



OPEN

In silico identification and functional validation of linear cationic α -helical antimicrobial peptides in the ascidian *Ciona intestinalis*

Yukio Ohtsuka[✉] & Hidetoshi Inagaki[✉]

We developed a computing method to identify linear cationic α -helical antimicrobial peptides (LCAMPs) in the genome of *Ciona intestinalis* based on its structural and physicochemical features. Using this method, 22 candidates of *Ciona* LCAMPs, including well-known antimicrobial peptides, were identified from 21,975 non-redundant amino acid sequences in *Ciona* genome database, Ghost database. We also experimentally confirmed the antimicrobial activities of five LCAMP candidates, and three of them were found to be active in the presence of 500 mM NaCl, nearly equivalent to the salt concentration of seawater. Membrane topology prediction suggested that salt resistance of *Ciona* LCAMPs might be influenced by hydrophobic interactions between the peptide and membrane. Further, we applied our method to *Xenopus tropicalis* genome and found 11 LCAMP candidates. Thus, our method may serve as an effective and powerful tool for searching LCAMPs that are difficult to find using conventional homology-based methods.

Abbreviations

LCAMP	Linear cationic α -helical antimicrobial peptide
MIC	Minimum inhibitory concentration
LPS	Lipopolysaccharide
EF1 α	Elongation factor 1 α
PPIA	Peptidylprolyl isomerase A

Antimicrobial peptides are crucial and widespread effector molecules of the innate immune system and are a part of the first line of host defense against invading pathogens^{1,2}. To date, thousands of antimicrobial peptides have been identified from bacteria to mammals and have been classified into several groups on basis of the structure and sequence features^{3–5}. Of these, linear cationic α -helical antimicrobial peptides (LCAMPs), the genes of which encode short amphipathic peptides without cysteine residues, are the most abundant and widespread in organisms. LCAMPs are known for their broad-spectrum activities including the ability to rapidly kill or neutralize bacteria, fungi, viruses, parasites, and even cancer cells^{6,7}. Therefore, they are considered promising lead candidates for the development of new peptide antibiotics⁸. There is a need for novel antibiotics due to the growing problem of microbial resistance to conventional antibiotics. LCAMPs might reduce the risk of emergence of antibiotic-resistant bacteria. In vitro development-resistance studies showed that pexiganan, an LCAMP, possesses a low potential for induction of bacterial resistance⁹.

Marine organisms are considered the richest source of novel antimicrobial peptides. They live in close proximity with pathogenic microbes. The estimated density of bacteria in seawater is approximately 10⁶ bacteria/mL¹⁰. In order to survive in this environment, they need to have a robust and effective immune system. Marine organisms are also exposed to extreme conditions, such as low temperature, high salinity and elevated pressure. The diversity of the marine environment has provided novel and unique sources of potential antimicrobial

Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan. ✉email: y-ohtsuka@aist.go.jp; h-inagaki@aist.go.jp

peptides¹¹. In fact, marine antimicrobial peptides were found to be structurally different from their analogs in terrestrial species¹². However, relatively limited numbers of antimicrobial peptides have been isolated from marine organisms, because antimicrobial peptides are a rapidly evolving group and display low sequence homology between orthologous antimicrobial peptides. Leoni et al. successfully identified a novel LCAMP family in marine mussels by bioinformatics analyses of available genomic and transcriptomic data¹³. To effectively and accurately identify novel antimicrobial peptides, there is a need to develop additional bioinformatics tools to survey protein sequence databases without use of sequence homology.

The ascidian *Ciona intestinalis* type A (hereinafter referred to as *C. intestinalis*) is a well-characterized marine invertebrate. Its genome has been sequenced¹⁴ and a large quantity of information about its transcriptome has been collected by EST and RNA-seq analyses^{15,16}. It has been concluded that among invertebrates, ascidians are the closest phylogenetic relatives to vertebrates. Moreover, genome-wide sequence analysis revealed that they have a set of genes for innate immune system, such as complement factors and Toll-like receptors, but lack an adaptive immune system that is present in higher vertebrates^{17,18}. These findings suggest that innate immune system for host defense is highly evolved in ascidians.

Previously, two salt-resistant LCAMPs, Ci-MAM-A and Ci-PAP-A, were identified from *C. intestinalis* hemocyte EST database by searching signal peptide and cationic helical region^{19,20}. In this study, to identify further novel LCAMPs in *C. intestinalis*, we developed an in silico screening method based on further criteria such as the size, amphipathicity and aggregation propensity. Consequently, we computationally predicted 22 potential LCAMP candidates in *Ciona* genome and experimentally confirmed five novel salt-resistant LCAMPs with broad-spectrum antimicrobial activity. This strategy was also successfully applied to *Xenopus tropicalis* genome, suggesting that our method could be applicable to the in silico screening of any genome.

Results

Genome-wide search for discovery of novel *Ciona* LCAMPs. We developed a computing method for the detection of LCAMPs in the *Ciona* genome based on the structural and physicochemical features of known LCAMP precursors: short length, secretory peptide, and cationic amphipathic α -helix of mature peptide. The *Ciona* genomic database called Ghost database, includes 21,975 non-redundant complete amino acid sequences, from which 1,949 sequences with 50–100 residues in length were extracted.

