Deliverable Report

Deliverable No:	D25.1	**Delivery Month:**	40
Deliverable Title	Marker genes of mucosal immunity in greater amberjack cloned and ways to increase their expression level determined		
WP No:	25	**WP Lead beneficiary:**	P5. UNIABDN
WP Title:	Fish health - Greater amberjack		
Task No:	25.3	**Task Lead beneficiary:**	P5. UNIABDN
Task Title:	Identification of immune markers		
Other beneficiaries:	P2. FCPCT		
Status:	**Delivered**	**Expected month:**	39
…….			

Lead Scientist preparing the Deliverable: Secombes, C.J. (UNIABDN)

Other Scientists participating: Milne, D.J. (UNIABDN), Montero, D. (FCPCT), Acosta, F. (FCPCT), Fernández-Montero A. (FCPCT).

Objective: Identify and clone marker genes key to the greater amberjack mucosal immune response and determine methods to increase their expression.

Description: Nucleotide sequence has been obtained (minimum of partial sequence) for all the genes described in the DOW with additional genes also discovered not mentioned in the DOW. These additional genes (such as TNFα, IL-1β, IL-8, IFN) allow a more complete analysis of immune responses in greater amberjack following immunostimulation or infection, including analysing the effects of immune stimulants incorporated into functional feeds (DL25.2/DL25.3).

Genes with sequence obtained

Sequence for 19 genes was obtained, via molecular cloning and sequencing, in order to develop QPCR primers to monitor gene expression of key immune genes in the mucosal immune response of the greater amberjack *Seriola dumerili* (see **Table 1**).

Table 1. Greater amberjack immune genes sequenced to date.

Cytokines	IL-1β	Anti-viral	Mx	Housekeeping Genes	EF-1a
	IL-8		IFN1		β-Actin
	IL-10		IFNγ	Antimicrobial Peptides	Piscidin
	IL-17A/F		iNOS		Defensin
	IL-17D	Immunoglobulin Related	IgM		Hepcidin
	IL-22		IgT		
	TNFα		RAG2		

Confirmation of gene sequence

Upon obtaining sequence, the genes were identified and confirmed to be the target genes. This was achieved using protein prediction software to identify protein structures and domains followed by BLAST analysis of the amino acid sequence of the target gene. Alignment of the target gene sequence with the resulting BLAST hits, in order to identify conserved protein motifs (eg TNFα domain for TNFα) or amino acids (eg conserved cysteine residues in IgM and IgT) was also undertaken. Finally, a phylogenetic tree was also constructed in order to confirm the identity of the gene and to give an insight into the isoform of the gene sequence.

NOTE

The following sections will use the target gene, hepcidin as an example for this process in all figures.

Conserved protein motif and residue identification was achieved by use of SIGNAL P, ExPasY Prosite software and current literature (see **Figure 1**). Signal P software identified the signal peptide region of the hepcidin, which is an essential component of secreted molecules, as anti-microbial peptides are known to be. ExPasy Prosite identified an antimicrobial peptide region spanning from the QSH iron binding motif, conserved in hepcidin proteins, until the c terminal of the protein indicating the active peptide. Within this active peptide region 8 conserved cysteine residues, which will form di-sulphide bonds, are found, a typical characteristic of the hepcidin protein. Further evidence supporting the identified active peptide region is the presence of an RXXR motif located before the active peptide site as this is a common indicator of an enzymatic cut site of the Furin-like family of enzymes. These conserved motifs and residues strongly indicate the sequence identified is that of a hepcidin protein.

MKAFSIAVAVTLVLAFICILESSAVPFHGVRELEEAGSNDTPVVARQEMSVSSWMMPNPV
RQKRQSHLSMCHWCCNCCTANKGCGFCCRF

Figure 1 – Amino acid sequence for greater amberjack hepcidin with conserved motifs and residues. Blue – signal peptide; yellow - enzyme cleavage site; light grey – conserved iron binding motif; dark grey – conserved cysteines; underlined – active peptide region.

