PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Identification of novel host defense peptides and the absence of α-defensins in the bovine genome

Christopher D. Fjell,[1]* Håvard Jenssen,[2] Patrick Fries,[3] Palok Aich,[3] Philip Griebel,[3] Kai Hilpert,[2] Robert E. W. Hancock,[2] and Artem Cherkasov[1]

[1] Division of Infectious Diseases, Department of Medicine, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

[2] Centre for Microbial Diseases and Immunity Research, University of British Columbia, Lower Mall Research Station, Vancouver, British Columbia, Canada

[3] Vaccine and Infectious Disease Organization, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

ABSTRACT

Host defense peptides (historically called antimicrobial peptides, AMPs) are key components in the mammalian innate immune system, and are responsible for both direct killing and immunomodulatory effects in host defense against pathogenic organisms. In order to identify novel host defense peptides by sequence analysis, we constructed the AMPer resource (http://www.cnbi2.com/cgi-bin/amp.pl) that utilizes hidden Markov models to recognize sequences of antimicrobial peptides. In the current work, we utilized the AMPer resource to search bovine expressed sequence tags from the NCBI dbEST project and the bovine genome sequence for novel host defense peptides. Of the 34 known bovine AMPs, 27 were identified with high confidence in the AMPs predicted from ESTs. A further potential 68 AMPs predicted from the EST data were found that appear to be novel giving a total estimate of 102 AMPs present in the genome. Two of these were cathelicidins and selected for experimental verification in RNA derived from bovine tissue. One predicted AMP, most similar to rabbit '15 kDa protein' AMP, was confirmed to be present in infected bovine intestinal tissue using PCR. These findings demonstrated the practical applicability of the developed bioinformatics approach and laid a foundation for future discoveries of gene-coded AMPs. No members of the α-defensin family were found in the bovine sequences. Since we could find no technical reasons these would be missed and no references to bovine α-defensins in the literature, this suggests that cattle lack this important family of host defense peptides.

Proteins 2008; 73:420–430.

Key words: antimicrobial peptides; AMPer; hidden Markov models; innate immune system.

INTRODUCTION

Host defense peptides (known also as antimicrobial peptides, AMPs) are natural peptides produced as part of the innate immune system of a broad range of organisms including mammalians, insects, amphibians, plants, and amoeboid protozoa among others.[1–4] As the problem of antibiotic resistance to conventional therapeutics by pathogenic microorganisms increases, AMPs have drawn significant scientific attention as a novel class of prospective anti-infective therapeutics.[5–8] They offer several advantages including fast target killing, broad range of activity, low toxicity, and minimal development of resistance in target organisms.[1,2,5–8] Their mechanisms of killing are diverse and include membrane disruption[9–13] as well as metabolic inhibitors of intracellular targets.[14] In addition to direct killing, certain host defense peptides play important roles in modulation of the innate immune response both in upregulation to enhance killing of pathogens, as well as downregulation to reduce detrimental conditions such as sepsis.[15] The relative importance of direct killing by AMPs versus immunomodulation is also unclear.[16]

Limited numbers of novel antimicrobial peptides have been identified previously with the help of computational approaches.[17–23] The majority of these studies[18–22] searched specifically for the presence of a particular class of antimicrobial peptide, the defensins, which belong to three subfamilies: the α-, β- and θ-defensins. The commonly used techniques to identify novel peptides using sequence analysis are as follows: comparing examples of a class of peptides to a novel sequence in a pairwise fashion (for example using a BLAST analysis[24]) or using a set of

Additional Supporting Information may be found in the online version of this article.

Grant sponsors: Canadian Institutes for Health Research (CIHR) and Genome BC for the Pathogenomics of Innate Immunity research program.

*Correspondence to: Christopher D. Fjell, Division of Infectious Diseases, Department of Medicine, Faculty of Medicine, University of British Columbia, 2733 Heather Street, Vancouver, British Columbia, V5Z 3J5, Canada. E-mail: cfjell@interchange.ubc.ca.

Received 6 November 2007; Revised 13 February 2008; Accepted 25 February 2008
Published online 28 April 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22059

similar peptides to construct a profile of the class of peptides and then deriving a statistical model of the class for searching novel sequence (for example using profile hidden Markov models[25]). Profile hidden Markov models (HMMs) have been used extensively for large-scale analysis of protein sequences[25,26] and we have previously developed the AMPer resource[27] (http://www.cnbi2.com/cgi-bin/amp.pl) that includes HMMs to describe and predict AMPs based on peptide sequence. AMPer includes all AMPs that were available in the Uniprot database and separately describes mature peptides and propeptides have been determined based on Uniprot annotations. These have been grouped into sets of related peptides, with each set used to produce one hidden Markov model specific to that subclass of AMP. AMPer includes 1045 mature peptides (with 146 corresponding models) and 253 propeptide sequences (with 40 corresponding models) derived from 970 Uniprot proteins.

Models from AMPer provide the means to perform high-throughput analysis to discover novel AMPs that are related to peptides that are currently known. This serves to identify additional peptides that may have antimicrobial activity and may suggest the absence of a class of peptide in an organism. As an example, we consider the α-defensins: there are currently no recognized α-defensins in the bovine genome. Phylogenetic analysis of defensins has suggested that all defensins in the mammalian lineage have been derived from a single ancestral β-defensin and that α-defensins arose from β-defensins by a process of gene duplication followed by diversification in response to the pathogens encountered in the particular ecological niche of the organism.[20,21,28] α-defensins were recently believed to be restricted to the primate and glires (rodents and lagomorphs) lineage[20,21,28]; however, more recent analysis of defensins from a broader range of mammals has identified α-defensins in opossum,[17] elephant and hedgehog tenrec,[22] and horse.[23]

In the current work, we used hidden Markov models from the AMPer resource to identify AMPs from bovine. For this work, we considered nucleic acid sequence from the draft genome sequence and expressed sequence tags (ESTs, single-pass sequences of cDNA created from mRNA). Our aim was to discover previously uncharacterized gene-coded bovine AMPs of all classes as well as to test the hypothesis that the bovine genome lacks α-defensins.

METHODS AND MATERIALS

Set of known antimicrobial peptides

We considered the set of known antimicrobial peptides to be derived from the 1135 proteins in Uniprot identified during construction of the AMPer resource (described previously[27]); these are the 980 protein IDs from AMSDb at the University of Trieste (http://www.bbcm.units.it/~tossi/pag1.htm) combined with additional proteins identified by AMPer that were found to have some support for antimicrobial or host defense activity in the literature. These are available at the AMPer web site (http://www.cnbi2.com/cgi-bin/amp.pl).

Creation of AMPer

The AMPer resource has been described previously.[27] Briefly, the 980 Uniprot protein IDs from AMSDb were considered to contain all known AMPs. The Uniprot database entries typically contain one entry for all products of a single gene. The mature peptides (the final active peptides that have antimicrobial activity) and propeptides (peptide regions that are present in the initial translated protein but are removed during post-translational modification and are not signal sequence) are described in Uniprot using a feature table. The mature and propeptide regions were identified from Uniprot annotations and treated separately. The peptides were compared to one another based on sequence similarity and grouped based on this similarity. For each group, a hidden Markov model was created using the HMMER software package. These models were used to iteratively scan Swiss-Prot to identify additional peptides that were not currently identified in the set of AMPs. Uniprot annotations were reviewed for proteins that were identified by AMPer; where annotations suggested antimicrobial activity, these were added to what was considered the set of known AMPs and used to update the AMPer hidden Markov models. Only the 146 hidden Markov models corresponding to mature peptides were used to search bovine sequence.