Firstly, to find LCAMP precursors from 1,949 sequences, we predicted signal peptides using SignalP 4.1 server. In total, 252 sequences were predicted to be secretory peptides, and the sequences corresponding to the signal peptides were eliminated from them. Furthermore, the secretion of their peptides into extracellular space was confirmed using TMHMM server, and then, eight sequences were eliminated as transmembrane proteins. Next, sequences containing cysteine residues were eliminated from the 244 subsequences without signal peptides, and cationic amphipathic α -helical structure was predicted using HeliQuest server. HeliQuest screening was performed for 165 subsequences without cysteine residues and 80 putative cationic amphipathic helical segments were identified. The secondary structure of HeliQuest-identified cationic amphipathic segments was further confirmed. Using PROTEUS server, it was found that 44 segments had the potential to form α -helical structures. Subsequently, the aggregation propensity of cationic amphipathic α -helical segments was predicted using AGGRESCAN, in which normalized average of aggregation propensity (Na⁴vSS) of the proteins was evaluated. Forty-four cationic amphipathic α -helical segments had Na⁴vSS values within a broad range of – 65.6 to 14.8. A previous study reported that antimicrobial peptides have Na⁴vSS values within the range of – 40 to 60²¹. Therefore, we investigated whether cationic amphipathic α -helical segments with Na⁴vSS value less than – 30 have antimicrobial activity. Four selected segments, KH.C1.365, KH.L63.9, KH.S1007.2, and KH.S930.2 with Na⁴vSS values of – 32.9, – 65.6, – 48.1 and – 44.4, respectively, exhibited no or weak antimicrobial activities against *Escherichia coli* (Supplementary Fig. S1). Consequently, as potential candidates for *Ciona* LCAMPs, we identified 24 segments with Na⁴vSS value between – 30 and 30. Finally, the prediction of subcellular localization was performed to eliminate cellular components from LCAMP candidates. Using DeepLoc-1.0 server, 22 genes were identified as LCAMP candidates (Table 1).

Ciona LCAMP candidates identified by our screening method include known LCAMPs such as Ci-MAM-family genes²⁰ and a number of uncharacterized proteins. In addition to *Ciona* genome, we predicted LCAMPs in *X. tropicalis* genome using our screening method. Eleven LCAMP candidates containing seven known LCAMPs²² were selected from 26,132 non-redundant protein sequences in the *X. tropicalis* genome (Supplementary Table S1), which suggests that our screening method is able to support the identification of LCAMPs from various genomes. LCAMP precursors are known to have a highly conserved signal peptide²³. Two conserved signal peptides were found in *Ciona* LCAMP candidates identified by our screening method (Fig. 1). In five members of Ci-MAM family, signal peptides composed of 21–22 amino acid residues were well-conserved (80% majority-rule consensus sequence: MDRKIVFALXLVXXLVXSXXXA). In addition, another conserved signal peptide composed of 23 amino acid residues (80% majority-rule consensus sequence: MNKSALLXLLXXGLXVLXEXXXA) was identified in six LCAMP candidates, suggesting that they belong to a novel LCAMP family. Previous studies showed that *Ciona* LCAMP precursors possess an anionic helical region in addition to a cationic amphipathic region^{19,20}. Using HeliQuest server, anionic helical sequences were found at the C-terminus of 17 *Ciona* LCAMP precursor candidates (Table 1), indicating that most *Ciona* LCAMP precursors are composed of a signal peptide, a cationic amphipathic region, and an anionic helical region in that order.

Antimicrobial activity of *Ciona* LCAMPs. Uncharacterized proteins in Table 1 are promising candidates of novel LCAMPs. To identify novel LCAMPs in *Ciona*, we tested antimicrobial activity in five LCAMP candidates, KH.C1.640, KH.C7.94, KH.S1531.4, KH.S908.1, and KH.S921.1. All synthesized peptides exhibited a broad-spectrum of antimicrobial activity (Table 2). They had potent antimicrobial activity against two anaerobic

Gene ID	Cationic amphipathic helix sequence	Anionic helix sequence	Putative protein product
KH.C1.100	WSSLGRSLLRLTHALKPLA	FAPEIEDELESEMMDRIMQQMADEKQ	Ci-MAM-A
KH.C1.1006	WTVTRYWNQKLANLLAGK		Uncharacterized protein
KH.C1.1192	DGWVRTGLAVARLVGRRRRRWNE- ANGL	RWNEANGLEKSSDAEETLSAAE- MEEVMQKIMDHQ	Ci-MAM-D ^a
KH.C1.1268	DWVRTAIGVAGLVGRRRRGGWN- QANGLKKFSSDA	LKKFSSDAEETLSAAEMEEVMQKIMDHQ	Ci-MAM-E ^a
KH.C1.453	ALRSVRTVARVGRAVLPHV	QQPTSQADMLEDALQAIEALMQE	Ci-PAP-A
KH.C1.640	WLSRRRSLFYWRRRRYDQ	RRYDQSAKDMNEDEPAETAISDM- LAKEDMPE	Uncharacterized protein
KH.C10.625	VFPSRRYGLWSSFRRIIRIHPQP		Uncharacterized protein
KH.C14.152	FRGLVRAGGKLVKEVLPIS	PWTKWNEQERMALADEIDAELMDLLDQ	Ci-META4
KH.C14.235	FRGLVRAGGKLVKEVLPIS	PWTKWNEQERMALADEIDAELMDLLEQ	Ci-META4
KH.C4.429	RLPQTVKNSLKHDQVIWKLKLVKGV		Uncharacterized protein
KH.C5.577	EHVFFSRRRRWTRWNQKVVVEDI	WNQKVVVEDILEAMEQSDMLH	Uncharacterized protein
KH.C7.94	FSRRRFDPSRRRIYVARRRSLAFAHRRRF- GDTA	NNPIKTDSSETSIIHSDQADDELQMMAE	Uncharacterized protein
KH.L136.4	WWLSGRRRVGRRRRRIIAY		Uncharacterized protein
KH.L14.3	YNARDLAKRNVGVSGQRVSI	LGDTADINEYLSRLLDYES	Uncharacterized protein
KH.S1531.4	INKKFRWHGKRKWWLRFVKQYSGNENIQ		Uncharacterized protein
KH.S2775.1	KGRSLKKIRQFWRKFKYKPF	NENIQRNFLNMERESIEEMMADEIFRNF	Uncharacterized protein
KH.S655.3	RRRRIAGKIGGGVAKTAAEL	TAAELAAEQALESSTGGGSWS	Ci-META6
		WNQQRKMENAMNDEMDEALLNLLKE	
KH.S775.4	SLLRVETKAILGTLALRRRTWENKASQ- QITP	ITPEMEEKLDAEMEKLMLQQLAEDQ	similar to Ci-MAM-A
KH.S775.5	SLAAGFKLLFKSWVHRRRTWNEKTTSD	NEKTTSDQFTAENDKPFDAEMDGLLNN- LETENQ	Ci-MAM-C ^a
KH.S908.1	KGGKFLNFKKAAKVGAKVGMAL	MAALGDEGEIEAFERLDTETQHAILAE- ALEN	Uncharacterized protein
KH.S921.1	TWPKNYWRKVWSKKNWRKFKFKH- WNQGGQNVED	KFKHWNQGGQNVEDMDLEDMLLWE	Uncharacterized protein
KH.S921.3	QAGVDRKFWTRKHWSQVGGKLRWN- QKQNVENM	RWNQKQNVENMDLDDEIQYYE	Uncharacterized protein

Table 1. The list of LCAMP candidates predicted from *Ciona* genome database. ^aNomenclature of LCAMPs by Fedders et al.¹⁹.

