CrossMark

Development of Antimicrobial Peptide Prediction Tool for Aquaculture Industries

Aditi Gautam[1] · Asuda Sharma[1] · Sarika Jaiswal[2] · Samar Fatma[2] ·
Vasu Arora[2] · M. A. Iquebal[2] · S. Nandi[3] · J. K. Sundaray[3] · P. Jayasankar[3] ·
Anil Rai[2] · Dinesh Kumar[2]

Abstract Microbial diseases in fish, plant, animal and human are rising constantly; thus, discovery of their antidote is imperative. The use of antibiotic in aquaculture further compounds the problem by development of resistance and consequent consumer health risk by bio-magnification. Antimicrobial peptides (AMPs) have been highly promising as natural alternative to chemical antibiotics. Though AMPs are molecules of innate immune defense of all advance eukaryotic organisms, fish being heavily dependent on their innate immune defense has been a good source of AMPs with much wider applicability. Machine learning-based prediction method using wet laboratory-validated fish AMP can accelerate the AMP discovery using available fish genomic and proteomic data. Earlier AMP prediction servers are based on multi-phyla/species data, and we report here the world's first AMP prediction server in fishes. It is freely accessible at http://webapp.cabgrid.res.in/fishamp/. A total of 151 AMPs related to fish collected from various databases and published literature were taken for this study. For model development and prediction, N-terminus residues, C-terminus residues and full sequences were considered. Best models were with

kernels polynomial-2, linear and radial basis function with accuracy of 97, 99 and 97 %, respectively. We found that performance of support vector machine-based models is superior to artificial neural network. This in silico approach can drastically reduce the time and cost of AMP discovery. This accelerated discovery of lead AMP molecules having potential wider applications in diverse area like fish and human health as substitute of antibiotics, immunomodulator, antitumor, vaccine adjuvant and inactivator, and also for packaged food can be of much importance for industries.

Keywords Antimicrobial peptides · Fish · Prediction · Support vector machine · Web-server

Introduction

Due to an increase in microbial disease incidences, in animals and plants, discovery for their antidote is simultaneously getting high priority in research. In the past years, microorganisms have developed resistance to antibiotics [1], thus also affecting both human health and agricultural economy [2]. The continuous increase of pathogens as well as resistance against antibiotics has led the researchers to search for new antimicrobial compounds from diverse sources. Among these compounds, antimicrobial peptides (AMPs) appear to be one of the most propitious candidates for clinical development in order to inhibit microbial activity because of their target specificity, speed of action and producing innate immunity in organism [3]. One of the most important factors for introducing AMPs as antibacterial agents is that it may not get resistance against bacteria [4] due to receptor-independent mechanism of action [5]. Antibiotics use in aquaculture is

Aditi Gautam, Asuda Sharma and Sarika Jaiswal have contributed equally to this work.

✉ Dinesh Kumar
 dineshkumarbhu@gmail.com

[1] Jaypee University of Information Technology, Solan, Himachal Pradesh, India

[2] Centre for Agricultural Bioinformatics, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

[3] Division of Fish Genetics and Biotechnology, ICAR-Central Institute of Freshwater Aquaculture, Bhubaneswar, Odisha, India

not preferred due to associated problem of resistance development and consumer health risk [6]. Due to associated toxicity, carcinogenicity, sensitivity or bio-accumulation leading to adverse human health, many of synthetic antibiotics, viz. nitrofurans and fluoroquinolones, and natural antimicrobial agents like malachite green are banned.

Aquaculture is also known as "underwater agriculture" which is important for food, employment and revenue [7]. It includes around 350 species having 34 finfish (piscean), 8 crustacean and 12 molluscan with productivity of more than 100,000 tonnes comprising economic value more than US$215 billion, which contributes around 0.3 % of the world economy [8–10]. But due to certain factors, economic loss in fisheries that has been reported in [8] is around US$50 billion per year. The most significant factor in fisheries deprivation is infectious diseases. The estimated loss in conservative figure is about tens of billions dollars over last 20 years [7].

AMPs are low molecular weight proteins of about 12–100 amino acid residues, usually positively charged. Since its 40–50 % residues are hydrophobic, it gets folded into an amphipathic structure making it soluble in aqueous and lipid [3]. AMPs are distinct molecules with the ability to kill targeted microbes using a mechanism in which some cationic peptides create electrostatic attraction for negatively charged phospholipids of microbial membranes and integrate into cell membrane of microbes resulting in membrane disintegration [11]. Genes of AMPs get expressed in different types of cells of host [12] supporting innate immunity also [13]. AMPs have been derived from many sources in which fish are the sources of many kinds of AMPs [14]. AMPs are of broad spectrum, acting against gram-positive and gram-negative bacteria, viruses and fungi. They have other roles like mediators of inflammation, cell proliferation, angiogenesis, wound healing, chemotaxis, immune induction and protease–antiprotease balance [15, 16]. They are also called "natural antibiotics" [15, 17].