Bovine genomic and EST sequences

We present here the results in the context of the current versions of the bovine genome and EST set. The bovine genome was downloaded from ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta/Btau200708xx/LinearScaffolds. Preliminary work used the draft bovine genome sequence was obtained from ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta/Btau20050310-freeze/linearScaffolds/. ESTs were obtained from the NCBI resource dbEST resource, downloaded on August 25, 2007 from ftp://ftp.ncbi.nih.gov/blast/db/FASTA/est_others.gz. Bovine ESTs (numbering 1,433,737) were identified as those containing the annotation 'Bos taurus cDNA.' Preliminary work used the same resource downloaded on October 2006. The EST sequences were translated into predicted protein sequences in all six reading frames using software from the BioJava project (http://www.biojava.org).

Prediction of AMPs in ESTs

Predicted protein sequences from ESTs were scanned using the 146 AMPer models for mature peptides using

the HMMER utility, hmmsearch.[25] Regions of sequence matched by a model ('predicted peptides') were examined to identify likely AMPs as follows. Predicted peptides that were less than 25% of the model length were excluded from consideration since they were considered to be unlikely candidates as AMPs and more likely represent conserved protein domains instead. Each matched EST was interpreted as a predicted AMP and assigned an identifier of the form DBEST_AMP_*n* where *n* is an integer. In addition, multiple ESTs may correspond to the same gene product and may differ due to sequencing errors and different lengths of sequencing reads (ESTs are single reads of a cDNA). Therefore, peptide sequences matched by a model were clustered into groups to represent a single predicted AMP based on similarity of the sequences. Specifically, predicted peptides were added to groups where each peptide was at least 90% identical to every other peptide in the group over the length of the peptide (or the smaller peptide if they varied in length). A pairwise BLAST (blastp) comparison was used.[24] Each group of similar predicted peptides was conservatively considered a single antimicrobial peptide. The longest predicted peptide was taken as the representative of each group of similar predicted peptides.

Prediction of AMPs in genomic sequence

The draft genome sequence of bovine was also scanned with the AMPer models of mature peptides using the HMMER utility, hmmsearch,[25] with the total number of sequences specified (using the parameters "-Z 922") to account for matches against the sequence database that spans many files. Genomic sequence contains introns, regions that are not translated into mRNA and hence not present in protein. However, the predicted protein sequence used for searching included intron sequence; therefore, the protein sequences matched by a model will be fragments of a mature peptide corresponding to exons. To account for intron and exon sequence within the genome, predicted peptides were constructed from multiple matching regions within 1000 amino acid positions of each other that cover the length of the AMPer model. Overlap between regions of matches for different models were not allowed. Predicted antimicrobial peptides based on genomic sequence were identified as GENOME_AMP_*n* where *n* is an integer.

Comparison of predicted AMPs to known AMPs

We wished to identify which of the predicted AMPs corresponded to known AMPs. The predicted AMPs were compared to known bovine AMPs using pairwise sequence comparison using the blastp algorithm of the BLAST package.[24] Significance of a match was taken as the *E*-value reported by blastp. Coverage of the two sequences was also calculated to assess the extent of the pairwise match, giving the extent of the matched region in comparison to the length of the known AMP and the AMPer model. Coverage is calculated as the alignment length divided by the maximum possible alignment length (the minimum sequence length between the known AMP sequence and the predicted AMP sequence). For each known bovine AMP, the best matching (lowest *E*-value) AMPs predicted from the EST data set was calculated. A match was considered good if the alignment had minimum 95% identity over minimum 95% coverage. For each AMP predicted from the EST data, the best matching known AMP (of any organism) and best matching known bovine AMP were calculated, taking the matches with lowest *E*-values as the best matches. These are reported on the web pages linked from the summary page at http://www.cnbi2.com/cgi-bin/amp.pl?dbests = hits. The on-line tools allow predicted AMPs to be viewed in the context of the multiple alignment (generated by ClustalW, v 1.83,[29]) containing the predicted AMPs of the model, all known AMPs of the model, the HMM consensus sequence for the model, best-matching AMPs to the predicted AMP and any AMPs predicted from the bovine genomic data that have significant match to the AMP predicted from dbEST data.

Identification of novel AMPs

For each class of AMP, the multiple sequence alignment (generated by ClustalW) was viewed and unique predicted AMPs were identified by eye, by requiring significant differences to be visible in the alignment between the predicted AMP, all other predicted AMPs and the known bovine AMPs. To determine whether the putative novel AMPs had been previously identified, we used the NCBI website (http://www.ncbi.nlm.nih.gov/BLAST/) to search for the sequences in the NCBI nr (nonredundant) databank which contains all nonredundant GenBank CDS translations, Refseq, PDB, SwissProt, PIR, and PRF.

Pairwise comparison of known AMPs to bovine sequence

The set of 1135 known AMPs were used to search for similar sequences in the translated bovine genome and ESTs sequences using blastp of the BLAST package. For genome scanning, the total number of sequences was corrected using the parameter "-z 922." The most significant matches (lowest *E*-values) are reported along with coverage calculated as the alignment length divided by the length of the known AMP. Only matches with *E*-values $<$1e-5 were considered, to restrict the matches to close matches and limit the number of results returned.

Analysis of AMP gene expression

Total RNA was extracted from bovine intestinal tissue and bovine peripheral blood mononuclear cells (PBMC)

Table I
Bovine Primers Used for qRT-PCR

Bovine gene	Accession number[a]	Primer direction	Primer sequence (5′–3′)
GAPDH	BC102589	Forward	AGATGGTGAAGGTCGGAGTG
		Reverse	GATCTCGCTCCTGGAAGATG
β-Actin	AF191490	Forward	CTAGGCACCAGGGCGTAATG
		Reverse	CCACACGGAGCTCGTTGTAG
DBEST_AMP_397	BF775065	Forward	TCGTGGTGGAGTTCAAATCA
		Reverse	GCTTGGAAGGCACTGGTACT
DBEST_AMP_416	BI537181	Forward	GGATTGGTGGAGGAAATCTG
		Reverse	GAATGGGCTGGTGAAACAGT

[a]Accession numbers are from NCBI (http://www.ncbi.nlm.nih.gov).

as described previously[30] and RNA was isolated using an RNeasy Mini Kit (Qiagen, Ontario, Canada). The intestinal samples were collected both prior to and 4 h after challenge with *S. typhimurium* using the infection model developed by Coombes *et al.*[31] Isolated RNA samples were eluted and stored in RNase-free water (Ambion, Austin, Texas) at −80°C until further use. The RNA concentration, integrity and purity were assessed determining the OD260/280 ratio with a BioPhotometer (Eppendorf, Hamburg Germany) in addition to analysis on a 1% agarose gel and Bioanalyzer (Agilent, USA).

Quantitative real-time PCR (qRT-PCR) was performed using Invitrogen's SuperScript™ III Platinum two-step qRT-PCR kit with SYBR-Green on the ABI 7300 Real Time PCR System (Applied Biosystems, Foster City, CA) as described previously.[32] Endogenous house keeping genes, GAPDH and β-actin, were used for normalization and determination of fold changes of the respective AMPs using the comparative threshold cycle method.[33] The qRT-PCR products were run on a 2% agarose gel to verify the presence of gene products.