A Ci-MAM family

KH.C1.100	MDRKIVFALLLLVSLNVSTVYG	C1.100	C1.1192	C1.1268	S775.4	S775.5
KH.C1.1192I...V.-.L..QATA	***	***			
KH.C1.1268	.N.....F.V.-.F..QAAA	25.4	79.3	***		
KH.S775.4A..F.FVI-.Q.T..SA	26.2	33.3	32.3	***	
KH.S775.5F.VF-.HI...SA	43.5	19.0	22.0	35.6	***
Consensus	*:***.* ::::.* : : . .	23.7				

B novel family

KH.L136.4	MNKSALLLLLIGLLVLTETTTA	L136.4	C1.1006	C10.625	C5.577	S921.1	S921.3
KH.C1.1006	.K..LI.V.FV.....A..SE.	***	***				
KH.C10.625VV.P...IG..N.	11.4	16.0	***			
KH.C5.577I.....P.AS	26.0	20.9	18.4	***		
KH.S921.1V...NSD.	19.6	16.3	15.7	23.5	***	
KH.S921.3V..V.....D.	14.3	18.8	18.9	25.0	40.8	***
Consensus	*:** :*:*:::* :** : :	18.0					

Figure 1. Multiple sequence alignment of the signal peptide and identity matrix in LCAMP families. Sequences of the predicted signal peptides in Ci-MAM (A) and novel (B) families were aligned using ClustalW. Percentage identity values for pairwise comparisons of the sequences without signal peptides of these family genes are shown.

Peptide	MIC (μM)			
	<i>E. coli</i> (NBRC 14237)	<i>S. aureus</i> (NBRC 12732)	<i>S. cerevisiae</i> (NBRC 10217)	<i>P. aeruginosa</i> (NBRC 12582)
KH.C1.640	<3.125	<3.125	12.5	Negative
KH.C7.94	6.25	6.25	6.25	Negative
KH.S1531.4	<3.125	<3.125	6.25	Negative
KH.S908.1	<3.125	<3.125	6.25	6.25
KH.S921.1	<3.125	<3.125	Negative	<3.125

Table 2. Antimicrobial properties of *Ciona* LCAMPs.

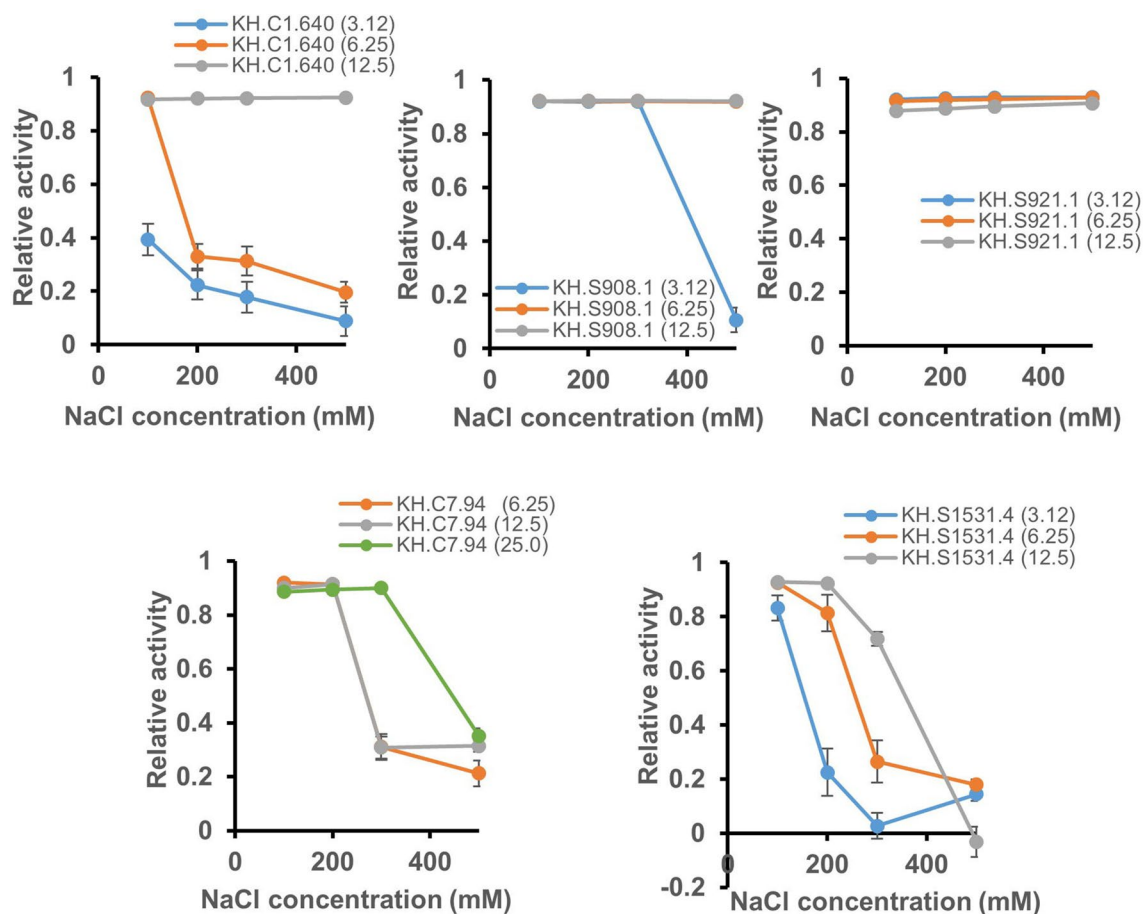


Figure 2. Effect of salt on the antimicrobial activity of *Ciona* LCAMPs. Antimicrobial activity of KH.C1.640, KH.C7.94, KH.S1531.4, KH.S908.1, and KH.S921.1 peptides against *E. coli* was assessed in the presence of 100, 200, 300, and 500 mM NaCl. Peptide concentrations were 3.12, 6.25, and 12.5 μM . Error bars show the standard deviation ($n=3$) from a single experiment.