As per World Organization for Animal Health, out of nine major diseases in fish, seven are the infectious disease, causing mortality. To combat this, antibiotics are used heavily, which often exceeds limits prescribed by statutory norms like tetracycline (0.1 mg/kg), oxytetracycline (0.1 mg/kg), trimethoprim (0.05 mg/kg) and oxolinic acid (0.3 mg/kg). Thus, there is a hard pressing need to look for alternates of antibiotics [18]. Various classes of AMPs found in fish include defensins, cathelicidins, hepcidins, histone-derived peptides and a fish-specific class of the cecropin family, called piscidins [13]. Since fishes are much more dependent on their innate immune defenses than mammals, they are potential rich source of therapeutic molecules against mammalian viral, pathogenic fungi, fibrosarcoma, bacterial bio-film in mastitis disease [19]. Discovery of such lead molecules having vast applications

and recombinant approach makes them of much potential industrious applications; thus, computational approach can be pivotal and can accelerate such investigations.

Earlier existing AMP prediction servers [20, 21] were generic and trained on all available multispecies wet laboratory-validated AMP data with various classifiers. Recently, species-specific approach without compromising prediction accuracy has been reported in cattle [22]. There is no such species-specific approach in case of fish reported so far. With the advent of low-cost sequencing technology, the plethora of genomic data is available. If machine learning techniques are used in user-friendly server mode, then the cost of AMP discovery will be drastically reduced by narrowing down the number of putative AMPs for antimicrobial assay through laboratory experiments. Thus, in silico search prediction is imperative. The present work aims at development of AMP prediction server using machine learning techniques, especially artificial neural networks (ANNs) and support vector machine (SVM) technique.

Materials and Methods

Data Collection and Preprocessing

There exist a number of AMP databases. An extensive search was made at the databases, viz. LAMP [23], CAMP [21, 24], PenBase [25], EROP [26] and APD2 [27] for experimentally validated AMPs specific to fishes and crustacean. A total of 308 AMPs were extracted after extensive search. To remove the redundancy, clustering was done with CD-HIT program [28] using "longest sequence first" list removal algorithm to remove sequences above certain identity threshold. Finally, 151 AMPs were selected as positive dataset for the study. Due to lack of experimentally proven non-antimicrobial peptides for fishes, peptides synthesized from fish mitochondria and other intracellular locations except the secretary proteins were considered as non-antimicrobial peptides [29] and taken as negative set. It is believed that eukaryotic mitochondrial organelle genome mimics prokaryotic genome features because of endosymbiont hypothesis endorsing prokaryotic origin of mitochondria during course of evolution [30].

Code in Perl was written to split the data into window size of 30, which is the average length of AMPs. The split sequences were validated in silico for their antimicrobial peptide property at antimicrobial peptide prediction servers like ABP2 [20] and CAMP [21]. A total of 31 features, selected on the basis of p values at 5 % level of significance based on Chi-square values obtained using STATISTICA version 6.0 software package [31], were considered for the model building. These features included
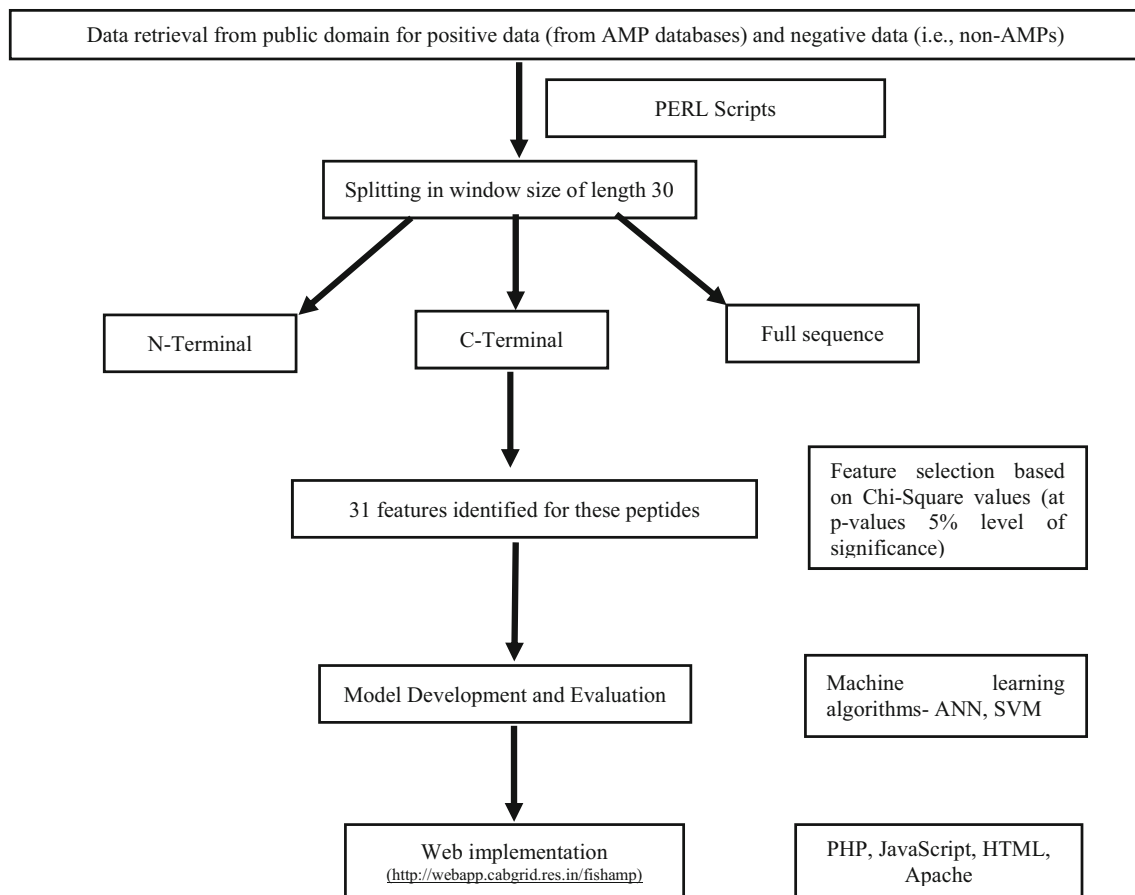
Fig. 1 Workflow of the fish AMP Web server