All primers used for qRT-PCR were designed using Primer3 v.0.3.0,[34] except β-actin that was designed earlier.[35] The primers are listed in Table I.

Informatics

All calculations were performed on a Linux or Mac OS X environment using custom Java, Python, Perl or BASH code. Data were stored in a MySQL database for manipulation and presentation via Perl CGI scripts on an Apache web server running on a Linux server at http://www.cnbi2.com.

RESULTS

Identification of host defense peptides

We used the AMPer models of mature peptides to identify known and potentially novel antimicrobial sequences of bovine using expressed sequence tags (from NCBI dbEST resource, http://www.ncbi.nlm.nih.gov/dbEST/,[36]) and genomic sequence (from the Baylor College of Medicine Human Genome Sequencing Center, http://www.hgsc.bcm.tmc.edu/projects/bovine). These were translated into all six reading frames and scanned with each of the 146 AMP models. Results are presented here using the dbEST resource containing 1,433,737 bovine ESTs (downloaded on August 25, 2007). The models of mature peptides produced 5628 matches with an *E*-value <10, consisting of 4591 unique ESTs. Of these, 2228 had an *E*-value <1 and cover at least 25% of the length of the model.

We identified unique sequences by clustering the matched protein using an all-vs-all comparison: each matched protein was compared to every other matched protein with blastp.[24] Where predicted peptides were at least 90% identical, we conservatively considered these to be the same antimicrobial peptide (at the risk of grouping together closely related peptides that are in fact distinct). By repeating this pairwise comparison, a total of 278 potential peptides were identified. From these 278 peptides, we selected those that were matched at high statistical significance (an HMM *E*-value $\leq$ 1e-5), resulting in the 124 potential peptides shown in Table II.

We mapped the 34 known bovine AMPs using the full protein sequence (Supplementary Table S1) to all predicted protein sequences from the ESTs using pairwise comparison (the blastp algorithm[24]) to identify the most likely ESTs corresponding to the bovine AMP. We similarly mapped the 34 known bovine AMPs to those predicted protein sequences identified by AMPer as containing an AMP (Table III). Since we expect these sequences to differ only due to artifacts such as sequencing errors, we called a match significant where there was at least 95% sequence identity between the known bovine AMP and the other sequence, and where the length of the matching region between two sequences (reported by blastp) was within 95% of the shorter sequence (this was meant to allow for the untranslated regions of the mRNA). A total of 27 known bovine AMPs had significant matches to ESTs. Since some AMPs are subsequences of other AMPs and ESTs may also be significantly shorter than the cDNA from which they are sequenced, it

Table II
Numbers of Predicted Antimicrobial Peptides by AMPer Model

AMPer Model	Peptide Families	Number of AMPs
17	15 kDa protein (rabbit)	1
66	Apolipoprotein A-II (multispecies, including bovine, horse, primate, mouse, rat)	12
106	Bactenecin (bovine, goat, sheep)	1
90	β-defensins (bovine β-defensins-7, -8, -9)	4
145	β-defensins (multispecies BD-1, -2)	7
144	β-defensin, LAP, TAP, Spheniscin (bovine, goat, sheep, penguin)	6
35	β-defensin, Circulin-B (chinchilla, mouse, chimpanzee)	1
117	BPI, LBP, (bactericidal permeability-increasing protein, lipopolysaccharide-binding protein) (human, bovine, rabbit, rat, mouse)	29
116	Cathelicidin (multispecies, including human, mouse, goat, rabbit)	1
87	Cathelicidin (horse and pig)	9
18	Cathelin-related (bovine, sheep)	1
133	Cysteine-rich antifungal protein (multispecies, mostly plant)	1
92	Eosinophil granule major basic protein (multispecies, including human, mouse, rat)	7
13	Granulysin, NK-lysin (human, pig)	5
27	Hemolin (insect)	2
64	Hepcidin (multispecies, human, pig, fish)	1
8	Uperin, Histone H1 (fish, amphibian)	5
12	Histone H2A (Hipposin from fish)	19
38	Histone H2A (Buforin from toad)	8
24	Myeloid antibacterial peptide (Bovine BMAP-27, pig PMAP-36)	2
95	Penaeidin, Liver-expressed AMP (shrimp, mammal)	1
39	Sperm-associated antigen 11 (rat, mouse)	1

An *E*-value threshold of 1e-5 was used to determine significance of an HMM match. Additional information is given to better identify the type of peptide as the "Peptide Families" column. These peptide families correspond to AMPer models and are shown along with an indication of the species represented by the AMPer model.

is difficult to determine uniquely which bovine AMPs were identified where multiple known bovine AMP sequences mapped to the same EST sequence. For four bovine AMPs the best matching EST was not unique (one other known bovine AMP also matched that EST most significantly of all ESTs). These are indicated by a '(2)' on four entries in Table III. Similarly, a total of 27 are also found to have significant matches to AMPs predicted by AMPer, though the list of known AMPs with no clearly matching predicted AMP are slightly different than those known AMPs with no clearly matching ESTs (24 had good matches to both ESTs and AMPer predictions).

Selection of predicted AMPs for confirmation

We manually examined the sequences of these predicted AMPs to identify peptides of interest for laboratory follow-up. Using on-line tools that we developed, we examined multiple alignments of these predicted AMPs alongside the following: the most similar known bovine AMP, the most similar AMP from any species (if different from bovine) and the peptides that were used to construct the AMPer model. (These are available from links on the bovine analysis pages at the AMPer site.) We chose two predicted AMPs for follow-up that appeared to be novel and belong to the cathelicidin family. Two ESTs corresponding to these predicted AMPs were identified for laboratory analysis of changes in expression due to infection as discussed below. The first predicted AMP that we sought to confirm was DBEST_AMP_248, matched by model 17. This peptide sequence was compared to all proteins in Uniprot (Swiss-prot and TrEMBL) using the on-line BLAST utility at http://www.expasy.org/tools/blast. Since we began this work, an entry containing DBEST_AMP_248 has been deposited in TrEMBL as A5PJH7_BOVIN (discussed below) based only on cDNA sequencing. The most similar peptide to DBEST_AMP_248 is an antimicrobial peptide found in rabbit, P15B_RABIT, designated as "15 kDa protein"[37] with 55% sequence identity and 99.2% coverage. The most similar known bovine AMP, Bactenecin-7 (BCTN7_BOVIN, now called CTHL3_BOVIN in the current version of Uniprot) has only 33% sequence identity and 95.8% coverage. On the basis of earlier data, in place of DBEST_AMP_248, we examined the predicted AMP, DBEST_AMP_397, and EST sequence gi|12122965|gb|BF775065.1 (a slightly shorter sequence within the same cluster of predicted AMP sequences as DBEST_AMP_248). As shown in Figure 1, the translated EST sequence (BF775065.1) shows good alignment with the 15 kDa protein sequence and poorer alignment with the bovine peptide BCTN7_BOVIN. In Figure 1, the underlined regions indicate the region of mature peptide corresponding to the active antimicrobial peptide.