The sequence was then subject to NCBI blast analysis, which provides better confirmation of the identity of the molecule by identifying currently known proteins of similar sequence. As can be seen in **Figure 2**, blast analysis identified hepcidin in species other than the greater amberjack to have the highest similarity and based on the nomenclature of a number of the hits we are able to ascribe the isoform of the gene present in addition to gene identity.

Gene Name	Max score	Total score	Query cover	E value	Identity	Accession
hepcidin *[Larimichthys crocea]*	165	165	100%	1e-51	87%	ABC18307.1
hepcidin-1 *[Micropterus salmoides]*	164	164	100%	3e-51	87%	ACD13023.1
hepcidin-1 *[Micropterus dolomieu]*	163	163	100%	9e-51	86%	ACD13025.1
hepcidin antimicrobial peptide 1 [*Pagrus auriga*]	162	162	100%	3e-50	83%	BAH03285.1

hepcidin 1 *[Dicentrarchus labrax]*	159	159	100%	5e-49	85%	AJU35234.1
hepcidin *[Eleginops maclovinus]*	157	157	100%	2e-48	83%	ABY84822.1
hepcidin *[Maylandia zebra]*	157	157	100%	2e-48	83%	XP_004551096.1

Figure 2. A summary of the results of the top 7 hits when the greater amberjack hepcidin amino acid sequence was queried.

To be sure there were no accidental hits in the blast analysis the amino acid sequence for all the top 7 hits of the blast analysis and the queried greater amberjack sequence were aligned, using MAFFT alignment software, as seen in **Figure 3**, to further confirm that sequences sufficiently resembled each other and that the conserved domains remained within these sequences and matched the query sequence.

In **Figure 3** it can be seen that when the query sequence is aligned with the blast hits, the regions described earlier as conserved motifs and residues are found within the other proteins. In some cases these regions are almost fully conserved throughout all species, as seen with the signal peptide with the exception of small mouth bass *Micropterus dolomieu* hepcidin where there is a change from an A to V at residue 3, and the enzyme cleavage site with the exception of redbanded seabream *Pagrus auriga* where the second residue of this site is changed from Q to E.

```
Greater_Amberja MKAFSIAVAVTLVLAFICILESSAVPFHGVRELEEAGSNDTPVVARQEMSVSSWMMPNPV
Large_yellow_cr MKAFSIAVAVTLVLAFICILESSAVPFTGVQELEEAGSNDTPVAAHQEMSMESWMMPNHI
largemouth_bass MKAFSIAVAVTLVLAFICILESSAVPFTGVQELEEAGSNDTPVAAHQEMSMESWMMPNHI
Smallmouth_bass MKVFSIAVAVTLVLAFICILESSAVPFTGVQELEEAGSNDTPVAAHQEMSMESWMMPNHI
European_seabas MKAFSIAVAVTLVLAFICILESSAVPFAGVQELEEAGSNDTPVAAHQEMSMESWMMPNHI
Patagonian_blen MKAFSIAVAVTLVLAFICILESSAVPFTGVQELEEAGGNDTPVALYQEMSMESRMMPNHI
zebra_mbuna_hep MKAFSIAVAVTLVLAFICILESSAIPFAGVQEPEEAGSNDTPVVAHQEMSAESSMMPNHI
Pagrus_auriga_h MKAFSIAVAVTLVLAFICILESSAVPLNGVQELEEAGSNDTPVAAHQEMSMESWMMPSRV
                **.**********************:*: **:* ****.*****.  **** .* ***. :

Greater_Amberja -RQKRQSHLSMCHWCCNCCTANKGCGFCCRF
Large_yellow_cr -RQKRQSHLSLCRWCCNCCKSNKGCGFCCRF
largemouth_bass -RQKRQSHLSLCRWCCNCCRGNKGCGFCCRF
Smallmouth_bass -RQKRQSHLSLCRWCCNCCKGNKGCGFCCRF
European_seabas SRQKRQSHLSLCRWCCNCCRGNKGCGFCCKF
Patagonian_blen -RQKRQSHLSLCRWCCNCCKGNKGCGFCCRF
zebra_mbuna_hep -RQKRQSHLSLCRWCCNCCRSNKGCGFCCKF
Pagrus_auriga_h -REKRQSHISMCYWCCNCCRANKGCGYCCKF
                 *:*****:*:* ****** .*****:**:*
```