physicochemical properties like amino acid composition (AAC) of all the amino acids, molecular weight, theoretical pI, number of carbon atoms, number of hydrogen atoms, number of sulfur atoms, number of oxygen atom, number of nitrogen atoms, half-life, instability index, aliphatic index, grand average of hydropathy (GRAVY), which were calculated. These parameters were calculated for batch of AMPs and non-AMPs using ProtParam module of BioPerl [32]. The workflow (Fig. 1) illustrates the procedure followed.

The study was divided into three parts, viz. prediction based on N-terminal, C-terminal and full sequence. C-terminus is more responsible for the interaction of negatively charged bacterial membrane and involved in penetration, while the N-terminus is involved in interaction with intracellular components, thus hindering the metabolic functions of bacteria [12].

Model Development

The various models were tried for N-terminal, C-terminal and full sequence using SVM light [33] and STATISTICA version 6.0 [31]. Neural network is a powerful machine learning technique widely used to solve a variety of problems in pattern recognition, prediction, optimization, associative memory, etc. The neurons basically consist of inputs (like synapses), that are multiplied by weights (respective signals strengths), and then computed by a mathematical function which determines the activation of neuron. The output of the artificial neuron is computed by another function, which sometimes depends on certain threshold. ANNs combine artificial neurons to process information [34]. This technique has also been applied in the present study for predicting antibacterial peptides from fishes. In order to develop ANN models, we used STATISTICA [31]. In order to build the prediction model, ANNs with back propagation algorithm have been used [35], but it was found that it overfits and provides underestimation of actual prediction error specially in case of small sample size. Standard square error (SSE) was used for the error function. The two most popular and widely used networks, namely multilayer perceptron (MLP) and radial basis function (RBF), are trained using all the three learning algorithms, viz. gradient descent algorithm (GDA), Broyden–Fletcher–Goldfarb–Shanno (BFGS) and conjugate gradient descent algorithm

(CGDA), with a view to minimizing sum of the squared error function of the network output. Several learning rates [35] are considered for training the networks as well as for adjusting the weights. For hidden units and output units, several activation functions, viz. identity, tanh, logistic, exponential and sine, are tried. Performance of the trained network is assessed by computing different measures.

Support vector machines (SVMs) are a group of supervised learning methods applied to classification or regression. It is an extension to nonlinear models of the generalized algorithm and is based on the statistical learning theory and optimization theory [36]. For **x,** a vector in *n*-dimensional input space and $K(\mathbf{x}_i, \mathbf{x}_j)$ being the kernel function, some choices of kernel function are as follows:

such that each set falls under test set. The developed models were evaluated using the evaluation measures, viz. sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), false-positive rate (FPR), false discovery rate (FDR), F1 score, accuracy and Matthews correlation coefficient (MCC). The sensitivity indicates the "quantity" of predictions, i.e., the proportion of real positives correctly predicted. The specificity indicates the "quality" of predictions, i.e., the proportion of true negatives correctly predicted. The PPV indicates the proportion of true positives in predicted positives, "the success rate," while NPV is the proportion of true negatives in predicted negatives. In terms of definition, FDR of a set of predictions is the expected percent of false predictions in the set of predictions. The F1 score is the measure of a test's accuracy. The MCC can be seen as the correlation coefficient between the observed and predicted binary classifications.

$$\begin{cases} \text{Sensitivity } = \text{TP}/(TP + FN) * 100 \\ \text{PPV} = \text{TP}/(TP + FP) * 100 \\ \text{FPR} = \text{FP}/(FP + TN) \\ F1 = 2\text{TP}/(2\text{TP} + \text{FP} + \text{FN}) \\ \text{Accuracy} = \dfrac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})} * 100 \\ \text{MCC} = \dfrac{(\text{TP} * \text{TN} - \text{FP} * \text{FN})}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} * 100 \end{cases}$$

$$\begin{aligned} \text{Specificity} &= \text{TN}/(\text{FP} + \text{TN}) * 100 \\ \text{NPV} &= \text{TN}/(\text{TN} + \text{FN}) * 100 \\ \text{FDR} &= \text{FP}/(\text{TP} + \text{FP}) = 1 - \text{PPV} \end{aligned}$$

(a) $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$ (Linear SVM)

(b) $K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d$
(Polynomial SVM of degree *d*)

(c) $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{-\gamma ||\mathbf{x}_i - \mathbf{x}_j||^2\right\}$
(Radial Basis function Kernel)

(d) $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh\left(\gamma \mathbf{x}_i^T \mathbf{x}_j + r\right)$ (Sigmoid)

where *r*, *d*, $\gamma > 0$ are the kernel parameters.