bacteria *E. coli* and *Staphylococcus aureus*. KH.S908.1 and KH.S921.1 peptides also showed potent antimicrobial activity against aerobic bacteria *Pseudomonas aeruginosa*. Furthermore, antifungal activity against *Saccharomyces cerevisiae* was observed in KH.C1.640, KH.C7.94, KH.S1531.4, and KH.S908.1 peptides. These results show that our screening method is capable of identifying novel LCAMPs.

Ci-PAP-A (KH.C1.453) and Ci-MAM-A (KH.C1.100) are known to exhibit antimicrobial activity in the presence of 100 and 450 mM NaCl, respectively^{19,20}. We examined the effects of five novel *Ciona* LCAMPs in the presence of NaCl on antimicrobial activity against *E. coli*. Reflective of the habitat of *C. intestinalis*, salt resistance was found in all peptides at the concentration of 12.5 μM (Fig. 2). In particular, KH.C1.640, KH.S908.1, and KH.S921.1 peptides retained potent antimicrobial activity in the presence of up to 500 mM NaCl. Previous studies suggest that helical stability, hydrophobicity and amphipathicity in LCAMPs are important for salt resistance^{24–27}. However, each property of *Ciona* LCAMPs cannot explain salt resistance per se, as shown in Table 3. Therefore, we predicted the membrane interaction property of *Ciona* LCAMPs using AmphipaSeek server, which predicts that a protein or peptide can be monotopic and anchored via an amphipathic helix inserted parallel to the membrane interface²⁸. The prediction showed that a membrane anchor region was found in higher salt resistant

Gene ID	Antimicrobial peptide sequence	Hydrophobicity	Hydrophobic moment	Net charge
KH.C1.640	WWLSRRRSSLFYWR	0.624	0.170	+ 4
KH.C7.94	SRRRFDFSRRIYVARRRSLAFAHRRR	- 0.033	0.118	+ 11
KH.S1531.4	TINKKFRWHGKRKWWLRFVKQ	0.316	0.151	+ 8
KH.S908.1	KGGKFLNFKKAAKVGAKVGMALG	0.294	0.319	+ 6
KH.S921.1	KLKTWPKNYWRKVWSKKNWRK-FVKKFKHW	0.272	0.141	+ 12
KH.C1.100 (Ci-MAM-A)	WRS LGRTLRL SHALKPLARRSGW	0.428	0.380	+ 6
KH.C1.453 (Ci-PAP-A)	ALRS AVRTVARV GRAVLPHVAI	0.449	0.350	+ 4

Table 3. Physicochemical properties of *Ciona* LCAMPs. The membrane anchor segments predicted by AmphipaseK are highlighted in bold. The hydrophobicity, hydrophobic moment, and net charge values calculated by HeliQuest.

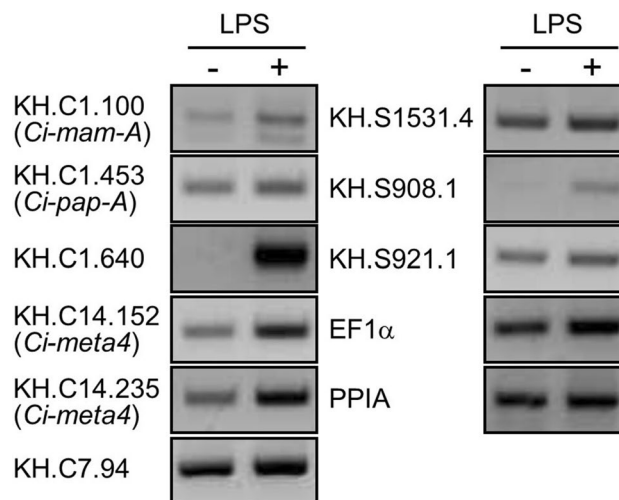


Figure 3. LCAMP mRNA expression in *Ciona* hemocytes stimulated with LPS. The expression of LCAMP genes in *Ciona* hemocytes treated (+) or untreated (-) with 0.1 mg/mL LPS for 1 h was analyzed by RT-PCR. EF1 α and PPIA were used as internal controls.

LCAMPs (Ci-MAM-A, KH.C1.640, KH.S908.1, and KH.S921.1) but not in low-salt resistant LCAMPs (Ci-PAP-A and KH.C7.94; Table 3). This suggests that in-plane membrane anchoring might be responsible for salt resistance.

LCAMP gene expression in *Ciona* hemocytes. To investigate the expression of *Ciona* LCAMPs in response to microbial infections, we performed RT-PCR analysis on *Ciona* hemocytes stimulated with bacterial lipopolysaccharide (LPS). First, the expression of *Ci-mam-A* (KH.C1.100), known as an inducible antimicrobial peptide²⁰, was examined using our culture system. As a previous study, *Ci-mam-A* expression was induced by LPS stimulation (Fig. 3). The expression of eight LCAMP genes in LPS-treated hemocytes was also further assessed by RT-PCR. *Ci-meta4* and two novel LCAMP genes (KH.C1.640 and KH.S908.1) were transcriptionally upregulated by LPS, whereas there were no changes in four LCAMP genes containing *Ci-pap-A* (Fig. 3). NF- κ B is activated in cells challenged with LPS and other inflammatory stimuli and is involved in the transcriptional activation of responsive genes²⁹. To confirm the involvement of *Ciona* LCAMPs in the immune response, we predicted NF- κ B binding sites in the upstream regions of each gene. Using Match-1.0 public, putative NF- κ B binding sites were identified in the upstream regions of LPS-inducible LCAMP genes, but not in the upstream regions of KH.C7.94 and KH.S1531.4 genes (Table 4). These suggest that four inducible LCAMPs (Ci-MAM-A, Ci-META4, KH.C1.640, and KH.S908.1) may play a role in the innate immune response.