Table III
Identification of Known Bovine Host Defense Peptides in dbEST Sequences

	Matched ESTs by BLAST				AMPer Predicted AMPs			
Known Bovine AMP	Matched EST[a]	% Ident	Match Length (coverage %)	Blast *E*-value	Matched Predicted AMP[a]	% Ident	Match Length (coverage %)	Blast *E*-value
APOA2_BOVIN	gi|75805025|gb|DT855734.1	100	100 (100.0)	7.20E-050	DBEST_AMP_1858	100	77 (100.0 %)	6.70E-041
BCTN1_BOVIN	gi|119554907|gb|EH155902.1	100	155 (100.0)	1.10E-086	DBEST_AMP_255	100	101 (100.0 %)	8.30E-058
BCTN5_BOVIN	gi|154772689|gb|EV792452.1	100	176 (100.0)	6.00E-102	DBEST_AMP_249	100	101 (100.0 %)	1.50E-057
BCTN7_BOVIN	gi|119563722|gb|EH164717.1	100	190 (100.0)	8.00E-111	DBEST_AMP_304	100	102 (100.0 %)	1.10E-057
BD01_BOVIN	gi|17892782|gb|BM257183.1	100	38 (100.0)	3.50E-019	DBEST_AMP_1047	100	36 (100.0 %)	7.70E-021
BD02_BOVIN	gi|7049236|gb|AW479130.1	100	40 (100.0)	5.30E-020	DBEST_AMP_1428	100	38 (100.0 %)	1.10E-021
BD03_BOVIN	gi|7049236|gb|AW479130.1 (2)	100	57 (100.0)	2.50E-029	DBEST_AMP_1428 (2)	100	38 (100.0 %)	7.40E-022
BD04_BOVIN	gi|154397167|gb|EV640446.1	100	63 (100.0)	6.20E-033	DBEST_AMP_901	100	36 (100.0 %)	3.00E-021
BD05_BOVIN	gi|17037442|gb|BM106372.1	100	64 (100.0)	2.50E-034	DBEST_AMP_860	100	37 (100.0 %)	8.90E-023
BD06_BOVIN	gi|119558511|gb|EH159506.1	100	40 (95.2)	9.00E-020	DBEST_AMP_2132	100	38 (100.0 %)	6.40E-022
BD07_BOVIN	gi|119561789|gb|EH162784.1	100	40 (100.0)	2.00E-019	DBEST_AMP_1576	100	38 (100.0 %)	1.50E-021
BD08_BOVIN	gi|119564671|gb|EH165666.1	100	38 (100.0)	3.90E-018	DBEST_AMP_308	100	38 (100.0 %)	1.50E-021
BD09_BOVIN	gi|119564671|gb|EH165666.1 (2)	98.18	55 (100.0)	9.20E-027	DBEST_AMP_308 (2)	97.37	38 (100.0 %)	5.10E-021
BD10_BOVIN	gi|42731051|gb|CK778738.1	98.39	62 (100.0)	1.00E-030	DBEST_AMP_139	100	36 (100.0 %)	1.50E-020
BD11_BOVIN	gi|119531287|gb|EH137278.1	100	60 (100.0)	4.60E-031	DBEST_AMP_209	100	37 (100.0 %)	6.70E-022
BD12_BOVIN	gi|74502222|gb|DT722637.1	97.37	38 (100.0)	6.60E-018	DBEST_AMP_1461	97.3	37 (100.0 %)	1.30E-020
BD13_BOVIN	gi|74502222|gb|DT722637.1 (2)	97.62	42 (100.0)	5.30E-020	DBEST_AMP_1461 (2)	97.3	37 (100.0 %)	1.20E-020
BDC7_BOVIN	gi|119531287|gb|EH137278.1 (2)	94.34	53 (100.0)	1.10E-024	DBEST_AMP_209 (2)	91.89	37 (100.0 %)	7.80E-020
BMA27_BOVIN	gi|120572158|gb|EH378295.1	99.22	128 (81.0)	1.80E-068	DBEST_AMP_383	98.98	98 (100.0 %)	6.80E-055
BMA28_BOVIN	gi|119558428|gb|EH159423.1	100	159 (100.0)	2.30E-087	DBEST_AMP_274	100	113 (100.0 %)	1.10E-063
BMA34_BOVIN	gi|61753367|emb|CR452179.2	100	165 (100.0)	1.70E-091	DBEST_AMP_478	100	129 (100.0 %)	8.30E-075
BPI_BOVIN	gi|119650848|gb|EH179456.1	99.61	254 (52.7)	2.00E-141	DBEST_AMP_1174	100	250 (100.0 %)	4.00E-146
CALT_BOVIN	gi|86366255|gb|DY165694.1	100	80 (100.0)	2.10E-040	DBEST_AMP_186	30.43	23 (28.8 %)	3.7
CAS2_BOVIN	gi|70828695|gb|DR712392.1	100	222 (100.0)	1.00E-124	DBEST_AMP_332	22.86	35 (31.2 %)	4.8
CCKN_BOVIN	gi|60967497|gb|DN524024.1	100	58 (100.0)	4.00E-027	DBEST_AMP_358	40.91	22 (21.8 %)	0.03
CMGA_BOVIN	gi|119653666|gb|EH182274.1	99.62	266 (59.2)	5.00E-152	–	–	–	–
EAP_BOVIN	gi|75771874|gb|DT822941.1	100	64 (100.0)	3.60E-033	DBEST_AMP_1091	100	36 (100.0 %)	3.20E-020
INDC_BOVIN	gi|119556821|gb|EH157816.1	100	144 (100.0)	2.00E-082	DBEST_AMP_542	100	100 (100.0 %)	2.70E-056
LAP_BOVIN	gi|154466011|gb|EV693095.1	100	64 (100.0)	6.80E-032	DBEST_AMP_746	100	38 (100.0 %)	1.10E-020
LBP_BOVIN	gi|82642070|gb|DV789175.1	93.44	380 (79.0)	0	DBEST_AMP_1816	93.44	380 (99.7 %)	0
LEAP2_BOVIN	gi|154538028|gb|EV742363.1	100	77 (100.0)	2.40E-039	DBEST_AMP_729	100	40 (100.0 %)	3.00E-021
PENK_BOVIN	gi|119686348|gb|EH206269.1	100	246 (93.5)	2.00E-144	DBEST_AMP_978	38.89	18 (60.0 %)	9.8
SCG1_BOVIN	gi|82827500|gb|DV893271.1	99.66	297 (46.0)	1.00E-177	DBEST_AMP_549	100	13 (100.0 %)	0
TAP_BOVIN	gi|154464382|gb|EV691466.1	100	64 (100.0)	1.10E-032	DBEST_AMP_753	100	36 (100.0 %)	9.20E-020