In the present study, SVM light has been used to predict AMP of fishes for using different kernel functions, viz. polynomial degree 2, 3, RBF and sigmoid. The whole analysis is performed for three types, viz. full sequence, N-terminus residues, C-terminus residues.

where TP = true positive (correctly identified as positive); TN = true negative (correctly identified as negative); FP = false positive (incorrectly identified as positive); FN = false negative (incorrectly identified as negative).

Results and Discussion

The positive (antimicrobial peptides) and negative (non-antimicrobial peptides) datasets were compared based on average percentage composition of amino acids (Fig. 2). It was seen that the amino acid G was very much dominant in AMPs. Also, the amino acids K and R showed significant presence in AMPs as compared to non-AMPS. On the other hand, residues like D, Q, E, S and V were remarkably present in non-AMPs as compared to AMPs, which is in accordance with the previously reported research [20].

ANN and SVM machine learning techniques were tried for the study in order to find out the models with best accuracy. For ANN, STATISTICA version 6.0 [31] was used. Supplementary Table 1 shows the features selected for model development along with *p* values. In this study, we have developed classification models using ANN and SVM.
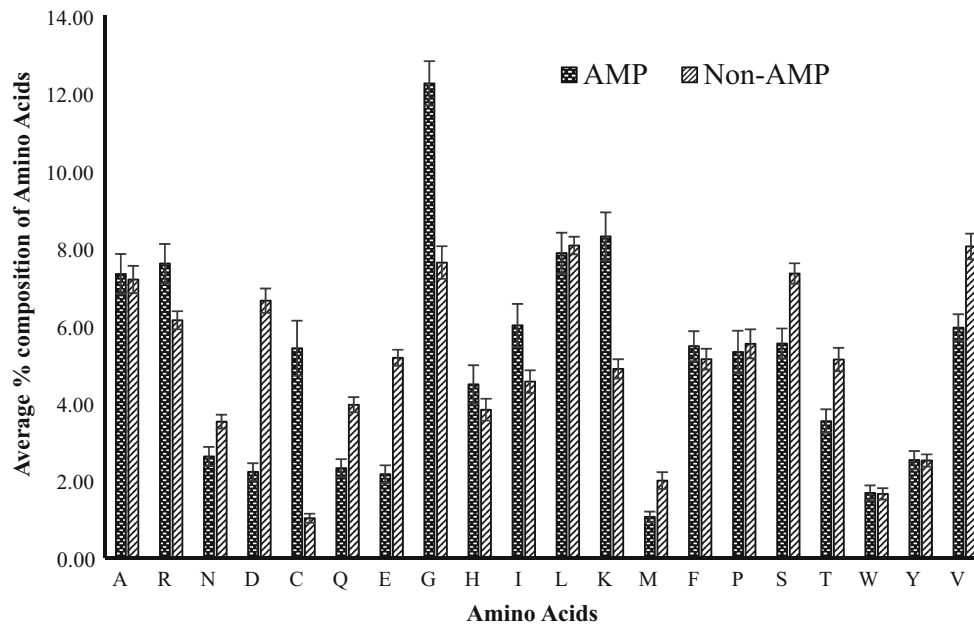
Model Evaluation

For error estimation, fivefold cross-validation technique [37] was applied where the whole data were divided into five sets, each having almost equal number of peptides. Further, four sets were clubbed and considered as training set, while the fifth set was taken as test set. This was repeated five times

Fig. 2 Comparison of average percentage of amino acids in antibacterial and non-antibacterial peptides (where all amino acids represent the standard codes)

Table 1 Performance of the models for full sequence residues using SVM and ANN methodology

Models	c	γ	No. of SVs	Sp	Sen	PPV	NPV	FPR	FDR	ACC	MCC	F1
SVM methodology												
Linear	19.75	–	58	0.93	0.90	0.93	0.91	0.07	0.07	0.92	0.84	0.92
Polynomial-2	57.25	0.03	119	0.93	0.92	0.93	0.92	0.07	0.07	0.93	0.85	0.92
Polynomial-3	66.50	0.05	144	0.91	0.94	0.91	0.94	0.09	0.09	0.92	0.84	0.92
RBF	45.00	0.01	66	0.98	0.96	0.98	0.96	0.02	0.02	0.97	0.94	0.97
Sigmoid	44.75	0.02	102	0.93	0.92	0.92	0.92	0.07	0.08	0.92	0.84	0.92
ANN methodology												
MLP 31-13-2	–	–	–	0.96	0.91	0.96	0.92	0.04	0.04	0.94	0.87	0.93

The models MLP 31-13-2, MLP 31-18-2 and MLP 31-15-2 with accuracy 94, 93 and 89 % were found to be best for full sequence, N-terminus residues and C-terminus residues, respectively. In case of full sequence, training algorithm BFGS, entropy error function, exponential activation function for hidden layer and for output layer softmax function were found to be the best. For N-terminus residues, activation function was exponential, while for C-terminus, it was tanh. Training algorithm and error functions for all three models were BFGS and entropy, respectively, whereas output layer function was softmax. The best models were selected on the basis of measures discussed in previous section, and the same are shown in Tables 1, 2 and 3, respectively, for full sequence, N-terminus residues and C-terminus residues.