Figure 3 – An alignment of the top blast hits when querying the greater amberjack hepcidin sequence with conserved motifs and residues identified. Blue – signal peptide; yellow - enzyme cleavage site; light grey – conserved iron binding motif; dark grey – conserved cysteines; underlined – active peptide region.

Finally, phylogenetic analysis was performed in an effort to conclusively identity the query sequence but also to give insight into the gene isoform when multiple genes exist. This was achieved using MEGA6 software for alignment and phylogenetic analysis. As can be seen in **Figure 4** the query sequence groups with hepcidin 1 of other species and is a strong indicator that the protein identification is correct.

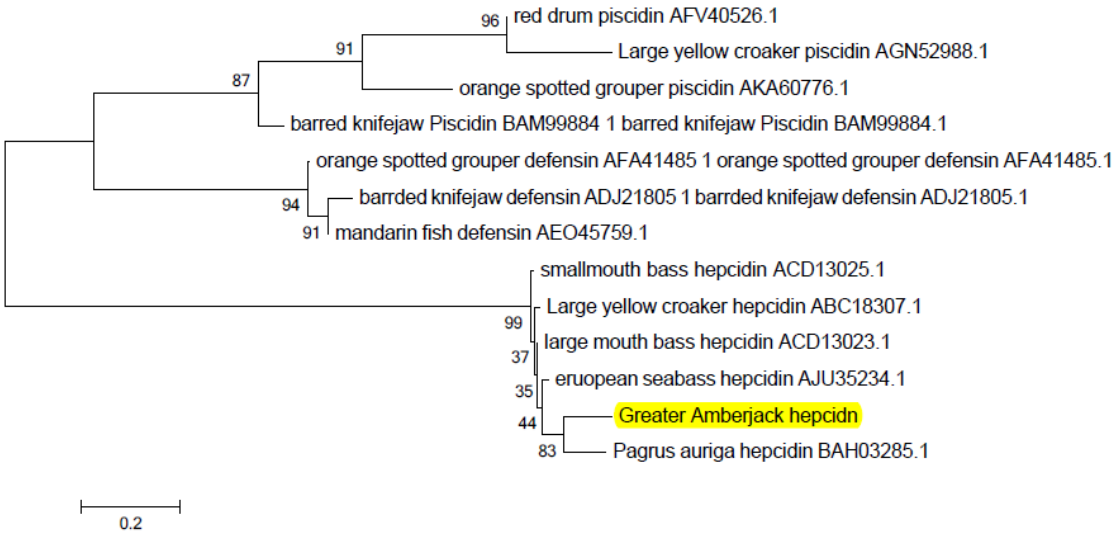

Figure 4. An unrooted phylogenetic tree showing that the greater amberjack sequence groups with the equivalent antimicrobial peptide of other species. The tree was constructed using Mega6 neighbour-joining method and the Jones-Taylor-Thorn matrix with pairwise deletion. The nodes show the bootstrap values of 10,000 replicates.

Generation of QPCR primers

Multiple QPCR primer sets were identified for each target gene and the best were selected for use in future work. This was achieved by taking the sequence information discovered, as mentioned above, and using IDT oligoanalyser software to determine the compatibility of primer pair candidates. Once suitable candidates were obtained they were used in the QPCR process with known amounts of CDNA to ensure a single PCR product was generated (**Figure 5**) and to obtain an efficiency value. Efficiency for all end result QPCR primer pairs is 95% or greater with the exception of the QPCR primer pair for hepcidin, which is 93% efficient (**Table 2**).