Discussion

In this study, we applied a computational approach to identify novel LCAMPs in *C. intestinalis* and identified 22 potential LCAMP candidates using their known physicochemical characteristics and subcellular localization predictions. Furthermore, we experimentally validated the broad-spectrum antimicrobial activity of five selected LCAMPs. Two of them were induced with LPS and putative NF- κ B binding sites were identified in the upstream region of their genes, suggesting that they may play a role in the innate immune response. However, our data do not rule out the involvement of the other three genes in the immune response, as they may be regulated by uncharacterized innate immune signaling pathways.

Gene ID	Positions from ATG start codon
KH.C1.100	- 99/- 108, - 480/- 493, - 1,291/- 1,300
KH.C1.453	- 523/- 532, - 545/- 554, - 1,121/- 1,131, - 1,210/- 1,221, - 1,235/- 1,248, - 1,635/- 1,648
KH.C1.640	- 286/- 299
KH.C14.152	- 37/- 46, - 243/- 253, - 1,449/- 1,458
KH.C14.235	- 1,446/- 1,455
KH.C7.94	N.D.
KH.S1531.4	N.D.
KH.S908.1	- 692/- 701, - 747/- 756
KH.S921.1	- 124/- 133, - 811/- 824, - 1,290/- 1,299

Table 4. Putative NF- κ B binding sites in the upstream region of *Ciona* LCAMP genes. *N.D. not detected.

In order to explore the proteins and peptides with similar function among different species, homology search tools such as BLAST have generally been used. Several defensin homologs were identified by BLAST search in plant and vertebrate genomes^{30–32}. In these studies, some defensins that could be missed by BLAST search were further identified by a hidden Markov model based on a conserved cysteine-rich defensin motif. In contrast to defensin with three unique cysteine frameworks, it is difficult to find LCAMPs among different species using commonly used homology search tools due to poor sequence conservation of LCAMPs, especially in the cationic amphipathic α -helical region³³. Therefore, computational approaches without sequence homology information have been designed to discover novel LCAMPs from sequence databases. In marine mussels, novel antimicrobial peptides, mytilalins, were identified in the genomes and transcriptomes using bioinformatics analyses¹³. In *Ciona*, two LCAMP families, Ci-MAM and Ci-PAP, were identified from a hemocyte EST database based on two structural criteria of LCAMPs, signal peptide, and cationic α -helical region^{19,20}. We developed an in silico method for screening LCAMPs based on seven criteria, size, cationicity, hydrophobicity, amphipathicity, helicity, aggregation propensity, and subcellular location. Thus, we identified peptides that potentially have antimicrobial activity based on the structural and physicochemical features and eliminated non-antimicrobial peptides which cannot come into contact with microbes using subcellular localization prediction. In the prediction of these properties, we used web-based tools that can handle a large number of sequences simultaneously (see “Materials and methods”). This enables screening of LCAMPs in genome databases. Consequently, we identified novel LCAMPs containing KH.S921.1 from *C. intestinalis* genome database using our in silico method. In addition to *C. intestinalis* genome, we applied this method to *X. tropicalis* and succeeded in identifying the cationic amphipathic region including known LCAMPs. Our newly developed method is effective in screening LCAMPs in the genomes of various species. However, our in silico method does not completely exclude false positives. Experimental validation is necessary to validate novel LCAMPs in other organisms using our screening method.

A number of LCAMP precursors are known to possess a signal peptide, cationic amphipathic region, and anionic proregion^{34,35}. In *C. intestinalis*, almost all candidates shared a typical primary structure composed of a signal peptide, a cationic amphipathic region, and an anionic helical proregion in that order. The signal peptide is highly conserved within the species, and this conservation tendency increases within the individual LCAMP families. For instance, the signal peptides of Ci-MAM family genes are well-conserved²⁰ (Fig. 1). We identified a novel conserved signal peptide composed of 23 amino acid residues, indicating that six LCAMP precursors containing KH.S921.1 with their conserved signal peptide form a novel LCAMP family. In this novel LCAMP family, the anionic proregion is short or absent, whereas in the Ci-MAM family, it is relatively conserved. LCAMP precursors require proteolytic processing in which anionic proregions are removed by proteolytic cleavage to attain their active forms. In frog, LCAMP precursors are processed at the RXXR-, KK-, or RR-motifs corresponding to common cleavage sites in proregions^{22,23}. However, in *Ciona* LCAMP candidates, acidic amino acid residues, E and D, were found around the boundary of putative cationic active region and proregion, but their basic processing motifs were not. In ascidian, LCAMP precursors may be processed at the N-terminal of the acidic amino acid residue. Previous studies showed that ascidian LCAMPs, Styelins and Clavanins, are cleaved at the N-terminal of D and DD, respectively^{36,37}.

Herein, we demonstrated salt resistance of *Ciona* LCAMPs. Many antimicrobial peptides lose their activities under physiological salt concentration, although they exhibit significant in vitro activity against bacteria in the presence of low salt concentration. Considering the therapeutic application, the salt-dependent inactivation of antimicrobial peptides is a major obstacle³⁸. Marine organisms, including *C. intestinalis*, are considered a good resource of salt-resistant antimicrobial peptides. So far, several salt-resistant antimicrobial peptides have been identified in marine organisms³⁹. We demonstrated that three synthetic peptides derived from KH.C1.640, KH.S908.1, and KH.S921.1 retained potent antimicrobial activity in the presence of up to 500 mM NaCl. Additionally, AmphipaSeek prediction strongly suggested these peptides appear to insert parallel to membrane interface. We propose that salt resistance is influenced by hydrophobic interaction between the peptide and membrane. The action mechanism of LCAMPs, such as dermaseptins and cecropins, is explained by “carpet model”⁴⁰. According to this model, the peptides first align parallel to the phospholipid bilayer membrane, remaining in contact with the lipid head groups and effectively coating the surrounding area. When a threshold concentration of the peptide is reached, the peptides self-assemble within the membrane plane and this reorientation leads to a local disturbance in membrane stability, causing the formation of large cracks, leakage of cytoplasmic

components, disruption of the membrane potential, and ultimately disintegration of the membrane. High-salt concentration tends to disturb the first step of the action mechanism, especially the electrostatic interaction between cationic peptides and negatively charged bacterial membranes. It seems likely that salt-resistant LCAMPs could be bound to the membrane surface by the hydrophobic interaction, even if the electrostatic interaction was disturbed under high-salt conditions.