The SVM models were developed using SVM light version 6.2. Various scripts in Perl were written to fine-

tune the parameters for final model selection. In comparison with ANN, SVM showed 97, 99 and 97 % accuracy for full sequence, N-terminus residues and C-terminus residues, respectively. This comparison was also evaluated using other studies where SVM methodology is reported to be better than ANN [20, 38].

The models were generated through SVM methodology considering five kernels, viz. linear, polynomial of degree 2, polynomial of degree 3, RBF and sigmoid function using criteria of maximum number of iterations or stop if error is <0.0001 for all the three sets (full sequence, N-terminus residues and C-terminus residues).

For full sequence, SVM model generated with RBF kernel function was considered to be best with *c* parameter as 45.0 using 66 numbers of support vectors. Further, the values of the other parameters, viz. sensitivity, specificity, accuracy, MCC, PPV, FPR, NPV, FDR and F1, were

Table 2 Performance of the models for N-terminus residues using SVM and ANN methodology

Models	c	γ	No. of SVs	Sp	Sen	PPV	NPV	FPR	FDR	ACC	MCC	F1
SVM methodology												
Linear	17.50	–	123	0.98	0.99	0.98	0.99	0.02	0.02	0.99	0.97	0.99
Polynomial-2	19.50	0.08	176	0.97	0.97	0.97	0.97	0.03	0.03	0.97	0.94	0.97
Polynomial-3	10.25	0.08	417	0.97	0.97	0.97	0.97	0.03	0.03	0.97	0.94	0.97
RBF	17.00	0.20	172	0.97	0.96	0.97	0.96	0.03	0.03	0.96	0.93	0.96
Sigmoid	8.00	0.11	240	0.97	0.97	0.97	0.97	0.03	0.03	0.97	0.94	0.97
ANN methodology												
MLP 31-18-2	–	–	–	0.94	0.93	0.94	0.93	0.06	0.06	0.93	0.87	0.93

Table 3 Performance of the models for C-terminus using SVM and ANN methodology

Models	c	γ	No. of SVs	Sp	Sen	PPV	NPV	FPR	FDR	ACC	MCC	F1
SVM methodology												
Linear	12.00	–	17	0.95	0.95	0.95	0.95	0.05	0.05	0.95	0.91	0.95
Polynomial-2	12.00	0.05	133	0.97	0.97	0.97	0.97	0.03	0.03	0.97	0.94	0.97
Polynomial-3	7.75	0.05	129	0.95	0.96	0.95	0.96	0.05	0.05	0.95	0.91	0.95
RBF	2.00	0.05	20	0.94	0.94	0.94	0.94	0.06	0.06	0.94	0.88	0.94
Sigmoid	1.50	0.31	81	0.92	0.94	0.93	0.94	0.08	0.07	0.93	0.87	0.93
ANN methodology												
MLP 31-15-2	–	–	–	0.89	0.89	0.89	0.89	0.11	0.11	0.89	0.78	0.89

obtained as 0.96, 0.98, 0.97, 0.94, 0.98, 0.96, 0.02, 0.02 and 0.97, respectively (Table 1).

Similarly, for N-terminus residues, linear kernel function was considered best with c parameter as 17.50. Values of sensitivity, specificity, accuracy, MCC, PPV, FPR, NPV, FDR and F1 were found to be 0.99, 0.98, 0.99, 0.97, 0.98, 0.99, 0.02, 0.02 and 0.99, respectively (Table 2). Moreover, polynomial kernel function with degree 2 was perceived to be the best for C-terminus residues with parameter c as 12. Values of sensitivity, specificity, PPV, NPV, FPR, FDR, accuracy, MCC and F1 were found to be 0.97, 0.97, 0.97, 0.94, 0.97, 0.97, 0.03, 0.03 and 0.97, respectively (Table 3).

The generic AMP prediction servers like ABP2 reported accuracy of 92.14 % by SVM methodology [20], while CAMP server has 93.2 % accuracy by random forest approach and 91.5 % by SVM approach [21]. The present fish-specific AMP prediction server is reported to give the predicted accuracy of 99 %. The higher accuracy may be due to the AMP diversification during course of evolution on specialized species-specific parameters.

AMPs predicted through this server can be modified for better antimicrobial activity to be used in industries. This can be done by modifying their polypeptide structure, thus changing its biological and chemical activity. Such modifications can be done at genetic level for novel antimicrobial peptides generation. Such approach is coming up,

for example, nisin AMP which is even approved by FDA. AMPs have wide applications in food additives, cosmetics, ointments, injections, etc., and thus characterization of the antimicrobial peptides is highly warranted [39].