Known bovine AMPs (left column) were mapped to EST seqeuences based on pairwise similarity using BLAST alone (middle columns) and to AMPs predicted by AMPer hidden Markov models (right columns).
[a]a '(2)' indicates that this is the second entry with the same identifier. The entries consisting of a dash (-) indicate no match was found.

```
P15A_RABIT     --------MAGVWKVLVVLVGLAVVACAIPRRRLRYEEVVAQALQFYNEGQQGQPLFRLLEATPPPS-LNSK-SRIPLNFRIKETVCIFTLDRQ
P15B_RABIT     --------MAGVWKVLVVLVGLAVVACAIPHRRLRYEEVVAQALQFYNEGQQGQPLFRLLEATPPPS-LNSK-SRIPLNFRIKETVCIFTLDRQ
BCTN7_BOVIN    METQRASLSLGRWSLWLLLLG--LVLPSASAQALSYREAVLRAVDRINERSSEANLYRLLELDPPPKDVEDRGARKPTSFTVKETVCPRTSPQP
BF775065.1     --------MAGAWKALVLVAGLAAVAC-VAQRGLSYEEIVTQALKFFNQGRRGQRIFGLLESTPPPPDLNS--TTIPLNFRIKETVCFLLWYRR
HMM_consensus  ----------------------------IPRRRLRYEEVVAQALQFYNEGQQGQPLFRLLEATPPPS-LNSK-SRIPLNFRIKETVCIFTLDRQ
DBEST_AMP_397  ----------------------------VAQRGLSYEEIVTQALKFFNQGRRGQRIFGLLESTPPPPDLNS--TTIPLNFRIKETVCFLLWYRR

P15A_RABIT     -PGNCAFREGGEERICRGAFVRRRWVRALTLRCDRDQRRQPEFPRVTRPAGPTA-----------------------------------------
P15B_RABIT     -PGNCAFREGGEERICRGAFVRRRRVRALTLRCDRDQRRQPEFPRVTRPAGPTA-----------------------------------------
BCTN7_BOVIN    -PEQCDFKENGLVKQCVGTITLDQSDDLFDLNCNELQSVRRIRPRPPRLPRPRPRPLPFPRPGPRPIPRPLPFPRPGPRPIPRPLPFPRPGPRPIPRPL
BF775065.1     RPRQCPFREGGEERNCTGSFFMLRQLRLLSLNCVPDRELE-------------------------------------------------------
HMM_consensus  -PGNCAFREGGEERICRGAFVRRRRVRALTLRCDRDQRRQPEFPRVTRPAGPTA-----------------------------------------
DBEST_AMP_397  RPRQCPFREGGEERNCTGSFFMLRQLRLLSLNC-----------------------------------------------------------
```

Figure 1

Multiple alignment of predicted host defense peptide DBEST_AMP_397. The predicted peptide DBEST_AMP_397 is shown aligned to all peptides in the AMPer cluster, the most similar AMP (P15B_RABIT), the most similar bovine AMP (BCTN7_BOVIN), and the EST that DBEST_AMP_297 was derived from (BF775065.1). Underlined sequence indicates the position of mature peptides within the proteins. BCTN7_BOVIN shows a poor alignment in residues and position of mature peptide compared to DBEST_AMP_397 and the rabbit AMPs. The consensus sequence of AMPer model 17 is also shown (HMM_consensus). BCTN7_BOVIN is now CTHL3_BOVIN in the current Uniprot database.

The second predicted novel AMP that we sought to confirm was identified from EST gi|15378291|gb|BI537181.1, as predicted peptide DBEST_AMP_416, matched by model 87. This predicted AMP matches a short region of the sequence for the known bovine AMP, Bactenecin-5 (BCTN5_BOVIN, now CTHL2_BOVIN in the current Uniprot). Examination of the translated EST sequence that was recognized by the AMPer model and produced DBEST_AMP_416 shows that it codes for a similar protein with differences near the N-terminus. The predicted sequence is shown in Figure 2, along with the proteins that were used to construct the AMPer model that recognized this peptide, and the closest matching known bovine AMP (BCTN5_BOVIN). We compared the EST sequence (232 nucleic acids) for this predicted AMP to the current bovine genome in Ensembl (http://www.ensembl.org) and did not find a significant match except to a short region of the genomic sequence for Bactenecin-5: 52 positions on chromosome 22 (49,818,207–49,818,362) matched the EST from positions 27–78. This region overlaps with Bactenecin-5 exon 4 (ENSBTAE00000175540) 49,818,093–49,818,356 and extends 6 positions into intron 3–4. Neighboring DNA regions on the chromosome did not contain additional flanking EST sequence that would be expected if the sequences were separated in the genome due to introns. However, the EST sequence matched a longer region of 77 nucleic acids (EST region 17–90) on a sequence contig from whole genome shotgun (gi|112113766|gb|AAFC03064548.1| Ctg60.CH240-439A19). This suggests that the predicted AMP from DBEST_AMP_416 is from a novel gene that has not yet been incorporated into the genome assembly. However, the

```
O62840_HORSE   METQRNTRCLGRWSPLLLLLGLVIPPATTQALSYKEAVLRAVDGLNQRSSDENLYRLLELDPLPKGDKDSDTPKPVSFMVKETVCPRIMK
O62841_HORSE   METQRDSCSLGRWSLLLLLLGLVIPLATTQTLSYKEAVLRAVDGLNQRSSDENLYRLLELDPLPKEDEDPDTPKPVSFTVKETVCPRTTQ
O62842_HORSE   METQRNTRCLGRWSPLLLLLGLVIPPATTQALSYKEAVLRAVDGLNQRSSDENLYRLLELDPLPKGDKDSDTPKPVSFMVKETVCPRIMK
ICTL_PIG       ----------------------------Q-LRYREAVLRAVDRLNEQSSEANLYRLLELDQPPKADEDPGTPKPVSFTVKETVCPRPTR
BCTN5_BOVIN    METQRASLSLGRCSLWLLLLGLVLPSASAQALSYREAVLRAVDQFNERSSEANLYRLLELDPTPNDDLDPGTRKPVSFRVKETDCPRTSQ
BI537181.1     --------------------------------------------------------------------------------------------
HMM_consensus  ----------------------------ALSYKEAVLRAVDGLNQRSSDENLYRLLELDPLPKEDEDSDTPKPVSFTVKETVCPRITK
DBEST_AMP_416  --------------------------------------------------------------------------------------------

O62840_HORSE   QTPEQCDFKENGLVKQCVGTVILGPVKDHFDVSCGEPQRVKRFGRLAKSFLRMRILLP--------------RRKILLAS---------
O62841_HORSE   QPLEECDFKENGLVKQCVGTVVLDPAKDYFDISCDKPQPIKRRHWFPLSFQEFLEQLRR------------FRDQLPFP----------
O62842_HORSE   QTPEQCDFKENGLVKQCVGTVILDPVKDYFDASCDEPQRVKRFHSVGSLIQRHQQMIRDKSEATRHGIRIITRPKLLLAS---------
ICTL_PIG       QPPELCDFKE----KQCVGTVTLNPSIHSLDISCNEIQSV----------------------------------------------------
BCTN5_BOVIN    QPLEQCDFKENGLVKQCVGTVTLDPSNDQFDINCNELQSVRFRPPIRRPPIRPPFYPPFRPPIRPPIFPPIRPPFRPPLGPFPGRR---
BI537181.1     --------KENGLVKQCVGTVTLDPSNDQFDINCNESVRFP--PPIRPPPIRPPFFPPIGPPIRPPIGPPIRPPIPPPIRPPFGPP---
HMM_consensus  QPPEQCDFKENGLVKQCVGTVILDPVKDYFDISCDEPQRVKRFHSLAKSFQRFRELLRR------------FRDKLLLAS---------
DBEST_AMP_416  --------KENGLVKQCVGTVTLDPSNDQFDINCNE----------------------------------------------------
```

Figure 2

Multiple alignment of predicted host defense peptide DBEST_AMP_416. The predicted peptide DBEST_AMP_416 is shown aligned to all peptides in the AMPer cluster, the most similar bovine AMP (BCTN5_BOVIN), and the EST that DBEST_AMP_416 was derived from (BI537181.1). Underlined sequence indicates the position of mature peptides within the proteins. The consensus sequence of AMPer model 87 is also shown (HMM_consensus). BCTN5_BOVIN is now CTHL2_BOVIN in the current Uniprot database.