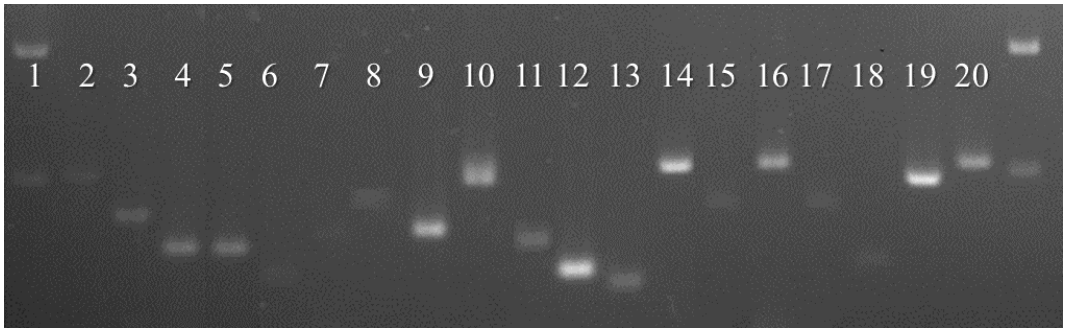

Figure 5. Gel containing PCR products for all amberjack QPCR primers. From left to right 1) ladder, 2) IL-1β, 3) IL-8, 4) IL-10, 5) IL-17A/F, 6) IL-17D, 7) IL-22, 8) RAG2, 9) IgM, 10) IgT, 11) Defensin, 12) Piscidin, 13) Hepcidin, 14) TNFα, 15) iNOS, 16) Mx, 17) IFNγ, 18) IFN1, 19) EF-1a, and 20) β-actin. Each increment on the ladder marks 200 bp.

Table 2. Properties of the QPCR primers used for analysis of the immune system of the greater amberjack. Ta (primer annealing temperature), Tm (primer melting temperature).

Gene	Product size (bp)	Efficiency (%)	Ta (°C)	Tm (°C)
IL-1β	205	99.5	54	89
IL-8	164	99.5	58	89
IL-10	134	99.75	58	84
IL-17A/F	120	99.6	62	88.5
IL-17D	111	99.6	62	91.5
IL-22	146	99.6	61	91
Defensin	133	97.8	58	87
Hepcidin	99	93.4	61	91
Piscidin	112	99.7	58	85
RAG2	162	99.4	55	86
IgM	148	97.8	58	86
IgT	196	98.85	55	87
Mx	211	96.8	61	86
INOS	151	99.05	60	88
IFN type 1	111	99.2	54	86
IFNγ	163	97	55	88
TNFα	212	99.1	55	91
EF-1a	194	99.05	60	89
β-Actin	212	96.8	61	90

Studies of immune stimulation *in vivo*

The following section will use the anti-microbial peptide (AMP) genes as an example of the in vivo studies undertaken to investigate modulation of immune genes relevant to mucosal responses/protection.

Molecules well established for stimulation of the immune response were used in this experiment. They included the PAMPs polyinosinic:polycytidylic acid (Poly I:C), lipopolysaccharide (LPS) and recombinant flagellin, the latter produced in our laboratory from the fish pathogen *Yersinia ruckeri.* These molecules have been shown to be capable of eliciting an immune response after 24 h post-injection in many fish species, such as rainbow trout and turbot, so this timing was used here for expression analysis.

Fish were injected intraperitoneally (ip) with a single dose (optimal for other fish species) of each stimulant or phosphate buffered saline (PBS) as a control. Each stimulant or control group was comprised of 10 fish. 24 h after injection the head kidney, spleen, intestine (gut) and gills were collected and placed in RNA later, and stored at -80^{o}C until they were processed. Processing involved RNA extraction and CDNA synthesis. Once processing was complete the resulting CDNA was used in QPCR to study the expression level of the immune genes and house-keeping genes, the latter for normalization. This allowed us to determine the gene expression of each of the target genes in the control and stimulant groups.