The present server can not only be used for discovery of novel AMPs using recently available genomic and proteomic data, but can also be used for future fish genome and proteome data. Fish AMPs are unique in comparison with other phyla in terms of much wider applications. In case of human diseases, Crohn's disease and Kostmann's syndrome treatment, AMPs from rainbow trout have shown promising result as immunomodulator [40]. Similarly, antitumor activity is reported from tilapia [41] and grouper fish [42, 43]. Fish AMPs from rainbow trout have been reported to increase efficacy of vaccination as an adjuvant [44]. The issues of formalin toxicity in developing inactivated vaccine can be resolved by substituting it with fish AMPs, e.g., grouper fish AMP has been reported for inactivation of NNV (nervous necrosis virus) [45]. Red sea bream fish AMPs have been used for non-contaminated coatings of food packages effective against both gram-negative and gram-positive bacteria [46]. Such wider application needs wet laboratory validation and its efficacy in different diverse applications. The comparisons of all the 3 cases (N-terminal, C-terminal and full sequences) have been made to show the trend of specificity and sensitivity using the training and test data (Fig. 3a–c).

Web Implementation

The best model based on above evaluation criteria for N-terminal, C-terminal and full sequence was implemented and is made available at http://webapp.cabgrid.res.in/fish amp/. This user-friendly server is designed with PHP using JavaScript having the interface composed of dynamic Web pages based on user input. This is supported by Apache, an open source Web server. The server has six menus, viz. Home, Algorithm, Submission, Links, Tutorial and Team. Submission page allows user to generate result using various input options like "Input Sequence" or "Upload File"

Fig. 3 **a** Comparison of various methods in terms of specificity and sensitivity using the training and test data for N-terminal (*Sp* specificity, *Sen* sensitivity, *Tr* training set, *Ts* test set, *Nt* N-terminal, *ANN* artificial neural network, *Lin* linear, *Pol-2* polynomial of degree 2, *Pol-3* polynomial of degree 3, *RBF* radial basis function, *Sig* sigmoid). **b** Comparison of various methods in terms of specificity and sensitivity using the training and test data for C-terminal (*Sp* specificity, *Sen* sensitivity, *Tr* training set, *Ts* test set, *Ct* N-terminal, *ANN* artificial neural network, *Lin* linear, *Pol-2* polynomial of degree 2, *Pol-3* polynomial of degree 3, *RBF* radial basis function, *Sig* sigmoid). **c** Comparison of various methods in terms of specificity and sensitivity using the training and test data for full sequence (*Sp* specificity, *Sen* sensitivity, *Tr* training set, *Ts* test set, *Full* full sequence, *ANN* artificial neural network, *Lin* linear, *Pol-2* polynomial of degree 2, *Pol-3* polynomial of degree 3, *RBF* radial basis function, *Sig* sigmoid)
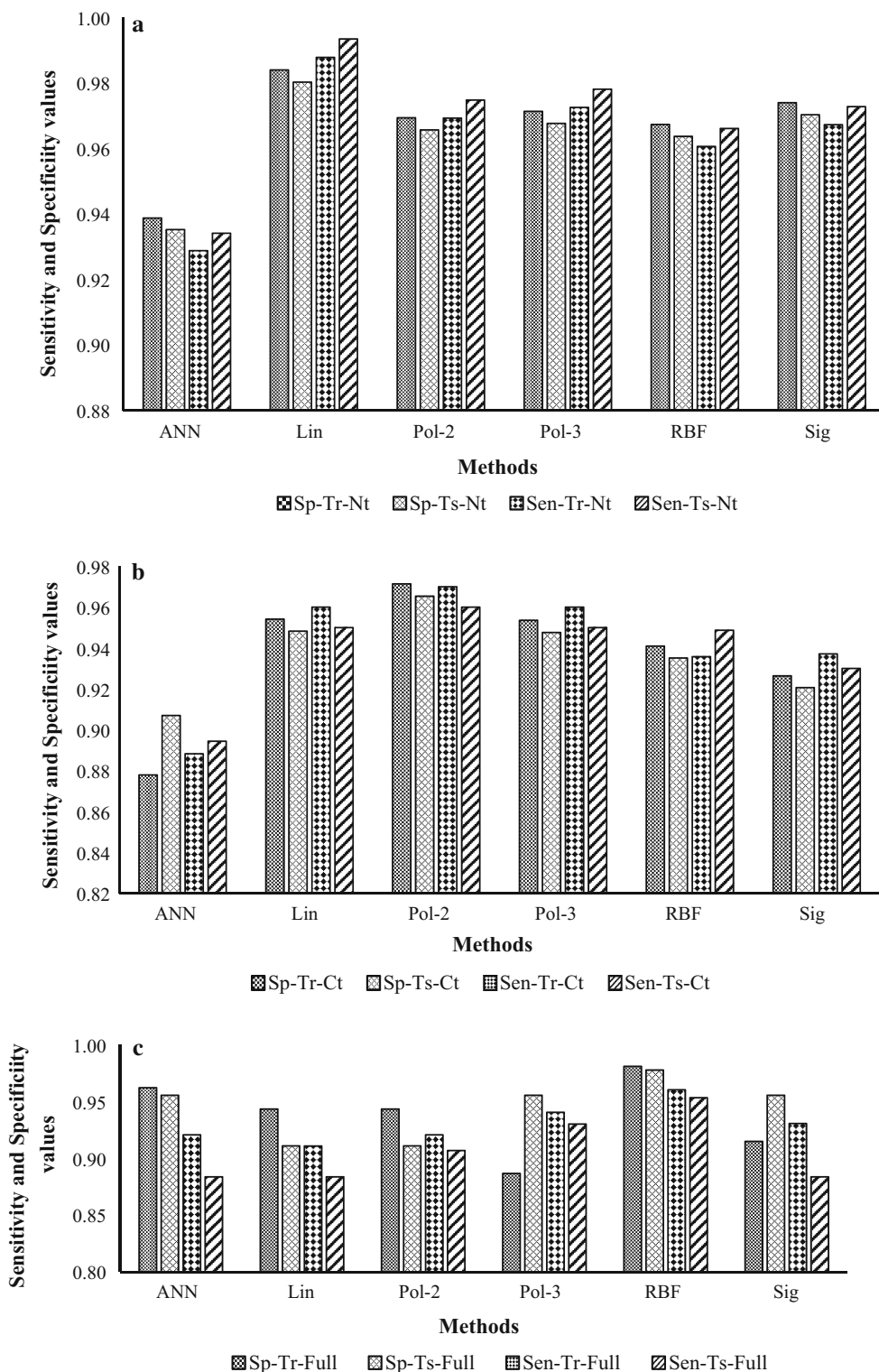
and can select either of the radio button captioned as C-terminus, N-terminus and Full Sequence. Tutorial page provides "User guide" which aids user to use this server in an effective manner.