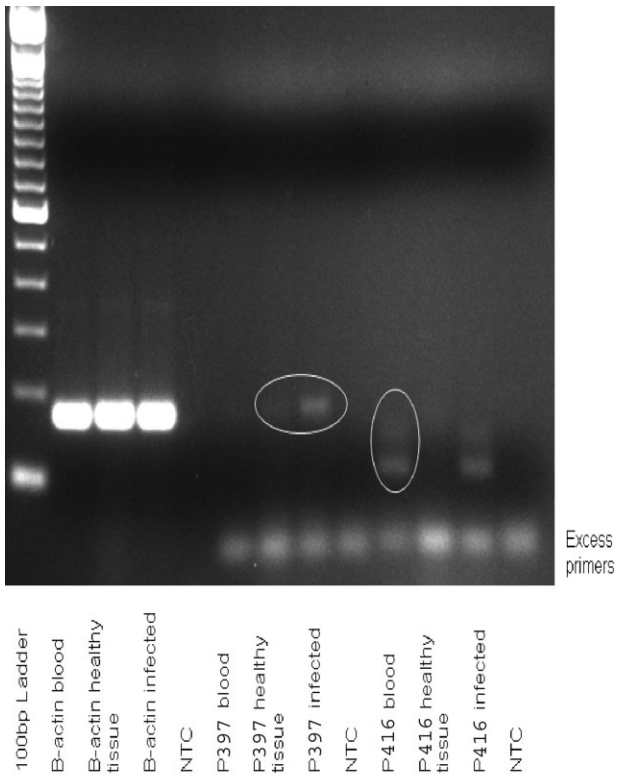

Figure 3

Gel image of qRT-PCR for putative AMPs in blood and tissue. The DBEST_AMP_397 (P397) and DBEST_AMP_416 (P416) products are visible. B-actin lanes are positive control lanes and NTC lanes are "no template" controls.

sequence was originally found in expressed sequence; therefore it appears to be a true gene rather than a pseudogene, despite not being able to identify the full gene sequence in the genome.

Analysis of predicted novel AMP gene expression

We designed primers to detect and amplify RNA corresponding to these two putative AMPs along with and two housekeeping genes (GAPDH and β-actin) that serve as positive controls. Quantitative real-time PCR (qRT-PCR) was performed using these primers on total RNA derived from PBMC, and tissue collected from the bovine small intestine (ileum). The intestinal tissue was sampled both prior to and 4 h after challenge with *S. typhimurium* with the *S. typhimurium* infection performed as described previously by Coombes *et al.*[31] Initial qRT-PCR products were run on agarose gel and showed faint bands (see Fig. 3). The qRT-PCR products were re-amplified using a 30 cycle Taq-man PCR protocol and visualized on gel (see Fig. 4). The DBEST_AMP_397 product is clearly visible and upregulated in response to bacterial infection in intestinal tissue. However, the DBEST_AMP_416 product cannot be distinguished from negative control lanes in Figure 4 and the presence of two bands rather than the expected single band in Figure 3 suggests the putative AMP product for DBEST_AMP_416 was not found.

Absence of α-defensins

Notably absent from Table II are any of the α-defensin peptide families (often described as simply "defensins"). There are several models in AMPer for mature peptides of this type including models 53, 98, 105, and 146 as well as subclasses such as cryptdins (model 75). For example, AMPer model 146 is built from a set of 45 α-defensin peptides from 42 different Swiss-Prot proteins taken from eight mammalian species. The model matches these 45 peptides with high statistical significance (*E*-values are all less than 1e-10 with only two greater than 1e-20; see AMPer web site). However, the most significant match in the bovine EST sequences is to gi|82672759|gb|DV812566.1 with an *E*-value of 3.6e-4.

The analysis described here tolerates the presence of introns and will combine neighboring regions identified

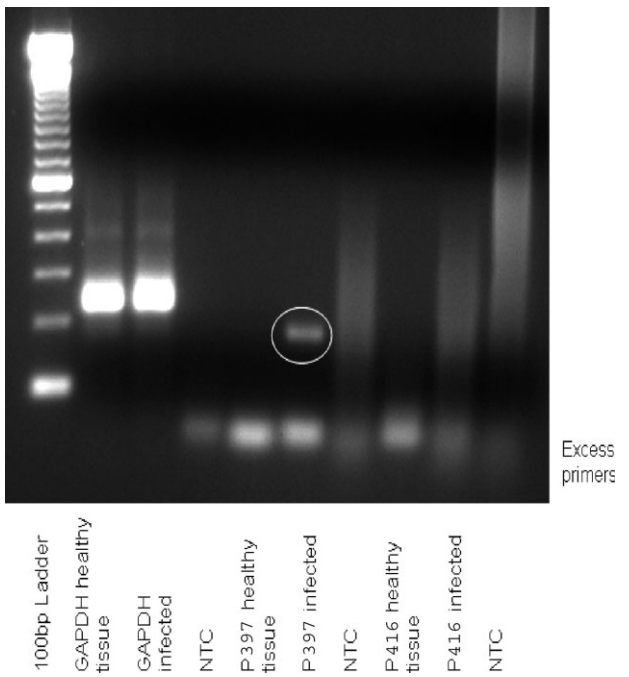

Figure 4

Gel image of putative AMPs following Taq-man re-amplification. The DBEST_AMP_397 (P397) product is clearly visible in the infected tissue but not healthy tissue. While a difference is observed for DBEST_AMP_416 (P416) between healthy and infected intestinal tissue, the P416 lane does not produce a useful band and is not distinguishable from NTC. GAPDH are positive control lanes and NTC lanes are "no template" controls.

by an HMM model to cover the length of the model and report the resulting peptide with a single ID. An example of an AMP containing introns that is correctly identified by AMPer is BD07_BOVIN. This BD07_BOVIN contains one intron of 1460 nucleotides (487 amino acids when translated) and is identified from EST sequence by DBEST_EST_292 (model 90). The predicted AMP based on genomic sequence, GENOME_AMP_169, is identical in sequence to BD07_BOVIN but short by 2 amino acids (length of 38 vs. 40) and produces an HMM *E*-value of 4e-23 (see web resources). In contrast, the most significant *E*-value for (α-defensin) model 146 against bovine genomic data is 4e-10 but the coverage of the model is low at only 69% and the predicted AMP sequence lacks the characteristic six-cysteine motif (see Supplementary Table S2 for predicted AMPs based on genomic data with *E*-values less than 1e-5).

DISCUSSION

Host defense peptides of the innate immune system are important components for control of infection. Historically, host defense peptides have been described as antimicrobial peptides (AMPs); however, the important role of modulation of the innate immune response has come to the fore recently.[38] Natural host defense peptides are considered to be lead compounds in the search for agents that beneficially modulate inflammatory responses both directed against a pathogen and to counter detrimental immune responses such as those involved in sepsis. The importance of these peptides in host defense and as the basis of possible novel therapeutics indicates the need for information about the numbers and types that are present to gain further understanding of their roles in innate immunity. In order to identify potentially novel host defense peptides, we used the hidden Markov models constructed for the AMPer resource to scan bovine expressed sequence tags and genomic sequence. The AMPer models represent groups of mature peptides as well as propeptides that are products of the parental prepropeptides due to processing after protein translation; there are 146 models of mature peptides and 40 models of propeptides representing classes and subclasses of peptides such as defensins and cathelicidins. In this study, we used the models for mature peptides only.