In this study, we demonstrated the effectiveness of a combination of computational and experimental approaches used to identify putative and novel LCAMPs from ascidian and frog genomes. LCAMP is a promising candidate for a new class of antibiotics, with a broad spectrum of antimicrobial activity, ease of synthesis, and a novel mechanism of action against microbes. Our *in silico* screening method will allow us to discover novel LCAMPs from genetic resources and shed light on the relationship between the structure and function of LCAMPs.

Materials and methods

Genome-wide *in silico* screening for antimicrobial peptides. The amino acid sequences used for discovery of novel LCAMPs were retrieved from two genome databases, *Ciona intestinalis* type A (recently called *Ciona robusta*) sequences from Ghost database (<https://ghost.zool.kyoto-u.ac.jp/cgi-bin/gb2/gbrowse/kh/>)⁴¹ and *Xenopus tropicalis* sequences from Xenbase (<https://www.xenbase.org/entry/>)⁴². LCAMP prediction was performed for sequences with 50–100 residues in length. LCAMP candidates were selected based on signal peptide sequence, hydrophobicity, hydrophobic moment, net charge, α -helix structure, and aggregation propensity. The signal peptide searches and cleavage site predictions were carried out using SignalP 4.1 server (<https://www.cbs.dtu.dk/services/SignalP-4.1/>)⁴³. Upon the removal of the signal peptide region, the sequences were used for the prediction of membrane-spanning regions using TMHMM server v.2.0 (<https://www.cbs.dtu.dk/services/TMHMM/>)⁴⁴ to eliminate transmembrane proteins. Subsequently, cationic amphipathic helix prediction was performed using HeliQuest (<https://heliquest.ipmc.cnrs.fr/>)⁴⁵. HeliQuest program slides an 18-residue window along the protein sequence and calculates hydrophobicity, hydrophobic moment, and net charge for each generated segment. Appropriate screening parameters were determined by analyses of cecropin-A (GenBank: AAA29185), moricin (BAB13508), Ci-PAP-A (ABR45664), Ci-MAM-A (ACA97856), dermaseptin SI (CAD92232), and PGLa (CAA25963): hydrophobicity between 0 and 0.6, hydrophobic moment between 0.1 and 1.0, and net charge between 3 and 10. Sequence statistics used in HeliQuest screening were as follows: minimal number of polar residues, minimal number of uncharged residues (serine, threonine, asparagine, glutamine, and histidine), minimal number of glycine residues, and maximal number of charged residues were 6, 1, 0, and 12, respectively, and cysteine was excluded. For identification of an anionic helical region in the LCAMP precursor, parameters were limited: hydrophobicity between -0.7 and 0.4 , hydrophobic moment between 0 and 0.5 , and net charge between -10 and -3 . Secondary structures of putative cationic amphipathic segments were further confirmed by PROTEUS (<https://www.proteus2.ca/proteus/>)⁴⁶, and their aggregation propensities were calculated using AGRRESCAN (<https://bioinf.uab.es/aggrescan/>)⁴⁷. Finally, subcellular localization was predicted using DeepLoc-1.0 (<https://www.cbs.dtu.dk/services/DeepLoc/>) to confirm the secretion of LCAMP candidates⁴⁸.

Sequence analysis. Peptide sequences were aligned using ClustalW. The prediction of NF- κ B binding sites upstream of the initiating methionine codon of each gene (less than 2,000 base pairs) was carried out using Match-1.0 public (<https://www.gene-regulation.com/pub/programs.html>)⁴⁹. The values of the cut-off for core and matrix similarities were 0.75 and 0.8, respectively.

Antimicrobial activities of the peptides. Peptides used in this study were chemically synthesized using the Fmoc method and obtained from Eurofins Genomics (Tokyo, Japan). The sequences of synthetic peptides were as follows:

KH.C1.365, RRRRRIAARIGSGVAQTGGELI;
 KH.C1.640, WWLSRRRSSLFYWR;
 KH.C7.94, SRRRFDfSRRRIYVARRRSLAFHRRR;
 KH.L63.9, FRRRRRRRRHWHHHHHYHYHHRRRRRRRRR;
 KH.S1531.4, TINKKFRWHGKRKWWLRFVKQ;
 KH.S908.1, KGGKFLNFKKAAKVGAKVGMALG;
 KH.S921.1, KLKTWPKNYWRKVWSKKNWRKFVKKFKHW;
 KH.S930.2, AERLMRKAADHWSNKMAKDVIWWEQ;
 KH.S1007.2, FGRRRRVPGRRRRWWNERAMNEI.

Antimicrobial activities of the peptides were measured as described previously⁵⁰. Synthetic peptides were added to the microbial suspension in the individual wells of a microplate at final concentrations of 3.125, 6.25, 12.5, and 25 μ M. The relative activities (RA) were calculated by the equation $RA = ((A_{PC} - A_{NC}) - A_{SA}) / (A_{PC} - A_{NC})$, where A_{PC} , A_{NC} , and A_{SA} represent the absorbance of positive control, negative control, and samples, respectively. The bacteria and yeast strains used were *Escherichia coli* (NBRC14237), *Staphylococcus aureus* (NBRC12732), *Pseudomonas aeruginosa* (NBRC12582), and *Saccharomyces cerevisiae* (NBRC10217), which were purchased from Biological Resource Center, National Institute of Technology and Evaluation (NBRC, Tokyo, Japan). The interaction between *Ciona* LCAMPs and membrane interface was predicted using Amphipaseek server (https://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=NPSA/npsa_amphipaseek.html)²⁸.