Intraperitoneal injection of PAMPs successfully stimulated expression of greater amberjack AMP genes, as can be seen in **Figure 6**. When compared to samples from the control, PBS injected fish, it can be seen that there is significant upregulation of anti-microbial peptides in a stimulant and tissue dependent manner. Flagellin was particularly potent at upregulating beta-defensin, with the highest expression seen in head kidney and spleen. Hepcidin was also clearly upregulated by LPS and flagellin, in head kidney and gill tissue. Overall, this study demonstrates that Poly I:C, LPS and flagellin can be used as stimulants of the immune response in vivo.

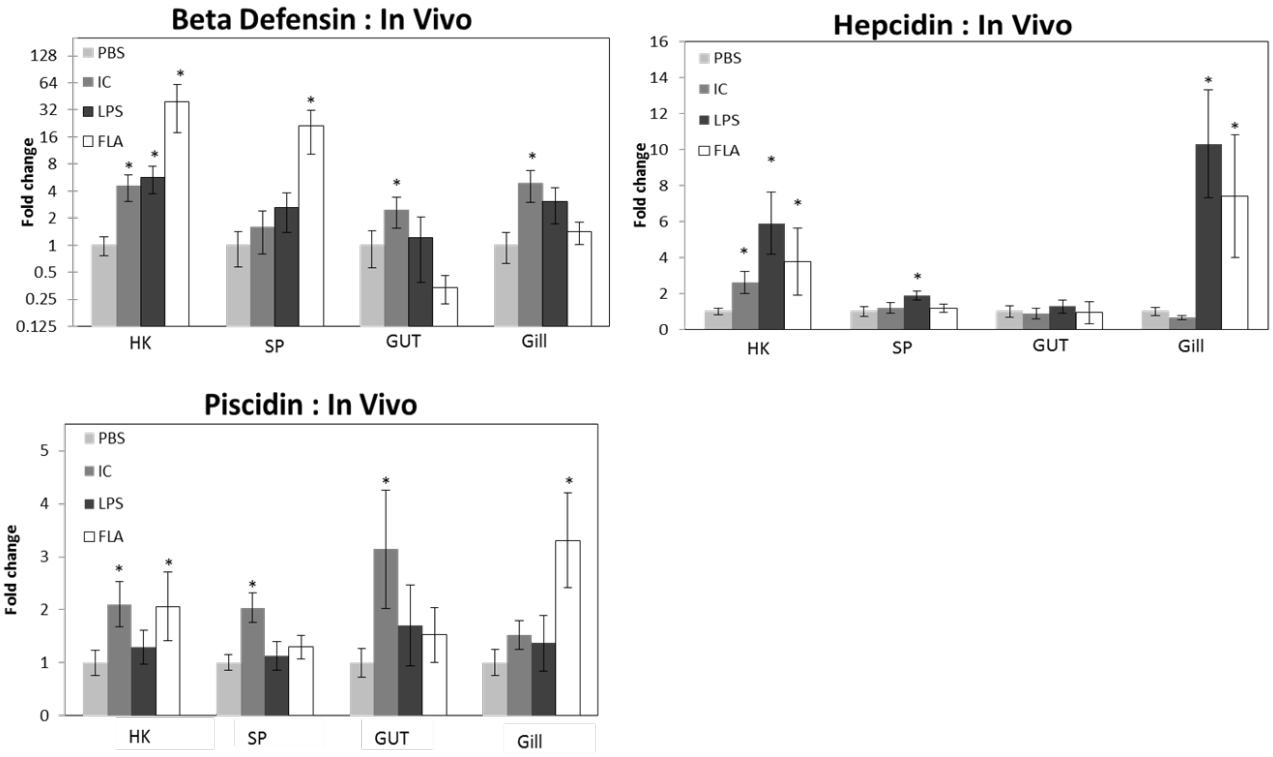

Figure 6. Gene expression of the identified greater amberjack anti-microbial peptides, in head kidney (HK), spleen (SP), intestine (GUT) and gills (Gill) after ip injection of phosphate buffered saline (PBS), Poly I:C (IC – 100 µg), lipopolysaccharide (LPS – 400 µg) and flagellin (FLA – 4 µg). Results are expressed as means ± SEM of the fold change calculated relative to the PBS injected fish samples. Asterisks indicate significant differences ($p \leq 0.05$) relative to the control group (Kruskal-Wallis test).