We are primarily concerned with identifying mature peptides for the purpose of structure-activity analysis. Therefore, we primarily relied upon EST sequences since they do not have the added complication of introns in predicted protein sequence. Since the same gene may lead to many ESTs, we sought to identify those unique sequences corresponding to a gene by grouping the predicted peptides based on sequence similarity. We chose a conservative threshold since we are interested in identifying novel AMPs, and are less interested in identifying close homologues of known bovine AMPs; in addition, EST sequencing is a single-pass process with sequencing errors of up to a few percent[36] so true matches are expected to not match perfectly. We considered EST sequence where the matched regions of these ESTs were more than 90% identical over the region of the pairwise match to belong to the same host defense peptide. This threshold yielded a total of 278 potential peptides of varying statistical significance. The HMM *E*-value represents the number of false positive matches expected at a given threshold; using an HMM *E*-value threshold of 1e-5 (ie. 1e-5 expected false positives for each of the 146 models) yields a prediction of up to 124 AMPs, including 32 matches to histone (from which the AMP buforin is derived[39]). There are 92 non-histone AMPs, a number that is feasible to review manually (Table II). As well, this *E*-value threshold is large enough that sequences belonging to more distant homologues would not be discarded, but at the risk of including peptides that are only distantly related to and not actually AMPs.

To determine which of these predicted AMPs correspond to known bovine AMPs, we compared the sequences using sequence similarity (blastp[24]) to find predicted peptide from both ESTs and peptide identified by AMPer models. Of the 34 known bovine AMPs (full length proteins, Table III), a total of 27 known bovine AMPs have significant matches to ESTs. As well, 27 known bovine AMPs have significant matches to AMPs predicted by AMPer. The known AMPs with no significant match to ESTs are slightly different than those known AMPs with no significant match to a peptide identified by AMPer. Several known bovine AMPs were not identified in the EST data presumably because they were not expressed in the tissues that were sampled for mRNA and used to construct the EST libraries. Of the three known AMPs (CALT_BOVIN, CAS2_BOVIN and CCKN_BOVIN) that appear to have been represented in the EST data set but missed by the AMPer search, only CCKN_BOVIN seems to have been missed due to inadequacy of the AMPer model: CALT_BOVIN and CAS2_BOVIN did not contribute mature peptides that were used in constructing AMPer models (for details of the AMPer construction algorithm see Ref. 27). Considering that a total of 95 non-histone AMPs were predicted and up to 27 known AMPs were found to have significant matching ESTs, there are up to 68 potentially novel AMPs identified in the EST set by the AMPer models at the threshold values we used.

We chose two predicted AMPs for follow-up that appear to be novel and belong to the cathelicidin family, a group of peptides of special interest to us. We chose two ESTs corresponding to these predicted AMPs for RT-PCR analysis of gene transcription as well as changes in gene expression following infection. (Note that since this work began, significantly more bovine sequence has become available and slightly different ESTs might have been chosen based on current data.) We demonstrated

that one of these, DBEST_AMP_397, is expressed in response to infection. When compared to all proteins found in Uniprot (both Swiss-prot and TrEMBL), this predicted peptide is most similar to the '15 kDa protein' AMP found in rabbit and of a class of AMP not previously described for bovine. Since our work began on AMPs in bovine, this peptide (DBEST_AMP_397) has been predicted based on sequencing of cDNA from a thymus sample and submitted to the TrEMBL database of Uniprot as A5PJH7_BOVIN (http://www.expasy.org/uniprot/A5PJH7) by the Mammalian Gene Collection project (http://mgc.nci.nih.gov/). Here, we report that we have independently identified this peptide using the AMPer resource and demonstrated that it is upregulated in the small intestine in response to infection. We did not find the second predicted AMP we attempted to confirm in the tissues we sampled, and we did not locate the genome location of its sequence in the current genome assembly. However, the sequence was found in whole genome shotgun sequence that was not incorporated into the current bovine assembly. Since it was originally found in expressed sequence, it appears to be a true gene rather than a pseudogene.

We did not identify any AMP sequences for α-defensins in bovine EST sequence, strongly suggesting that α-defensins are not present in the EST dataset we used. In addition, when we scanned translated genomic sequence we also did not find evidence for α-defensins. The analysis we performed did account for the presence of introns in constructing AMP predictions: For example, β-defensins were found reliably despite the presence of intron sequence. Since we cannot account for the lack of α-defensins identified using the AMPer models due to any technical deficiencies (and additionally we cannot find reference to any bovine α-defensins in the literature), we conclude that these results indicate that the bovine genome lack this important class of host defense peptide. Other mammalian species such as mouse are known to lack neutrophil-derived α-defensins.[40] Previous reports have speculated that α-defensins are found only in the primate and glires (rodents and lagomorphs) lineage,[20,21,28] while more recent reports have identified α-defensins in a wide range of diverse mammals such as opossum,[17] elephant and hedgehog tenrec,[22] and the horse,[23] a close evolutionary cousin to bovine. This suggests that the bovine genome has lost α-defensins from an ancestor through evolution, rather than being on a lineage where α-defensins were never present.

CONCLUSIONS

We have used the HMM models from the AMPer resource to scan the draft bovine genome and bovine expressed sequence tags from the dbEST data set. To additionally describe the peptides, we have identified the most similar known AMP for each predicted peptide. The AMPer models identified 27 of the 34 known bovine antimicrobial peptides. An additional 68 potential peptides were identified that appear to be previously unidentified AMPs, for a total of 102 AMPs. We sought to experimentally verify two of these that belong to the cathelicidin family. One of these, DBEST_AMP_397, was clearly identified in qRT-PCR product and was found to be upregulated in bovine intestinal tissue following challenge with *S. typhimurium*. One other putative AMP (DBEST_AMP_416) was not confirmed in blood mononuclear cells and small intestine. In addition to the identification of unrecognized AMPs our results suggest that bovine lacks α-defensins.

The novel antimicrobial peptide, DBEST_AMP_397, was also predicted by the Mammalian Gene Collection project as part of an effort to provide full-length clones to investigators for a limited number of organisms (human, rat, mouse and bovine). This serves to confirm the utility of the AMPer approach to identifying novel AMPs: by examining the large resource of low quality EST sequence, we have identified a novel peptide that was added to the major sequence databases only recently, after a high quality cDNA sequencing project. This suggests that a large number of additional peptides might be identified from publicly available data that will not be added to major databases for some time. These results indicate the effectiveness of *in silico* screening with software resources such as AMPer that are tailored to specific interests of the community, in this case, investigators examining peptides of the innate immune system. The hidden Markov models used by AMPer are freely available to investigators and straightforward to use (see http://www.cnbi2.com/cgi-bin/amp.pl). Future work on AMPer will include automation of the steps involved in the study described here, and its application to larger numbers of organisms.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the Canadian Institutes of Health Research (CIHR) and from Genome BC for the Pathogenomics of Innate Immunity research program. CDF is supported by a Doctoral Research Award from CIHR. KH received a CIHR postdoctoral fellowship. REWH was the recipient of a Canada Research Chair.

