell Biology* **20**, 625–635.

Armstrong, P. B. (2006). Proteases and protease inhibitors: a balance of activities in host-pathogen interaction. *Immunobiology* **211**, 263–281. doi: 10.1016/j.imbio.2006.01.002.

Ataev, G. L. and Coustau, C. (1999). Cellular response to *Echinostoma caproni* infection in *Biomphalaria glabrata* strains selected for susceptibility/resistance. *Developmental and Comparative Immunology* **23**, 187–198. doi: 10.1016/S0145-305X(99)00023-3.

Bouchut, A., Roger, E., Coustau, C., Gourbal, B. and Mitta, G. (2006*a*). Compatibility in the *Biomphalaria glabrata/Echinostoma caproni* model: potential involvement of adhesion genes. *International Journal for Parasitology* **36**, 175–184. doi: 10.1016/j.ijpara.2005.09.009.

Bouchut, A., Sautiere, P. E., Coustau, C. and Mitta, G. (2006*b*). Compatibility in the *Biomphalaria glabrata/ Echinostoma caproni* model: potential involvement of proteins from hemocytes revealed by a proteomic approach. *Acta Tropica* **98**, 234–246. doi: 10.1016/j.actatropica.2006.05.007.

Butzke, D., Hurwitz, R., Thiede, B., Goedert, S. and Rudel, T. (2005). Cloning and biochemical characterization of APIT, a new l-amino acid oxidase from *Aplysia punctata*. *Toxicon* **46**, 479–489. doi: 10.1016/j.toxicon.2005.06.005.

Duclermortier, P., Lardans, V., Serra, E., Trottein, F. and Dissous, C. (2005). *Biomphalaria glabrata* embryonic cells express a protein with a domain homologous to the lectin domain of mammalian selectins. *Parasitology Research* **85**, 481–486.

Guillou, F., Mitta, G., Dissous, C., Pierce, R. and Coustau, C. (2004). Use of individual polymorphism to validate potential functional markers: case of a candidate lectin (BgSel) differentially expressed in susceptible and resistant strains of *Biomphalaria glabrata*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **138**, 175–181. doi: 10.1016/j.cbpc.2004.03.010.

Harada, Y., Hosoiri, Y. and Kuroda, R. (2004). Isolation and evaluation of dextral-specific and dextral-enriched cDNA clones as candidates for the handedness-determining gene in a freshwater gastropod, *Lymnaea stagnalis*. *Development Genes and Evolution* **214**, 159–169. doi: 10.1007/s00427-004-0392-6.

Hoffmann, W. and Schwarz, H. (1996). Ependymins: meningeal-derived extracellular matrix proteins at the blood-brain barrier. *International Review of Cytology* **165**, 121–158.

Humbert, E. and Coustau, C. (2001). Refractoriness of host haemocytes to parasite immunosuppressive factors as a putative resistance mechanism in the *Biomphalaria glabrata-Echinostoma caproni* system. *Parasitology* **122**, 651–660.

Jung, Y., Nowak, T. S., Zhang, S. M., Hertel, L. A., Loker, E. S. and Adema, C. M. (2005). Manganese superoxide dismutase from *Biomphalaria glabrata*. *Journal of Invertebrate Pathology* **90**, 59–63.

Kanost, M. R. (1999). Serine proteinase inhibitors in arthropod immunity. *Developmental and Comparative Immunology* **23**, 291–301. doi: 10.1016/S0145-305(99)00012-9.

Langand, J. and Morand, S. (1998). Heritable non-susceptibility in an allopatric host-parasite system: *Biomphalaria glabrata* (Mollusca)-*Echinostoma caproni* (platyhelminth Digenea). *Journal of Parasitology* **84**, 739–742.

Langand, J., Jourdane, J., Coustau, C., Delay, B. and Morand, S. (1998). Cost of resistance, expressed as a delayed maturity, detected in the host-parasite system *Biomphalaria glabrata/Echinostoma caproni*. *Heredity* **80**, 320–325.

Lewis, F. A., Richards, C. S., Knight, M., Cooper, L. A. and Clark, B. (1993). *Schistosoma mansoni*: analysis of an unusual infection phenotype in the intermediate host snail *Biomphalaria glabrata*. *Experimental Parasitology* **77**, 349–361. doi: 10.1006/expr.1993.1092.