Studies of immune stimulation *in vitro*

The following section will use the study of AMPs in splenocytes as an example of the in vitro work performed to date.

As for the in vivo experiment, Poly I:C, LPS and flagellin were used as stimulants, with PBS as a control. For this in vitro stimulation experiment cells were isolated from 4 tissues (head kidney, spleen, intestine, gill) to generate primary cell cultures of each tissue. The primary cell cultures were then exposed to the stimulant, or PBS as control, for 4 h, 12 h or 24 h, to allow the kinetics of any gene induction to be studied over time.

After the designated time period the cells were harvested, stored in RNAlater, and subsequently processed as described above. The resulting CDNA was used in QPCR to determine the relative levels of gene expression of the target immune/AMP genes.

Greater amberjack leucocytes were successfully stimulated in vitro, as demonstrated by the changes in gene expression in the primary cell cultures (**Figures 7 and 8**). When comparing to the PBS exposed control cells, it can be seen that there was significant upregulation of AMP genes in a stimulant and time dependent manner. Flagellin was the most potent stimulant, affecting defensin, hepcidin and piscidin expression. Interestingly there was a temporal difference in the upregulation seen, with the effects on hepcidin and piscidin occurring relatively early in splenocytes vs head kidney cells. LPS was also a good inducer of hepcidin in splenocytes and all three AMPs in head kidney cells. Lastly, some small effects of Poly I:C stimulation on AMP expression were also seen.

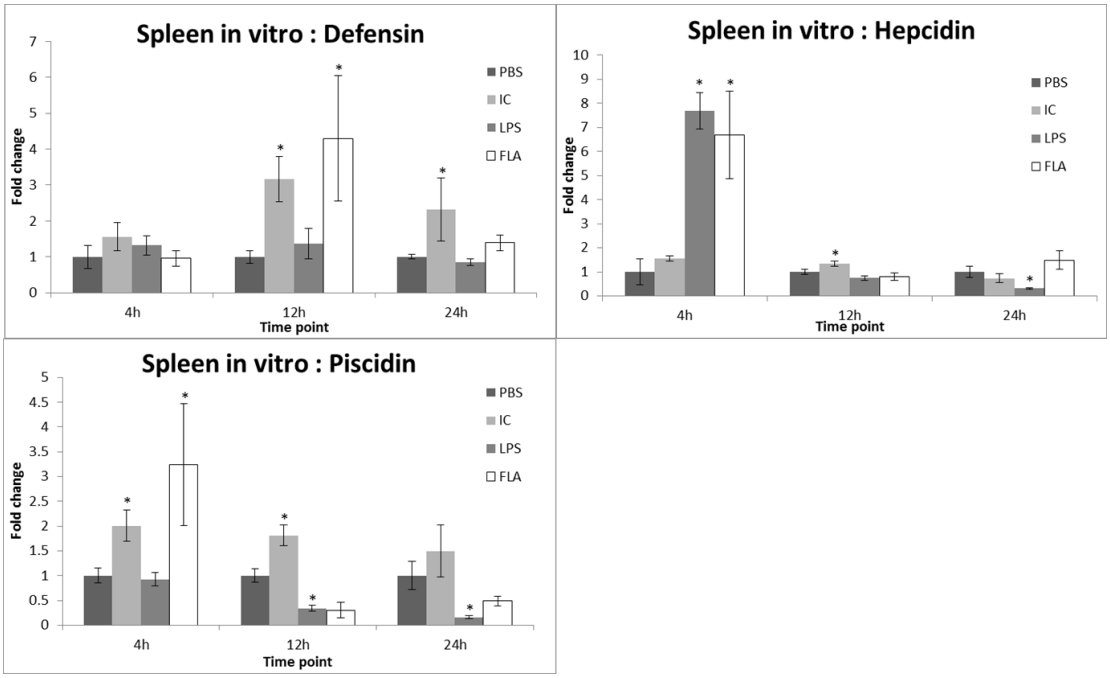

Figure 7. Gene expression of the identified greater amberjack anti-microbial peptides in splenocyte primary cell cultures, at 3 different times post-stimulation with Poly I:C (IC – 100 µg/ml), lipopolysaccharide (LPS – 50 µg/ml) and flagellin (FLA – 4 µg/ml) or phosphate buffered saline (PBS) as control. Results are expressed as means ± SEM of the fold change calculated relative to the PBS exposed cells.