WEB RESOURCES

AMPer: http://www.cnbi2.com/cgi-bin/amp.pl
AMSDb at University of Trieste: http://www.bbcm.units.it/~tossi/pag1.htm
Baylor College of Medicine Human Genome Sequencing Center, bovine genome: http://www.hgsc.bcm.tmc.edu/projects/bovine
BioJava: http://www.biojava.org
NCBI dbEST: http://www.ncbi.nlm.nih.gov/dbEST/

REFERENCES

1. Sima P, Trebichavsky I, Sigler K. Mammalian antibiotic peptides. Folia Microbiol 2003;48:123–137.
2. Sima P, Trebichavsky I, Sigler K. Non-mammalian vertebrate antibiotic peptides. Folia Microbiol 2003;48:709–724.
3. Simmaco M, Mignogna G, Barra D. Antimicrobial peptides from amphibian skin: what do they tell us? Biopolymers 1998;47:435–450.
4. Khush RS, Leulier F, Lemaitre B. Drosophila immunity: two paths to NF-kappaB. Trends Immunol 2001;22:260–264.
5. Hancock RE, Lehrer R. Cationic peptides: a new source of antibiotics. Trends Biotechnol 1998;16:82–88.
6. Hancock RE, Chapple DS. Peptide antibiotics. Antimicrob Agents Chemother 1999;43:1317–1323.
7. Hancock RE. Concerns regarding resistance to self-proteins. Microbiology 2003;149:3343–3344.
8. Marshall SH, Arenas G. Antimicrobial peptides: A natural alternative to chemical antibiotics and a potential for applied biotechnology. Electron J Biotechnol 2003;6:271–284.
9. Bechinger B. Structure and function of channel-forming peptides: magainins, cecropins, melittin and alamethicin. J Membr Biol 1997; 156:197–211.
10. Bechinger B. The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy. Biochim Biophys Acta 1999;1462:157–183.
11. Blondelle SE, Lohner K, Aguilar MI. Lipid-induced conformation and lipid-binding properties of cytolytic and antimicrobial peptides: determination and biological specificity. Biochim Biophys Acta 1999; 1462:89–108.
12. Shai Y. Mechanism of the binding, insertion and destabilization of phospholipids bilayer membranes by α-helical antimicrobial and cell non-selective membrane-lytic peptides. Biochim Biophys Acta 1999;1462:55–70.
13. Epand RM, Vogel HJ. Diversity of antimicrobial peptides and their mechanisms of action. Biochimica Biophysica Acta 1999;1462:11–28.
14. Brogden KA. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat Rev Microbiol 2005;3:238–250.
15. Mookherjee N, Hancock RE. Cationic host defence peptides: innate immune regulatory peptides as a novel approach for treating infections. Cell Mol Life Sci 2007;64:922–933.
16. Bowdish DM, Davidson DJ, Hancock RE. A re-evaluation of the role of host defence peptides in mammalian immunity. Curr Protein Pept Sci 2005;6:35–51.
17. Belov K, Sanderson CE, Deakin JE, Wong ES, Assange D, McColl KA, Gout A, de Bono B, Barrow AD, Speed TP, Trowsdale J, Papenfuss AT. Characterization of the opossum immune genome provides insights into the evolution of the mammalian immune system. Genome Res 2007;17:982–991.
18. Scheetz T, Bartlett JA, Walters JD, Schutte BC, Casavant TL, McCray PB Jr. Genomics-based approaches to gene discovery in innate immunity. Immunol Rev 2002;190:137–145.
19. Schutte BC, Mitros JP, Bartlett JA, Walters JD, Jia HP, Welsh MJ, Casavant TL, McCray PB Jr. Discovery of five conserved β-defensin gene clusters using a computational search strategy. Proc Natl Acad Sci USA 2002;99:2129–2133.
20. Patil AA, Cai Y, Sang Y, Blecha F, Zhang G. Cross-species analysis of the mammalian β-defensin gene family: presence of syntenic gene clusters and preferential expression in the male reproductive tract. Physiol Genomics 2005;23:5–17.
21. Patil A, Hughes AL, Zhang G. Rapid evolution and diversification of mammalian α-defensins as revealed by comparative analysis of rodent and primate genes. Physiol Genomics 2004;20:1–11.
22. Lynn DJ, Bradley DG. Discovery of α-defensins in basal mammals. Dev Comp Immunol 2007;31:963–967.
23. Looft C, Paul S, Philipp U, Regenhard P, Kuiper H, Distl O, Chowdhary BP, Leeb T. Sequence analysis of a 212 kb defensin gene cluster on ECA 27q17. Gene 2006;376:192–8.
24. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402.
25. Durbin R, Eddy SR, Krogh A, Mitchison G. Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge, UK: Cambridge University Press; 1998.
26. Sonnhammer EL, Eddy SR, Birney E, Bateman A, Durbin R. Pfam: multiple sequence alignments and HMM-profiles of protein domains. Nucleic Acids Res 1998;26:320–322.
27. Fjell CD, Hancock RE, Cherkasov A. AMPer: a database and an automated discovery tool for antimicrobial peptides. Bioinformatics 2007;23:1148–1155.
28. Xiao Y, Hughes AL, Ando J, Matsuda Y, Cheng JF, Skinner-Noble D, Zhang G. A genome-wide screen identifies a single β-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. BMC Genomics 2004; 5:56.
29. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
30. Aich P, Heather LW, Neil AR, Shakiba J, Radhey SK, Ashley AB, Andrew AP, Lorne AB, Mitchell SA, Philip JG. Microarray analysis of gene expression following preparation of sterile intestinal "loops" in calves. Can J Animal Sci 2005;85:13–22.
31. Coombes BK, Coburn BA, Potter AA, Gomis S, Mirakhur K, Li Y, Finlay BB. Analysis of the contribution of Salmonella pathogenicity islands 1 and 2 to enteric disease progression using a novel bovine ileal loop model and a murine model of infectious enterocolitis. Infect Immun 2005;73:7161–7169.
32. Mookherjee N, Wilson HL, Doria S, Popowych Y, Falsafi R, Yu JJ, Li Y, Veatch S, Roche FM, Brown KL, Brinkman FS, Hokamp K, Potter A, Babiuk LA, Griebel PJ, Hancock RE. Bovine and human cathelicidin cationic host defense peptides similarly suppress transcriptional responses to bacterial lipopolysaccharide. J Leukoc Biol 2006;80:1563–1574.
33. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.
34. Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. Bioinformatics methods and protocols: methods in molecular biology, Vol. 132. Totowa, NJ: Humana Press; 2000. pp 365–386.
35. Whale TA, Wilson HL, Tikoo SK, Babiuk LA, Griebel PJ. Pivotal Advance: Passively acquired membrane proteins alter the functional capacity of bovine polymorphonuclear cells. J Leukoc Biol 2006;80: 481–491.
36. Boguski MS, Lowe TMJ, Tolstoshev CM. dbEST—database for "expressed sequence tags." Nat Genet 1993;4:332–333.
37. Levy O, Weiss J, Zarember K, Ooi CE, Elsbach P. Antibacterial 15-kDa protein isoforms (p15s) are members of a novel family of leukocyte proteins. J Biol Chem 1993;268:6058–6063.
38. Jenssen H, Hamill P, Hancock REW. Peptide antimicrobial agents. Clin Microbiol Rev 2006;19:491–511.
39. Kim HS, Yoon H, Minn I, Park CB, Lee WT, Zasloff M, Kim SC. Pepsin-mediated processing of the cytoplasmic histone H2A to strong antimicrobial peptide buforin. I. J Immunol 2000;165:3268–3274.
40. Eisenhauer PB, Lehre RI. Mouse neutrophils lack defensins. Infect Immun 1992;60:3446–3447.