Hemocyte collection and stimulation. *Ciona* adults were provided by the Maizuru Fisheries Research Station of Kyoto University and by the Misaki Marine Biological Station of the University of Tokyo through the National BioResource Project of MEXT, Japan.

The animals were bled by removal of the tunic and puncture of the heart, and washed with sterile Ca²⁺- and Mg²⁺-free artificial seawater (CMF-ASW; 463 mM NaCl, 11 mM KCl, 25.5 mM Na₂SO₄, 2.15 mM NaHCO₃, 5 mM HEPES (pH 8.0), 5 mM EDTA). The coelomic fluid (hemolymph) and wash solution were collected and filtered through a nylon net. After centrifugation at 1,000×g for 30 s, the hemocyte pellet was washed with CMF-ASW and resuspended in CMF-ASW. Coelomic fluid was further centrifuged and the resulting supernatant (hemolymph) was used for the cultivation of hemocytes.

Hemocytes were plated in a 24-well plate and cultured at 16 °C in filtered seawater containing 20% hemolymph in the presence of penicillin (final concentration of 50 units/mL) and streptomycin (final concentration of 50 µg/mL). After incubation for 36 h, the cells were cultured in 20% hemolymph/seawater without antibiotics and then stimulated with 0.1 mg/mL LPS from *E.coli* (Sigma-Aldrich, St. Louis, MO, USA) for 1 h at 16 °C.

Reverse transcription-polymerase chain reaction (RT-PCR) analysis. Total RNA was extracted from uninduced and LPS-induced cells and was reverse transcribed using PrimeScript RT reagent Kit (TAKARA Bio Inc., Shiga, Japan) with oligo dT primer. The primers used in this study are mentioned in Supplementary Table S2. PCR cycles were as follows: initial denaturation at 94 °C for 2 min; 35 or 40 cycles at 94 °C for 30 s, 48 °C for 30 s, and 68 °C for 30 s. EF1α (elongation factor 1α) and PPIA (peptidylprolyl isomerase A) were used as internal control for RT-PCR analyses. The PCR products were separated by electrophoresis on 2% agarose gel and stained with ethidium bromide.

Received: 18 February 2020; Accepted: 10 July 2020

Published online: 28 July 2020

References

- Bulet, P., Stocklin, R. & Menin, L. Anti-microbial peptides: From invertebrates to vertebrates. *Immunol. Rev.* **198**, 169–184 (2004).
- Zaslloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389–395 (2002).
- Boman, H. G. Antibacterial peptides: Basic facts and emerging concepts. *J. Intern. Med.* **254**, 197–215 (2003).
- Waghu, F. H., Barai, R. S., Gurung, P. & Idicula-Thomas, S. CAMPR3: A database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res.* **44**, D1094–1097 (2016).
- Wang, G., Li, X. & Wang, Z. APD3: The antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.* **44**, D1087–1093 (2016).
- Hoskin, D. W. & Ramamoorthy, A. Studies on anticancer activities of antimicrobial peptides. *Biochim. Biophys. Acta.* **1778**, 357–375 (2008).
- Jenssen, H., Hamill, P. & Hancock, R. E. Peptide antimicrobial agents. *Clin. Microbiol. Rev.* **19**, 491–511 (2006).
- Marr, A. K., Gooderham, W. J. & Hancock, R. E. Antibacterial peptides for therapeutic use: Obstacles and realistic outlook. *Curr. Opin. Pharmacol.* **6**, 468–472 (2006).
- Ge, Y. *et al.* In vitro antibacterial properties of pexiganan, an analog of magainin. *Antimicrob. Agents. Chemother.* **43**, 782–788 (1999).
- Ammerman, J. W., Fuhrman, J. A., Hagstrom, A. & Azam, F. Bacterioplankton growth in seawater: I. Growth kinetics and cellular characteristics in seawater cultures. *Mar. Ecol. Prog. Ser.* **18**, 31–39 (1984).
- Kang, H. K., Seo, C. H. & Park, Y. Marine peptides and their anti-infective activities. *Mar. Drugs* **13**, 618–654 (2015).
- Charlet, M. *et al.* Innate immunity. Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, *Mytilus edulis*. *J. Biol. Chem.* **271**, 21808–21813 (1996).
- Leoni, G. *et al.* Myticalins: A novel multigenic family of linear, cationic antimicrobial peptides from marine mussels (*Mytilus* spp.). *Mar. Drugs* **15**, 20 (2017).
- Dehal, P. *et al.* The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. *Science* **298**, 2157–2167 (2002).
- Brozovic, M. *et al.* ANISEED 2015: A digital framework for the comparative developmental biology of ascidians. *Nucleic Acids Res.* **44**, D808–818 (2016).
- Satou, Y. *et al.* Improved genome assembly and evidence-based global gene model set for the chordate *Ciona intestinalis*: New insight into intron and operon populations. *Genome Biol.* **9**, R152 (2008).
- Parrinello, N., Cammarata, M. & Parrinello, D. In *Advances in Comparative Immunology* (ed. Cooper, E. L.) 521–590 (Springer, Berlin, 2018).
- Azumi, K. *et al.* Genomic analysis of immunity in a Urochordate and the emergence of the vertebrate immune system: “Waiting for Godot”. *Immunogenetics* **55**, 570–581 (2003).
- Fedders, H. & Leippe, M. A reverse search for antimicrobial peptides in *Ciona intestinalis*: Identification of a gene family expressed in hemocytes and evaluation of activity. *Dev. Comp. Immunol.* **32**, 286–298 (2008).
- Fedders, H., Michalek, M., Grotzinger, J. & Leippe, M. An exceptional salt-tolerant antimicrobial peptide derived from a novel gene family of haemocytes of the marine invertebrate *Ciona intestinalis*. *Biochem. J.* **416**, 65–75 (2008).
- Torrent, M., Andreu, D., Nogues, V. M. & Boix, E. Connecting peptide physicochemical and antimicrobial properties by a rational prediction model. *PLoS One* **6**, e16968 (2011).
- Roelants, K. *et al.* Origin and functional diversification of an amphibian defense peptide arsenal. *PLoS Genet.* **9**, e1003662 (2013).
- Vanhoye, D., Bruston, F., Nicolas, P. & Amiche, M. Antimicrobial peptides from hylid and ranin frogs originated from a 150-million-year-old ancestral precursor with a conserved signal peptide but a hypermutable antimicrobial domain. *Eur. J. Biochem.* **270**, 2068–2081 (2003).
- Deslouches, B. *et al.* De novo generation of cationic antimicrobial peptides: Influence of length and tryptophan substitution on antimicrobial activity. *Antimicrob. Agents Chemother.* **49**, 316–322 (2005).
- Friedrich, C., Scott, M. G., Karunaratne, N., Yan, H. & Hancock, R. E. Salt-resistant alpha-helical cationic antimicrobial peptides. *Antimicrob. Agents Chemother.* **43**, 1542–1548 (1999).
- Park, I. Y. *et al.* Helix stability confers salt resistance upon helical antimicrobial peptides. *J. Biol. Chem.* **279**, 13896–13901 (2004).
- Rydlo, T., Rotem, S. & Mor, A. Antibacterial properties of dermaseptin S4 derivatives under extreme incubation conditions. *Antimicrob. Agents Chemother.* **50**, 490–497 (2006).