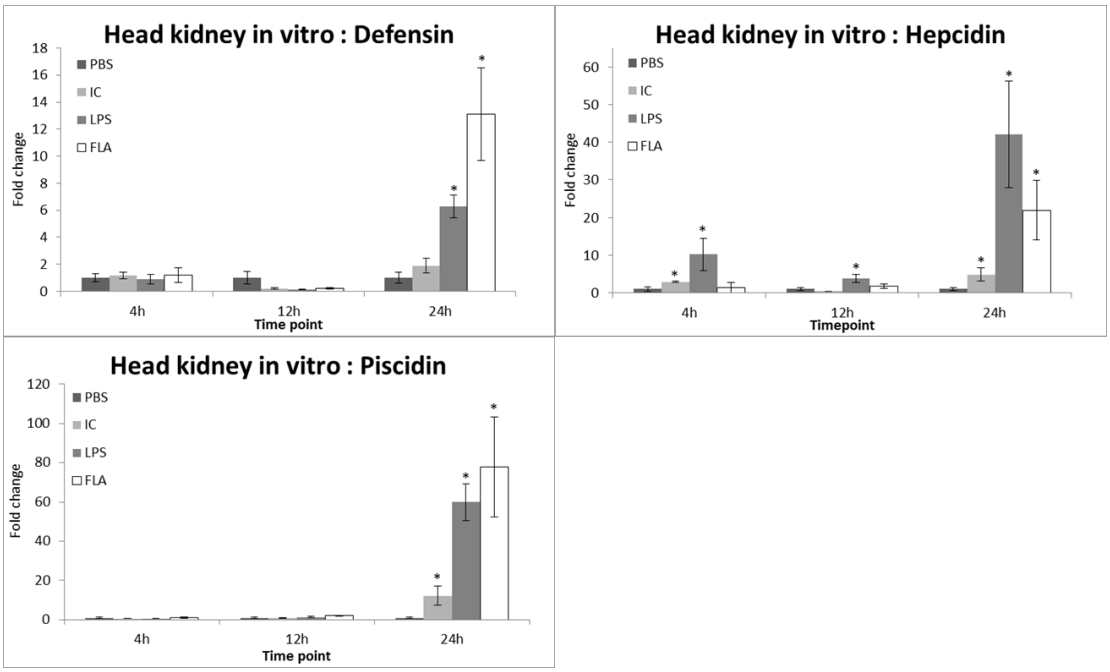

Figure 8. Gene expression of the identified greater amberjack anti-microbial peptides in head kidney primary cell cultures, at 3 different times post-stimulation with Poly I:C (IC – 100 µg/ml), lipopolysaccharide (LPS – 50 µg/ml) and flagellin (FLA – 4 µg/ml) or phosphate buffered saline (PBS) as control. Results are expressed as means ± SEM of the fold change calculated relative to the PBS exposed cells. Asterisks indicate significant differences ($p \leq 0.05$) relative to the control group (Kruskal-Wallis test).

Conclusions

These studies have sequenced a panel of relevant greater amberjack immune genes, to allow future study of mucosal immune responses in this species. QPCR assays have been optimised for each gene. In vivo and in vitro studies using PAMP stimulation have demonstrated that the expression of these genes can be modulated by such stimulants, and so these molecules are good markers for the effects of treatments (eg dietary or other) to increase disease resistance.

Deviations

None.

Co-funded by the Seventh Framework Programme of the European Union