Conclusions

ANN and SVM methodologies have successfully been used for prediction of species-specific antimicrobial peptide prediction in fishes for the first time. This is the world's first prediction server in fishes, which is freely accessible at http://webapp.cabgrid.res.in/fishamp/ for global research community. This methodology has an accuracy of 97 % for fish-specific AMP prediction of unknown peptides. This in silico approach can drastically reduce the time and cost of AMP discovery without compromising accuracy. This server can accelerate the discovery rate of lead AMP molecules having potential wider applications in diverse area like fish and human health as substitute of antibiotics, immunomodulator, antitumor, vaccine adjuvant and inactivator, and packaged food.

Availability and Requirements

The antimicrobial peptide prediction tool for aquaculture industries is freely accessible for research purposes for nonprofit and academic organizations at http://webapp.cabgrid.res.in/fishamp/.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Davies J, Davies D (2010) Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev 74(3):417–433
2. Cabello FC (2006) Heavy use of prophylactic antibiotics in aquaculture: a growing problem for human and animal health and for the environment. Environ Microbiol 8:1137–1144
3. Silva NC, Sarmento B, Pintado M (2013) The importance of antimicrobial peptides and their potential for therapeutic use in ophthalmology. Int J Antimicrob Agents 41(1):5–10
4. Bradshaw J (2003) Cationic antimicrobial peptides: issues for potential clinical use. BioDrugs 17(4):233–240
5. Rakers S, Niklasson L, Steinhagen D, Kruse C, Schauber J, Sundell K et al (2013) Antimicrobial peptides (AMPs) from fish epidermis: perspectives for investigative dermatology. J Invest Dermatol 133:1140–1149
6. Alderman DJ, Hastings TS (2003) Antibiotic use in aquaculture: development of antibiotic resistance-potential for consumer health risks. Int J Food Sci Technol 33:139–155
7. Pridgeon JW, Klesius PH (2012) Major bacterial diseases in aquaculture and their vaccine development. Cab Rev 7(48):1–16
8. FAO (2009) Fishstat Plus. Food and Agricultural Organisation of the United Nations, Rome
9. FAO (2012) Food and Agriculture Organization of the United Nations. Fisheries Department. The state of world fisheries and aquaculture. Food and Agriculture Organization of the United Nations, Rome
10. Lafferty KD, Harvell CD, Conrad JM, Friedman CS, Kent ML, Kuris AM et al (2015) Infectious diseases affect marine fisheries and aquaculture economics. Annu Rev Mar Sci 7:471–496
11. Cederlund A, Gudmundsson GH, Agerberth B (2011) Antimicrobial peptides important in innate immunity. FEBS J 278(20):3942–3951
12. Sarika IM, Rai A (2012) Biotic stress resistance in agriculture through antimicrobial peptides. Peptides 36:322–330
13. Diamond G, Beckloff N, Weinberg A, Kisich KO (2009) The roles of antimicrobial peptides in innate host defense. Curr Pharm Des 15(21):2377–2392
14. Rajanbabu V, Chen JY (2011) Applications of antimicrobial peptides from fish and perspectives for the future. Peptides 32(2):415–420
15. Beisswenger C, Bals R (2005) Functions of antimicrobial peptides in host defense and immunity. Curr Protein Pept Sci 6(3):255–264
16. Izadpanah A, Gallo RL (2005) Antimicrobial peptides. J Am Acad Dermatol 52:381–390
17. Gordon YJ, Romanowski EG, McDermott AM (2005) A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. Curr Eye Res 30(7):505–515
18. Austin B (2012) Infectious disease in aquaculture: prevention and control. Woodhead Pub. Ltd, Oxford
19. Karen EB, Dunman PM, McAleese F (2004) Global gene expression in Staphylococcus aureus biofilms. J Bacteriol 186(14):4665–4684
20. Snehlata SB, Raghava GPS (2007) Analysis and prediction of antibacterial peptides. BMC Bioinform 8:263
21. Thomas S, Karnik S, Barai RS, Jayaraman VK, Thomas SI (2010) CAMP: a useful resource for research on antimicrobial peptides. Nucleic Acids Res 38:D774–D780
22. Sarika IM, Arora V, Rai A, Kumar D (2015) Species specific approach to the development of web-based antimicrobial peptides prediction tool for cattle. Comput Electron Agric 111:55–61
23. Zhao X, Wu H, Lu H, Li G, Huang Q (2013) LAMP: a database linking antimicrobial peptides. PLoS ONE 8(6):e66557
24. Waghu FH, Gopi L, Barai RS, Ramteke P, Nizami B, Thomas SI (2014) CAMP: collection of sequences and structures of antimicrobial peptides. Nucleic Acids Res 42:D1154–D1158
25. Gueguen Y, Garnier J, Robert L, Lefranc MP, Mougenot I, de Lorgeril J et al (2006) PenBase, the shrimp antimicrobial peptide penaeidin database: sequence-based classification and recommended nomenclature. Dev Comp Immunol 30(3):283–288
26. Zamyatnin AA, Borchikov AS, Vladimirov MG, Voronina OL (2006) The EROP-Moscow oligopeptide database. Nucleic Acids Res 34:D261–D266
27. Wang G, Li X, Wang Z (2009) APD2: the updated antimicrobial peptide database and its application in peptide design. Nucleic Acids Res 37:D933–D937
28. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659
29. Kumar M, Verma R, Raghava GPS (2006) Prediction of mitochondrial proteins using support vector. J Biol Chem 281(9):5357–5363
30. Martin W, Mentel M (2010) The origin of mitochondria. Nat Educ 3(9):58
31. StatSoft, Inc. (2001) STATISTICA (Data Analysis Software System). Version 6.0. www.statsoft.com

32. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD et al (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, New York, pp 571–607

33. Joachims T (1999) Making large-Scale SVM learning practical. In: Schölkopf B, Burges C, Smola A (eds) Advances in Kernel methods—support vector learning. MIT-Press, Cambridge, pp 1–22

34. Haykin S (1994) Neural networks: a comprehensive foundation. Macmillan Publishing, New York

35. Cheng B, Titterington DM (1994) Neural networks: a review from a statistical perspective. Stat Sci 9(1):2–30

36. Vapnik V (2000) The nature of statistical learning theory. Springer Science & Business Media, Berlin

37. Efron B (1983) Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc 78:316–631

38. Bhasin M, Raghava GPS (2004) Prediction of CTL epitopes using QM, SVM and ANN techniques. Vaccine 22(23–24):3195–3204

39. Pathan FK, Venkata DA, Panguluri SK (2010) Recent patents on antimicrobial peptides. Recent Pat DNA Gene Seq 4(1):10–16

40. Kindrachuk J, Napper S (2010) Structure-activity relationships of multifunctional host defence peptides. Mini Rev Med Chem 10:596–614

41. Chen JY, Lin WJ, Lin TL (2009) A fish antimicrobial peptide, tilapia hepcidin TH2-3, shows potent antitumor activity against human fibrosarcoma cells. Peptides 30:1636–1642

42. Chen JY, Lin WJ, Wu JL, Her GM, Hui CF (2009) Epinecidin-1 peptide induces apoptosis which enhances antitumor effects in human leukemia U937 cells. Peptides 30:2365–2373

43. Lin WJ, Chien YL, Pan CY, Lin TL, Chen JY et al (2009) Epinecidin-1, an antimicrobial peptide from fish (Epinephelus coioides) which has an antitumor effect like lytic peptides in human fibrosarcoma cells. Peptides 30:283–290

44. Chiou PP, Khoo J, Bols NC, Douglas S, Chen TT (2006) Effects of linear cationic alpha-helical antimicrobial peptides on immune-relevant genes in trout macrophages. Dev Comp Immunol 30:797–806

45. Wang YD, Kung CW, Chi SC, Chen JY (2010) Inactivation of nervous necrosis virus infecting grouper (Epinephelus coioides) by epinecidin-1 and hepcidin 1-5 antimicrobial peptides, and downregulation of Mx2 and Mx3 gene expressions. Fish Shellfish Immunol 28:113–120

46. Fulmer PA, Lundin JG, Wynne JH (2010) Development of antimicrobial peptides (AMPs) for use in self-decontaminating coatings. ACS Appl Mater Interfaces 2:1266–1270