28. Sapay, N., Guermeur, Y. & Deleage, G. Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier. *BMC Bioinform.* **7**, 255 (2006).
29. Baldwin, A. S. Jr. The NF-kappa B and I kappa B proteins: New discoveries and insights. *Annu. Rev. Immunol.* **14**, 649–683 (1996).
30. Lynn, D. J. *et al.* Bioinformatic discovery and initial characterisation of nine novel antimicrobial peptide genes in the chicken. *Immunogenetics* **56**, 170–177 (2004).
31. Schutte, B. C. *et al.* Discovery of five conserved beta-defensin gene clusters using a computational search strategy. *Proc. Natl. Acad. Sci. USA* **99**, 2129–2133 (2002).
32. Silverstein, K. A., Graham, M. A., Paape, T. D. & VandenBosch, K. A. Genome organization of more than 300 defensin-like genes in Arabidopsis. *Plant Physiol.* **138**, 600–610 (2005).
33. Nicolas, P., Vanhoye, D. & Amiche, M. Molecular strategies in biological evolution of antimicrobial peptides. *Peptides* **24**, 1669–1680 (2003).
34. Novkovic, M., Simunic, J., Bojovic, V., Tossi, A. & Juretic, D. DADP: The database of anuran defense peptides. *Bioinformatics* **28**, 1406–1407 (2012).
35. Tincu, J. A. & Taylor, S. W. Antimicrobial peptides from marine invertebrates. *Antimicrob. Agents Chemother.* **48**, 3645–3654 (2004).
36. Zhao, C., Liaw, L., Lee, I. H. & Lehrer, R. I. cDNA cloning of Clavanins: Antimicrobial peptides of tunicate hemocytes. *FEBS Lett.* **410**, 490–492 (1997).
37. Zhao, C., Liaw, L., Lee, I. H. & Lehrer, R. I. cDNA cloning of three cecropin-like antimicrobial peptides (Styelins) from the tunicate, *Styela clava*. *FEBS Lett.* **412**, 144–148 (1997).
38. Mohanram, H. & Bhattacharjya, S. Salt-resistant short antimicrobial peptides. *Biopolymers* **106**, 345–356 (2016).
39. Sperstad, S. V. *et al.* Antimicrobial peptides from marine invertebrates: Challenges and perspectives in marine antimicrobial peptide discovery. *Biotechnol. Adv.* **29**, 519–530 (2011).
40. Shai, Y. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochim. Biophys. Acta* **1462**, 55–70 (1999).
41. Satou, Y., Kawashima, T., Shoguchi, E., Nakayama, A. & Satoh, N. An integrated database of the ascidian, *Ciona intestinalis*: Towards functional genomics. *Zool. Sci.* **22**, 837–843 (2005).
42. Karimi, K. *et al.* Xenbase: A genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Res.* **46**, D861–D868 (2018).
43. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
44. Moller, S., Croning, M. D. & Apweiler, R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**, 646–653 (2001).
45. Gautier, R., Douguet, D., Antonny, B. & Drin, G. HELIQUEST: A web server to screen sequences with specific alpha-helical properties. *Bioinformatics* **24**, 2101–2102 (2008).
46. Montgomerie, S., Sundararaj, S., Gallin, W. J. & Wishart, D. S. Improving the accuracy of protein secondary structure prediction using structural alignment. *BMC Bioinform.* **7**, 301 (2006).
47. Conchillo-Sole, O. *et al.* AGGRESKAN: A server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinform.* **8**, 65 (2007).
48. Almagro Armenteros, J. J., Sonderby, C. K., Sonderby, S. K., Nielsen, H. & Winther, O. DeepLoc: Prediction of protein subcellular localization using deep learning. *Bioinformatics* **33**, 3387–3395 (2017).
49. Kel, A. E. *et al.* MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* **31**, 3576–3579 (2003).
50. Steinberg, D. A. & Lehrer, R. I. Designer assays for antimicrobial peptides. Disputing the “one-size-fits-all” theory. *Methods Mol. Biol.* **78**, 169–186 (1997).

Acknowledgements

We thank Dr. Reiko Yoshida, Dr. Satoe Aratake, and other members working under the National BioResource Project (MEXT, Japan) for providing experimental animals and Dr. Satoshi Ueno for his technical help on this project. This work was supported by a grant from National Institute of Advanced Industrial Science and Technology.

Author contributions

Y.O. and H.I. designed and performed the experiments, and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-69485-y>.

Correspondence and requests for materials should be addressed to Y.O. or H.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020