Lie, K. J. (1982). Swellengrebel lecture: survival of *Schistosoma mansoni* and other trematode larvae in the snail *Biomphalaria glabrata*. A discussion of the interference theory. *Tropical and Geographical Medicine* **34**, 111–122.

Lockyer, A. E., Noble, L. R., Rollinson, D. and Jones, C. S. (2004). *Schistosoma mansoni*: resistant specific infection-induced gene expression in *Biomphalaria glabrata* identified by fluorescent-based differential display. *Experimental Parasitology* **107**, 97–104. doi: 10.1016/j.exppara.2004.04.004.

Loukas, A., Mullin, N. P., Tetteh, K. K., Moens, L. and Maizels, R. M. (1999). A novel C-type lectin secreted by a tissue-dwelling parasitic nematode. *Current Biology* **9**, 825–828. doi: 10.1016/S0960-9822(99)80366-2.

Luo, T., Zhang, X., Shao, Z. and Xu, X. (2003). PmAV, a novel gene involved in virus resistance of shrimp *Penaeus monodon*. *FEBS Letters* **551**, 53–57. doi: 10.1016/S0014-5793(03)00891-3.

Miller, A. N., Raghavan, N., FitzGerald, P. C., Lewis, F. A. and Knight, M. (2001). Differential gene expression in haemocytes of the snail *Biomphalaria glabrata*: effects of *Schistosoma mansoni* infection. *International Journal for Parasitology* **31**, 687–696. doi: 10.1016/S0020-7519(01)00133-3.

Nimmrich, I., Erdmann, S., Melchers, U., Chtarbova, S., Finke, U., Hentsch, S., Hoffmann, I., Oertel, M., Hoffmann, W. and Muller, O. (2001). The novel ependymin related gene UCC1 is highly expressed in colorectal tumor cells. *Cancer Letters* **165**, 71–79. doi: 10.1016/S0304-3835(01)00390-1.

Nowak, T. S., Woodards, A. C., Jung, Y., Adema, C. M. and Loker, E. S. (2004). Identification of transcripts generated during the response of resistant *Biomphalaria glabrata* to *Schistosoma mansoni* infection using suppression subtractive hybridization. *Journal of Parasitology* **90**, 1034–1040.

Raghavan, N., Miller, A. N., Gardner, M., FitzGerald, P. C., Kerlavage, A. R., Johnston, D. A., Lewis, F. A. and Knight, M. (2003). Comparative gene analysis of *Biomphalaria glabrata* hemocytes pre- and post-exposure to miracidia of *Schistosoma mansoni*. *Molecular and Biochemical Parasitology* **126**, 181–191. doi: 10.1016/S0166-6851(02)00272-4.

Richards, C. S. (1975). Genetic factors in susceptibility of *Biomphalaria glabrata* for

different strains of *Schistosoma mansoni*. *Parasitology* **70**, 231–241.

Richards, C. S. and Shade, P. C. (1987). The genetic variation of compatibility in *Biomphalaria glabrata* and *Schistosoma mansoni*. *Journal of Parasitology* **73**, 1146–1151.

Shagin, D. A., Barsova, E. V., Bogdanova, E., Britanova, O. V., Gurskaya, N., Lukyanov, K. A., Matz, M. V., Punkova, N. I., Usman, N. Y., Kopantzev, E. P., Salo, E. and Lukyanov, S. A. (2002). Identification and characterization of a new family of C-type lectin-like genes from planaria *Girardia tigrina*. *Glycobiology* **12**, 463–472.

Takamatsu, N., Shiba, T., Muramoto, K. and Kamiya, H. (1995). Molecular cloning of the defense factor in the albumen gland of the sea hare *Aplysia kurodai*. *FEBS Letters* **377**, 373–376. doi: 10.1016/0014-5793(95)01375-X.

Vergote, D., Bouchut, A., Sautiere, P. E., Roger, E., Galinier, R., Rognon, A., Coustau, C., Salzet, M. and Mitta, G. (2005). Characterisation of proteins differentially present in the plasma of *Biomphalaria glabrata* susceptible or resistant to *Echinostoma caproni*. *International Journal for Parasitology* **35**, 215–224. doi: 10.1016/j.ijpara.2004.11.006.

Yuasa, H. J., Furuta, E., Nakamura, A. and Takagi, T. (1998). Cloning and sequencing of three C-type lectins from body surface mucus of the land slug, *Incilaria fruhstorferi*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **119**, 479–484. doi: 10.1016/S0305-0491(98)00008-X.